BUDAPESTI KÖZGAZDASÁGTUDOMÁNYI EGYETEM

Puskás Csaba, Szabó Imre, Tallos Péter

LINEÁRIS ALGEBRA

JEGYZET

BUDAPEST, 1997

A szerzők Lineáris Algebra, illetve Lineáris Algebra II. c. jegyzeteinek átdolgozott kiadása.

Előszó

Ez a jegyzet a Budapesti Közgazdaságtudományi Egyetem hallgatóinak második féléves lineáris algebrai tanulmányait szeretné segíteni. A jegyzet az "alapszintű" matematika oktatásban résztvevő hallgatók lineáris algebra tananyagát tartalmazza. A lineáris algebra eredetileg lineáris egyenletrendszerek megoldásával foglalkozott, ezért először csak a mátrixaritmetika és determinánselmélet tartozott tárgyához. Döntő hatással volt fejlődésére, az a felismerés, hogy a mindennapi értelemben vett tér geometriájának általánosításaként kapott vektorterek elmélete a lineáris egyenletrendszerek problémakörét más megvilágításba helyezi. Ebben a jegyzetben a vektorterek elméletének elemeit tárgyaljuk, és a mátrixaritmetika ennek a célnak a szolgálatába van állítva. Úgy érezzük, hogy így könnyebben megmutatkozik mind a tételek mélyebb értelme és az azok közötti kapcsolat. Ez a felépítés lehetővé teszi, hogy az itt nyert eredményeket mind a matematikán belül, mind más tudományterületeken is alkalmazzák.

Szóljunk néhány szót a jegyzet szerkezetéről és jelölésmódjáról. Először a későbbiekben sokat használt mátrixaritmetika elemeit gyűjtöttük össze, majd bemutatjuk az absztrakt vektortereket, és legfontosabb tulajdonságaikkal jellemezzük azokat. Ezután rátérünk a lineáris leképezések és transzformációk tárgyalására. Ezek reprezentációja teremti meg a kapcsolatot a mátrixaritmetikával. Ezt követően már eleget tudunk ahhoz, hogy a lineáris egyenletrendszerek megoldását elegánsan kezelhessük. Ezután az euklideszi terek tárgyalása következik, majd a lineáris transzformációk sajátértékeinek és sajátvektorainak meghatározására adunk módszert. Végül a többváltozós függvények lokális szélsőértékeinek meghatározásakor elengedhetetlen kvadratikus alakok és azok definitségének vizsgálata következik. Az utolsó hatodik fejezet a többváltozós függvénytan elemeinek lineáris algebrai eszközökkel való tárgyalását tartalmazza.

A bevezetett fogalmak többségét számozott definíciókban adjuk meg, néha azonban a gördülékenység érdekében csak dőltbetűs szedéssel hívjuk fel rájuk a figyelmet. A tételek és állítások tripla számozása megmutatja, hogy mely fejezet, melyik pontjának hányadik tételéről vagy állításáról van szó.

A Faktorterek című szakasz $**$ jelzéssel van ellátva, ami azt jelzi, hogy ismerete nélkül is érthető a további anyag, de úgy gondoltuk, hogy elolvasása hozzájárulhat a vektorterek elméletének jobb megértéséhez.

A jegyzet első öt fejezetét Puskás Csaba, míg az utosó hatodik fejezetet Szabó Imre és Tallos Péter írták.

Itt hívjuk fel a figyelmet arra, hogy az egy-egy pontot lezáró feladatok és gyakorlatok nem pótolhatják a feladatgyűjteményt. Ebből a szempontból ez a jegyzet meglehetősen hiányos.

Tudjuk, hogy minden igyekezetünk ellenére még mindig maradtak hibák, elírások, bár a kollégáink nagyon sokat felfedeztek és azokat természetesen kijavítottuk.

A jegyzetet szedési munkái a TEX kiadványszerkesztő szoftver LATEX változatával, az ábrák pedig a PICTEX szoftverrel készültek.

Budapest, 1997. február 6.

Puskás Csaba, Szabó Imre, Tallos Péter

Tartalomjegyzék

1. Fejezet

Vektorterek és elemi tulajdonságaik

Ebben a fejezetben először a mátrix fogalmával, a mátrixokkal végzett műveletekkel, és azok tulajdonságaival ismerkedünk meg. Ezután a köznapi értelemben vett sík vektorain értelmezett összeadás és vektorok valós számokkal való szorzásának tulajdonságait vizsgáljuk. A szerzett tapasztalatok birtokában definiáljuk az absztrakt vektortér fogalmát. Már itt szeretnénk arra biztatni az olvasót, hogy a későbbiekben bevezetett fogalmak és állítások pontos megértése érdekében mindig vizsgálja meg, hogy a szóbanforgó fogalomnak, illetve állításnak mi a geometriai jelentése a síkon és a térben. A geometriai modell, bár nem helyettesíti a bizonyítást, de segíti a megértést.

A közgazdász hallgatók számára a valós koordinátaterek ismerete a legfontosabb, mégis, ahol a minden vektortérre jellemző tulajdonságokat tárgyaljuk, a tisztább fogalomalkotás érdekében nem használjuk a vektorok koordinátáit.

1.1 Mátrixaritmetika

A lineáris algebrai problémák numerikus megoldása az esetek nagy többségében mátrixaritmetikai operációk elvégzését kívánja a feladat megoldójától. Ezért ebben az első pontban a mátrix fogalmával és az ezekkel végzett műveletekkel fogunk megismerkedni. Mátrixokkal már számtalanszor találkozott az olvasó, csak nem biztos, hogy azokat mátrixnak nevezték. Számítástechnikában például a mátrix név helyett a tömb név használatos, míg a mindennapi életben egyszerűen csak számtáblázatról beszélünk. A mátrix valójában nem más, mint egy téglalap alakú számtáblázat. A mátrix elemei tehát számok, és a számokkal végzett műveletekre fogjuk visszavezetni a mátrixokkal végrehajtandó műveleteket. Meg kell itt jegyezni, hogy a matematika különböző területein értelmeznek nemcsak számokból felépülő mátrixokat is, de a mi céljainknak tökéletesen megfelelnek a számmátrixok. Persze még meg kell azt is mondanunk, hogy milyen számok lesznek a vizsgálandó mátrixok elemei. Erre a kérdésre azt lehet válaszolni, hogy az esetek többségében valós számok, és minden esetben olyan számok, amelyeken ugyanolyan tulajdonságú műveletek végezhetők, mint a valós számok halmazán, azaz valamilyen számtestest elemeiből építjük majd fel a mátrixokat. Az algebrai struktúrákkal való ismerkedés során, már találkoztunk a test fogalmával, most mégis újra megadjuk annak definícióját.

1.1.1 Definíció. *Az $\mathbb{F}$ kétműveletes algebrai struktúrát testnek nevezzük, ha a műveletek — melyeket összeadásnak, illetve szorzásnak hívunk és jelölésükre a $+$, illetve $\cdot$ jeleket használjuk – - eleget tesznek az alább felsorolt tulajdonságoknak. Bármely $\alpha, \beta, \gamma \in \mathbb{F}$-re:*

1. $\alpha + \beta = \beta + \alpha$ *(az összeadás kommutatív),*

2. $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ *(az összeadás asszociatív),*

3. $(\exists\, 0 \in \mathbb{F}) : 0 + \alpha = \alpha + 0 = \alpha$ *(van zéróelem $\mathbb{F}$-ben),*

4. $(\forall\, \alpha \in \mathbb{F}) : (\exists\, (-\alpha) \in \mathbb{F}) : \alpha + (-\alpha) = (-\alpha) + \alpha = 0$ *(minden elemnek van negatívja),*

a. $\alpha \cdot \beta = \beta \cdot \alpha$ *(a szorzás kommutatív),*

b. $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$ *(a szorzás asszociatív),*

c. $(\exists\, 1 \in \mathbb{F}) : 1 \cdot \alpha = \alpha \cdot 1 = \alpha$ *(van egységelem $\mathbb{F}$-ben),*

d. $(\forall\, \alpha(\neq 0) \in \mathbb{F}) : (\exists\, \alpha^{-1} \in \mathbb{F}) : \alpha \cdot \alpha^{-1} = \alpha^{-1} \cdot \alpha = 1$ *(minden nemzéró elemnek van reciproka),*

A. $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$ *és* $(\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$ *(a szorzás az összeadásra vonatkozóan disztributív).*

A jól ismert számhalmazaink közül a racionális számok, a valós számok és a komplex számok alkotnak testet a szokásos összeadás és szorzás műveletekkel. Vannak azonban véges testek is, például véges testet kapunk, ha tetszőleges p prímszám esetén a $\{0, 1, \ldots, (p-1)\}$ halmazon úgy definiáljuk két elem összegét/szorzatát, hogy képezzük előbb a közönséges összegüket/szorzatukat, majd a kapott eredmény p-vel való osztása utáni maradékát vesszük. $p = 5$ esetén az alábbi táblázatokban bemutatjuk az összeadás és szorzás fenti értelmezését:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

és

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Az olvasó könnyen ellenőrizheti, hogy a test definíciójában megkövetelt tulajdonságok mind egyike teljesül. A jólismert számtestek és a most mutatott véges testek egy lé nyeges tulajdonságban különböznek: nevezetesen, ha $\alpha \neq 0$ a test egy tetszőleges eleme és n egy pozitív egész, akkor az $n \cdot \alpha$ — amin olyan n tagú összeget kell értenünk, melynek minden tagja α — a racionális, valós, vagy komplex számok testében sohasem nulla, míg a véges testekben lehet nulla. Azt a legkisebb pozitív n egész számot, amelyre $n \cdot \alpha = 0$ teljesül a test tetszőleges $\alpha \neq 0$ elemére, a test *karakterisztikájának* nevezzük, és ha ilyen pozitív egész szám nem létezik, akkor a testet 0-*karakterisztikájúnak* mondjuk. A $\{0,1,2,3,4\}$ halmazon a fenti műveleti táblákkal értelmezett test karakterisztikája öt, például az $1+1+1+1+1=0$, de nem nehéz ellenőrizni, hogy a $2,3,4$ elemekre is teljesül, hogy ötszörösük nulla az adott testben.

Tévedés lenne azt gondolni, hogy csak a racionális, valós, vagy a komplex számok teste 0-karakterisz tikájú, például az $a+b\sqrt{2}$ alakú számok halmaza, ahol a és b racionálisak — a valós számhalmaznak valódi, de a racionális számok halmazánál bővebb részhalmaza — a szokásos összeadás és szorzás műveletekkel 0-karakterisztikájú test. Persze lineáris algebrai tanulmányaink során legtöbbször számon valós számot és néhányszor komplex számot fogunk érteni, de az általánosíthatóságot azzal jelezzük, hogy a szóbanforgó számtestet $\mathbb{F}$-fel fogjuk jelölni..

1.1.2 Definíció. *Legyen $\mathbb{F}$ egy tetszőleges számtest és jelölje I az $\{1,\dots,m\}$ és J pedig az $\{1,\dots,n\}$ természetes számok halmazát. Az $\mathbf{A}: I\times J\to\mathbb{F}$ alakú függvényeket $m\times n$ tipusú $\mathbb{F}$ test feletti mátrixoknak nevezzük.*

Azonnal megjegyezzük, hogy amint azt a sorozatok esetében is tettük, amelyek a természetes számok halmazán értelmezett függvények, de a függvényértékek rendszerével reprezentáltuk azokat, a mátrixokat is a függvényértékek rendszerével reprezentálhatjuk, amelyeket m sorból és n oszlopból álló téglalap alakú táblázatba rendezünk:

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

Ennek megfelelően azon, hogy az $\mathbf{A}$ mátrix tipusa $m\times n$ azt kell érteni, hogy m sora és n oszlopa van. A mátrixok jelölésére vastagon szedett nagy betűket fogunk használni, de sokszor egy úgynevezett általános elemével is hívatkozhatunk a mátrixra, amelyet szögletes zárójelek közé írunk. Például, a fenti $\mathbf{A}$ mátrixra az $[\alpha_{ij}]$ jellel is utalhatunk. Itt α_{ij} persze az i-dik sorban és j-dik oszlopban lévő $\mathbb{F}$-beli szám. Mielőtt a mátrixokkal műveleteket hajtanánk végre, meg kell mondnunk, hogy mikor tekintünk két mátrixot egyenlőnek.

Az $\mathbf{A}$ és $\mathbf{B}$ mátrixok egyenlők, ha tipusaik azonos és az azonos pozícióban lévő elemeik egyenlők. Formalizálva a fentieket, az

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

és

$$\mathbf{B}=\begin{bmatrix}\beta_{11} & \beta_{12} & \dots & \beta_{1\ell}\\ \beta_{21} & \beta_{22} & \dots & \beta_{2\ell}\\ \vdots & \vdots & \ddots & \vdots\\ \beta_{k1} & \beta_{k2} & \dots & \beta_{k\ell}\end{bmatrix}$$

mátrixokra

$$\mathbf{A}=\mathbf{B}\Leftrightarrow m=k,\quad n=\ell \text{ és } \forall i(=1,\dots,m);\ j(=1,\dots,n):\quad \alpha_{ij}=\beta_{ij}\,.$$

Ezek után már definiálhatjuk mátrixok összeadását a következőképpen: két azonos tipusú $\mathbf{A} = [\alpha_{ij}]$ és $\mathbf{B} = [\beta_{ij}]$ mátrix összeadását az

$$[\alpha_{ij}] + [\beta_{ij}] \stackrel{\text{def}}{=} [\alpha_{ij} + \beta_{ij}]$$

egyenlőség értelmezi. Tehát két mátrix összege csak akkor van értelmezve, ha tipusaik megegyeznek, és az összegmátrix i-dik sorának j-dik eleme az összeadandó mátrixok i-dik sorában és j-dik oszlopában lévő elemek összege.

Például, az

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 6 \\ -1 & 9 & -2 \end{bmatrix} \quad \text{és} \quad \mathbf{B} = \begin{bmatrix} 1 & -2 & 3 \\ 4 & -5 & 6 \end{bmatrix}$$

mátrixok összege az

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 3 & -2 & 9 \\ 3 & 4 & 4 \end{bmatrix}.$$

Értelmezzük mátrixoknak számmal való szorzatát is az alábbi definiáló egyenlőséggel:

$$\lambda\mathbf{A} = \lambda[\alpha_{ij}] \stackrel{\text{def}}{=} [\lambda\alpha_{ij}]\,.$$

Szavakkal megfogalmazva ugyanezt egy $\mathbf{A}$ mátrix λ számmal való szorzatát megkapjuk, ha az $\mathbf{A}$ mátrix minden elemét megszorozzuk a λ számmal.

Például az

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 6 \\ -1 & 9 & -2 \end{bmatrix}$$

mátrix 3-szorosa:

$$3 \cdot \mathbf{A} = \begin{bmatrix} 6 & 0 & 18 \\ -3 & 27 & -6 \end{bmatrix}.$$

Az összeadásra és a számmal való szorzásra támaszkodva képezhetjük azonos tipusú mátrixok úgynevezett lineáris kombinációját a következőképpen:

1.1.3 Definíció. *Az $\mathbf{A}_1, \ldots, \mathbf{A}_k$ azonos tipusú mátrixok $\lambda_1, \ldots, \lambda_k$ skalárokkal képzett lineáris kombinációján a*

$$\lambda_1\mathbf{A}_1 + \cdots + \lambda_k\mathbf{A}_k$$

mátrixot értjük.

A mátrixok szorzásának definiálásakor fontos szerepet játszanak azok a mátrixok, amelyeknek csak egyetlen sora vagy csak egyetlen oszlopa van. Az $1 \times n$ tipusú mátrixokat sorvektoroknak, míg az $m \times 1$ tipusúakat oszlopvektoroknak is fogjuk nevezni és jelölésükre vastagon szedett kisbetűket fogunk használni. Időnként, ha külön is hangsúlyozni kivánjuk, hogy sorvektorról van szó, azt a $*$ felső indexszel jelöljük.

Értelmezzük két azonos elemszámú vektor belső szorzatát, vagy másnéven skaláris szorzatát az alábbi módon:

1.1.4 Definíció. *Az $\mathbf{a} = [\alpha_1, \ldots, \alpha_p]$ és a $\mathbf{b} = [\beta_1, \ldots, \beta_p]$ vektorok belső szorzatán az*

$$\mathbf{a} \cdot \mathbf{b} \stackrel{\text{def}}{=} \sum_{i=1}^{p} \alpha_i\beta_i$$

számot értjük.

A definícióból világos, hogy csak azonos elemszámú vektorok belső szorzata létezik és a gyakran használt skaláris szorzat elnevezés onnan ered, hogy a szorzat eredménye egy szám, amelynek szinonímája a skalár.

Két $\mathbf{A}$ és $\mathbf{B}$ mátrix $\mathbf{A} \cdot \mathbf{B}$ szorzatát csak abban az esetben értelmezzük, ha képezhető az $\mathbf{A}$ mátrix sorainak és a $\mathbf{B}$ mátrix oszlopainak a belső szorzata, vagy ami ugyanaz, ha az $\mathbf{A}$ oszlopainak a száma egyenlő a $\mathbf{B}$ sorainak a számával.

1.1.5 Definíció. *Jelölje az* $m \times n$ *tipusú* $\mathbf{A}$ *mátrix* i*-dik sorát* $\mathbf{a}_i$ $(i = 1, \ldots m)$ *és az* $n \times p$ *tipusú* $\mathbf{B}$ *mátrix* j*-dik oszlopát* $\mathbf{b}_j$ $(j = 1, \ldots, p)$. *Ekkor az* $\mathbf{A} \cdot \mathbf{B}$ *szorzatmátrix értelmezve van, tipusa* $m \times p$ *és a szorzat* i*-dik sorában a* j*-dik elem* $\mathbf{a}_i \cdot \mathbf{b}_j$, *minden* $i (= 1, \ldots m)$ *és* $j (= 1, \ldots, p)$ *indexpárra.*

Részletesen kiírva a definícióban leírtakat:

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_1 \cdot \mathbf{b}_p \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_2 \cdot \mathbf{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m \cdot \mathbf{b}_1 & \mathbf{a}_m \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_m \cdot \mathbf{b}_p \end{bmatrix}.$$

Amint az a fenti definícióból is kiderül, a mátrixok szorzásánál a tényezők sorrendje lényeges, hiszen a baloldali tényező sorvektorainak és a jobboldali tényező oszlopvektorainak kell képezzük a belső szorzatait, hogy megkapjuk a szorzatmátrix elemeit. Ezt hangsúlyozandó célszerű a fenti szorzatmátrixot

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} \mathbf{a}_1^* \cdot \mathbf{b}_1 & \mathbf{a}_1^* \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_1^* \cdot \mathbf{b}_p \\ \mathbf{a}_2^* \cdot \mathbf{b}_1 & \mathbf{a}_2^* \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_2^* \cdot \mathbf{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m^* \cdot \mathbf{b}_1 & \mathbf{a}_m^* \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_m^* \cdot \mathbf{b}_p \end{bmatrix}.$$

alakban írni.

Ezt a jelölést az is indokolja, hogy amíg egy sorvektornak (sormátrixnak) és egy oszlopvektornak (oszlopmátrixnak) a szorzata csak akkor létezik, ha azonos elemszámúak, addig egy oszlopvektornak és egy sorvektornak a szorzata a fenti szorzásdefiníció szerint mindig létezik. Valóban, ha

$$\mathbf{a} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} \quad \text{és} \quad \mathbf{b}^* = [\beta_1, \beta_2, \ldots, \beta_n],$$

akkor szorzatuk

$$\mathbf{a} \cdot \mathbf{b}^* = \begin{bmatrix} \alpha_1\beta_1 & \alpha_1\beta_2 & \ldots & \alpha_1\beta_n \\ \alpha_2\beta_1 & \alpha_2\beta_2 & \ldots & \alpha_2\beta_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_m\beta_1 & \alpha_m\beta_2 & \ldots & \alpha_m\beta_n \end{bmatrix}$$

egy $m \times n$ tipusú mátrix, amelyet az $\mathbf{a}$ és $\mathbf{b}^*$ vektorok *diadikus szorzatának* nevezünk.

1.1.1 A mátrixműveletek tulajdonságai

Az előzőekben definiáltunk néhány mátrixokkal végrehajtható műveletet. Ebben az alpontban megvizsgáljuk, hogy a számokkal végzett műveleteink tulajdonságai közül melyek maradnak érvényesek és melyek vesztik érvényüket a mátrixműveletek esetében.

A mátrixok összeadásának tulajdonságai

1. Az $\mathbf{A} = [\alpha_{ij}]$ és $\mathbf{B} = [\beta_{ij}]$ azonos tipusú mátrixok összege

$$[\alpha_{ij}] + [\beta_{ij}] = [\alpha_{ij} + \beta_{ij}] = [\beta_{ij} + \alpha_{ij}] = [\beta_{ij}] + [\alpha_{ij}]$$

független a tényezők sorendjétől, hiszen az összeadás elemeik összadására van visszavezetve és az elemeik egy test elemei, amelyben az összeadás kommutatív. Ebből következik, hogy a mátrixok összeadása is kommutatív művelet.
2. Teljesen hasonlóan kapjuk, hogy a mátrixok összeadása asszociatív is.
3. Amint a számok összeadásánál, létezik egy neutrális (reprodukáló) elem a nulla, ugyanúgy tetszőleges $m \times n$ tipusú mátrixok esetén az úgynevezett nullmátrix (jelölése: $\mathbf{O} = [0]$), amelynek minden eleme nulla, neutrális elem lesz a mátrixok összeadására nézve.
4. Bármely $\mathbf{A} = [\alpha_{ij}]$ mátrixnak létezik az ellentettje is a $-\mathbf{A} = [-\alpha_{ij}]$ mátrix, amely eleget tesz az $\mathbf{A} + (-\mathbf{A}) = \mathbf{O}$ egyenlőségnek.

Összefoglalva a fentiekben felsorolt tulajdonságokat azt mondhatjuk, hogy az $m \times n$ tipusú $\mathbb{F}$ test feletti mátrixok az összeadás művelettel kommutatív csoportot alkotnak.

A mátrixok skalárral való szorzásának tulajdonságai

1. Ha két skalár összegével $(\lambda + \kappa)$-val szorozzuk az $\mathbf{A} = [\alpha_{ij}]$ mátrixot, akkor a

$$(\lambda + \kappa)[\alpha_{ij}] = [(\lambda + \kappa)\alpha_{ij}] = [\lambda\alpha_{ij} + \kappa\alpha_{ij}] = [\lambda\alpha_{ij}] + [\kappa\alpha_{ij}]$$

egyenlőségek mutatják, hogy ugyanazt kapjuk, mint az $\mathbf{A}$ λ-szorosának és az $\mathbf{A}$ κ-szorosának az összegét, tehát fennáll az

$$(\lambda + \kappa)\mathbf{A} = \lambda\mathbf{A} + \kappa\mathbf{A}$$

egyenlőség.
2. Egy λ skalárral szorozva két mátrix összegét, a

$$\lambda([\alpha_{ij}] + [\beta_{ij}]) = \lambda[\alpha_{ij} + \beta_{ij}] = [\lambda(\alpha_{ij} + \beta_{ij})] = [\lambda\alpha_{ij} + \lambda\beta_{ij}] = [\lambda\alpha_{ij}] + [\lambda\beta_{ij}]$$

számolás szerint ugynazt kapjuk mintha előbb az egyik, majd a másik mátrixot szoroznánk a skalárral, majd pedig az így kapott mátrixokat összeadnánk. Ezt a tulajdonságot tehát így formalizálhatjuk:

$$\lambda(\mathbf{A} + \mathbf{B}) = \lambda\mathbf{A} + \lambda\mathbf{B}\,.$$

3. Tekintsük most egy mátrixnak skalárok szorzatával való szorzatát. A

$$(\lambda\kappa)[\alpha_{ij}] = [(\lambda\kappa)\alpha_{ij}] = [\lambda(\kappa\alpha_{ij})] = \lambda[\kappa\alpha_{ij}] = \lambda(\kappa[\alpha_{ij}])$$

egyenlőségsorozat mutatja, hogy előbb az egyik skalárral szorozva kapunk egy mátrixot, majd ezt kell szorozni a másik skalárral. Tekintve, hogy testbeli elemek szorzása kommutatív, ezt a tulajdonságot így formalizálhatjuk:

$$(\lambda\kappa)\mathbf{A} = \lambda(\kappa\mathbf{A}) = \kappa(\lambda\mathbf{A})\,.$$

4. Ha a test 1-gyel jelölt egységelemével szorozzuk a mátrixot, akkor annak minden eleme változatlan marad, azaz

$$1\mathbf{A} = \mathbf{A}\,.$$

A mátrixok szorzásának tulajdonságai

1. Amint az már a definícióból is azonnal következik, a mátrixok szorzásművelete nem kommutatív, sőt az általános esetben az $m \times n$ tipusú $\mathbf{A}$ és az $n \times p$ tipusú $\mathbf{B}$ mátrixok $\mathbf{A} \cdot \mathbf{B}$ szorzata létezik, ugyanakkor $m \neq p$ esetén a $\mathbf{B} \cdot \mathbf{A}$ szorzat mégcsak nem is létezik. Nem kommutatív a szorzás akkor sem, ha a tényezők úgynevezett kvadratikus (négyzetes) mátrixok, amelyekben a sorok és oszlopok száma megegyezik, amint azt az alábbi egyszerű példa mutatja. Legyen

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{és} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix},$$

Ekkor az

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \quad \text{míg a} \quad \mathbf{B} \cdot \mathbf{A} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}.$$

2. Talán meglepő ezek után, hogy a mátrixok szorzása is asszociatív. Ezt bizonyítandó legyenek $\mathbf{A}$ $m \times n$ tipusú, $\mathbf{B}$ $n \times p$ tipusú és $\mathbf{C}$ pedig $p \times q$ tipusú mátrixok:

$$\mathbf{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{np} \end{bmatrix}$$

és

$$\mathbf{C} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1q} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{p1} & \gamma_{p2} & \dots & \gamma_{pq} \end{bmatrix},$$

Ekkor az $(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C}$ mátrix i-edik sorának j-edik eleme, amint azt a

$$\sum_{s=1}^{p} \left(\sum_{t=1}^{n} \alpha_{it}\beta_{ts} \right) \gamma_{sj} = \sum_{s=1}^{p} \sum_{t=1}^{n} (\alpha_{it}\beta_{ts})\gamma_{sj} =$$

$$\sum_{s=1}^{p}\sum_{t=1}^{n}\alpha_{it}(\beta_{ts}\gamma_{sj}) = \sum_{t=1}^{n}\alpha_{it}\left(\sum_{s=1}^{p}\beta_{ts}\gamma_{sj}\right),$$

egyenlőségsorozat bal- és jobboldala mutatja megegyezik az $\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})$ mátrix i-edik sorának j-edik elemével tetszőleges $i(=1,\ldots,m)$ és $j(=1,\ldots,q)$ indexpárra és ez éppen a szorzás asszociatív voltát igazolja.
3. A mátrixok szorzása nem-kommutatív lévén, egyáltalán nem meglepő, hogy általában nincs olyan mátrix, amellyel akár balról, akár jobbról szorozva egy $m\times n$ $(m\neq n)$ tipusú mátrixot, azt változatlanul hagyja. Létezik viszont bármely mátrixhoz külön balodali és külön jobboldali egységelem. Tetszőleges rögzített n természetes számra jelölje $\mathbf{E}_n$ azt az $n\times n$-es mátrixot, amelynek ϵ_{ij} eleme 1, ha $i=j$ és 0, ha $i\neq j$, azaz

$$\mathbf{E}_n = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}.$$

Akkor, amint az könnyen ellenőrizhető,

$$\mathbf{E}_m\cdot\mathbf{A}=\mathbf{A} \quad \text{és} \quad \mathbf{A}\cdot\mathbf{E}_n=\mathbf{A},$$

tehát $\mathbf{E}_m$ baloldali– és $\mathbf{E}_n$ pedig jobboldali egységelem. Persze a négyzetes mátrixok esetében a baloldali– és jobboldali egységelem megegyezik.
4. A testet alkotó számok szorzásánál azt is megfigyelhettük, hogy minden nemnulla számnak van multiplikatív inverze, amellyel a számot szorozva az egységelemet kapjuk. A mátrixok szorzása általában nem invertálható, még akkor sem, ha csupán kavadratikus mátrixokra szorítkozunk. Ennek igazolásához további lineáris algebrai ismeretekre van szükség, ezért későbbre halasztjuk.

A mátrixok szorzásának és összeadásának kapcsolata

Az alábbiakban igazoljuk, hogy a mátrixok szorzása disztributív, azok összeadására nézve. Legyen

$$\mathbf{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \ldots & \beta_{np} \end{bmatrix}$$

és

$$\mathbf{C} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1p} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{np} \end{bmatrix}$$

tetszőleges mátrixok. Az $\mathbf{A}\cdot(\mathbf{B}+\mathbf{C})$ mátrix i-edik sorának j-edik eleme a

$$\sum_{t=1}^{n}\alpha_{it}(\beta_{tj}+\gamma_{tj}) = \sum_{t=1}^{n}(\alpha_{it}\beta_{tj}+\alpha_{it}\gamma_{tj}) = \sum_{t=1}^{n}\alpha_{it}\beta_{tj} + \sum_{t=1}^{n}\alpha_{it}\gamma_{tj}$$

egyenlőségsorozat bal– és jobboldala szerint megegyezik az $\mathbf{A}\cdot\mathbf{B}+\mathbf{A}\cdot\mathbf{C}$ mátrix i-edik sorának j-edik elemével tetszőleges $i(=1,\ldots,m)$ és $j(=1,\ldots,p)$ indexpár esetén, és éppen ezt kellett megmutatnunk.

1.2 Speciális mátrixok

Ebben az a pontban néhány nevezetes mátrixszal ismerkedünk meg, amelyekkel későbbi tanulmányaink során még találkozni fogunk.

Már az előzőekben szó volt a kvadratikus, vagy más néven négyzetes mátrixokról, amelyekben a sorok és oszlopok száma megegyezik. A kvadratikus mátrixok sorainak a számát a mátrix *rendjének* fogjuk nevezni.

A négyzetes mátrixok fődiagonálisán az azonos sor– és oszlopindexű elemeinek az összességét értjük. Egy kvadratikus mátrixot diagonális mátrixnak nevezünk, ha a fődiagonálisán kívüli elemei mind zérók. Ha $\mathbf{A}$ n-edrendű diagonális mátrix, akkor gyakran csupán fődiagonálisbeli elemeinek feltüntetésével, $<\alpha_{11},\alpha_{22},\ldots,\alpha_{nn}>$ módon jelöljük. Vegyük észre, hogy minden n-re az $\mathbf{E}_n$-nel jelölt n-edrendű egységmátrix is diagonális mátrix, amelyben minden fődiagonálisbeli elem 1.

Egy n-edrendü $\mathbf{A}=[\alpha_{ij}]$ mátrixot em szimmetrikusnak mondunk, ha invariáns sorainak és oszlopainak felcserélésére nézve, azaz formalizálva ugyanezt, ha $\alpha_{ij}=\alpha_{ji}$ minden $i,j(=1,\ldots,n)$ indexpár esetén.

Jelöljük $\mathbf{A}^*$-gal azt a mátrixot, amelyet úgy kapunk $\mathbf{A}$-ból, hogy annak sorait oszlopaival cseréljük fel, tehát, ha

$$\mathbf{A}=\begin{bmatrix}\alpha_{11}&\alpha_{12}&\ldots&\alpha_{1n}\\ \alpha_{21}&\alpha_{22}&\ldots&\alpha_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ \alpha_{m1}&\alpha_{m2}&\ldots&\alpha_{mn}\end{bmatrix}\quad\text{akkor}\quad\mathbf{A}^*=\begin{bmatrix}\alpha_{11}&\alpha_{21}&\ldots&\alpha_{m1}\\ \alpha_{12}&\alpha_{22}&\ldots&\alpha_{m2}\\ \vdots&\vdots&\ddots&\vdots\\ \alpha_{1n}&\alpha_{2n}&\ldots&\alpha_{mn}\end{bmatrix}.$$

Az így kapott $\mathbf{A}^*$ mátrixot az $\mathbf{A}$ *transzponáltjának* nevezzük.

A transzponált fogalmának az ismeretében azt mondhatjuk, hogy egy $\mathbf{A}$ mátrix pontosan akkor szimmetrikus, ha megegyezik a transzponáltjával, azaz $\mathbf{A}=\mathbf{A}^*$.

Ha az $\mathbf{A}$ mátrix elemei történetesen komplex számok, akkor az $\mathbf{A}^*$-gal jelölt mátrix jelentése kicsit módosul. Ekkor $\mathbf{A}^*$ azt a mátrixot jelöli, amelyet úgy kapunk $\mathbf{A}$-ból, hogy sorait és oszlopait felcseréljük és minden elemet komplex konjugáltjára cseréljük. Ebben az esetben az $\mathbf{A}^*$ mátrixot az $\mathbf{A}$ mátrix *adjungáltjának* hívjuk. Ha az $\mathbf{A}$ mátrix megegyezik az adjungáltjával, akkor *önadjungáltnak* nevezzük. Tekintettel arra, hogy a valós számok halmaza tekinthető úgy, mint a komplex számok részhalmaza, beszélhetünk valós elemű mátrixok adjungáltjairól is, persze egy valós elemű mátrix adjungáltja és transzponáltja ugyanaz. Például az önadjungált valós elemű mátrixok a szimmetrikus mátrixok. Ez az oka annak, hogy jelölésben nem teszünk különbséget egy mátrix adjungáltja és transzponáltja között. Mint azt látni fogjuk, komplex vektorterekben ugyis csak az adjungálásnak van igazi szerepe, a transzponálásnak nincs.

Mielőtt a permutáló mátrix fogalmával megismerkednénk, fel kell hívjuk az olvasó figyelmét a mátrixok szorzásával kapcsolatos néhány érdekességre. Ha az

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

mátrixot megszorozzuk egy n elemű

$$\mathbf{x} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$$

oszlopvektorral, akkor az

$$\mathbf{A}\cdot\mathbf{x} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{bmatrix} \cdot \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} \xi_1\alpha_{11} + \xi_2\alpha_{12} + \dots + \xi_n\alpha_{1n} \\ \xi_1\alpha_{21} + \xi_2\alpha_{22} + \dots + \xi_n\alpha_{2n} \\ \vdots + \vdots + \ddots + \vdots \\ \xi_1\alpha_{m1} + \xi_2\alpha_{m2} + \dots + \xi_n\alpha_{mn} \end{bmatrix} =$$

$$\xi_1\begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{bmatrix} + \xi_2\begin{bmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{bmatrix} + \dots + \xi_n\begin{bmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{bmatrix}$$

oszlopvektort kapjuk, tehát a szorzat az $\mathbf{A}$ mátrix oszlopainak az $\mathbf{x}$ vektor komponenseivel alkotott lineáris kombinációja.

Jelöljük $\mathbf{e}_i$-vel $i(= 1, \dots, n)$ azt az n elemű oszlopvektort, amelynek i-edik komponense 1, míg minden más komponense 0, amelyet a továbbiakban i-edik egységvektornak fogunk nevezni. A fentiek alapján állíthatjuk, hogy ha egy $m \times n$ tipusú $\mathbf{A}$ mátrixot az i-edik egységvektorral szorozzuk jobbról, akkor eredményül az $\mathbf{A}$ mátrix i-edik oszlopát kapjuk.

Teljesen hasonló egyszerű számolással mutathatjuk meg, hogy ha az $\mathbf{y}^* = [\eta_1, \eta_2, \dots, \eta_m]$ sorvektorral szorozzuk az $\mathbf{A}$ mátrixot balról, akkor a szorzat a

$$\eta_1\mathbf{a}_1^* + \eta_2\mathbf{a}_2^* + \dots + \eta_m\mathbf{a}_m^*$$

sorvektor, ahol $\mathbf{a}_i^*$ $i(= 1, \dots, m)$ az $\mathbf{A}$ mátrix i-edik sorát jelöli. Tehát egy sorvektor és egy mátrix szorzata a mátrix sorainak a sorvektor komponenseivel, mint skaláregyütthatókkal képzett lineáris kombinációja. Ennek megfelelően, ha $\mathbf{e}_j^*$ m elemű j-edik egységvektor, akkor az $\mathbf{e}_j^* \cdot \mathbf{A}$ szorzat az $\mathbf{A}$ mátrix j-edik sora.

Tekintsük most két mátrix az

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^* \\ \mathbf{a}_2^* \\ \vdots \\ \mathbf{a}_m^* \end{bmatrix} \quad \text{és a } \mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p] \quad \mathbf{A}\cdot\mathbf{B} = \begin{bmatrix} \mathbf{a}_1^*\cdot\mathbf{b}_1 & \mathbf{a}_1^*\cdot\mathbf{b}_2 & \dots & \mathbf{a}_1^*\cdot\mathbf{b}_p \\ \mathbf{a}_2^*\cdot\mathbf{b}_1 & \mathbf{a}_2^*\cdot\mathbf{b}_2 & \dots & \mathbf{a}_2^*\cdot\mathbf{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m^*\cdot\mathbf{b}_1 & \mathbf{a}_m^*\cdot\mathbf{b}_2 & \dots & \mathbf{a}_m^*\cdot\mathbf{b}_p \end{bmatrix}$$

szorzatát. Vegyük észre, hogy a szorzat oszlopait, illetve sorait

$$[\mathbf{A}\mathbf{b}_1, \mathbf{A}\mathbf{b}_2 \ldots, \mathbf{A}\mathbf{b}_p] = \begin{bmatrix} \mathbf{a}_1^*\mathbf{B} \\ \mathbf{a}_2^*\mathbf{B} \\ \vdots \\ \mathbf{a}_m^*\mathbf{B} \end{bmatrix}$$

alakban is írhatjuk, tehát a szorzat minden oszlopa a baloldali mátrix oszlopainak lineáris kombinációja, illetve a szorzat minden sora a jobboldali mátrix sorainak lineáris kombinációja.

Ha egy n-edrendű kvadratikus mátrix megkapható az n-edrendű $\mathbf{E}_n$ egységmátrix oszlopainak (vagy sorainak) átrendezésével akkor *permutáló* mátrixnak nevezzük.

Az elnevezés a fenti észrevételek alapján érthető, hiszen, ha az $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ mátrixot megszorozzuk a $\mathbf{P} = [\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_n}]$ permutáló mátrixszal (itt az $i_1, i_2, \ldots, i_n$ számok az $1, 2, \ldots, n$ számok valamely permutációja), akkor az eredmény

$$\mathbf{A} \cdot \mathbf{P} = [\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_n}]\,,$$

tehát olyan mátrix, amelynek oszlopai az $\mathbf{A}$ mátrix oszlopainak átrendezésével (permutációjával) kapható. Hasonlóan, egy $\mathbf{A}$ mátrixot balról szorozva egy permutáló mátrixszal, az az $\mathbf{A}$ mátrix sorait permutálja.

A következő fogalom megfogalmazása előtt megjegyezzük, hogy a mátrixok szorzásának értelmezése alapján bevezethetjük kvadratikus mátrixok hatványozását az alábbi rekurzív definícióval: ha $\mathbf{A}$ n-edrendű mátrix, akkor $\mathbf{A}^0 \stackrel{\text{def}}{=} \mathbf{E}_n$ és bármely $k \geq 1$-re legyen $\mathbf{A}^k \stackrel{\text{def}}{=} \mathbf{A} \cdot \mathbf{A}^{k-1}$.

A valós számok viselkedésétől eltérően egy nemnulla mátrix valamely hatványa, lehet nullmátrix, amint azt az alábbi példa mutatja. Ha

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \text{akkor} \quad \mathbf{A}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{O}\,.$$

Egy n-edrendű $\mathbf{A}$ mátrixot k-adfokban *nilpotensnek* nevezünk, ha $\mathbf{A}^{k-1} \neq \mathbf{O}$ de $\mathbf{A}^k = \mathbf{O}$.

A következő elnevezés már arra utal, hogy a mátrixok szoros kapcsolatban vannak bizonyos leképezésekkel. Egy kvadratikus $\mathbf{A}$ mátrixot vetítő vagy más néven *projektív* mátrixnak nevezünk, ha eleget tesz az $\mathbf{A}^2 = \mathbf{A}$ feltételnek.

Ezzel a bevezetővel egyelőre felfüggesztjük a mátrixokkal kapcsolatos ismereteink gyarapítását, azzal az igérettel, hogy lineáris algebrai tanulmányaink során lépten nyomon találkozni fogunk azokkal és lesz alkalmunk további tanulmányozásukra.

1.3 Vektorok a síkon

A lineáris algebra, vagy kifejezőbb nevén a vektorterek elmélete a középiskolában tanult vektoralgebra, illetve koordinátageometria általánosítása. Ezért ebben a bevezető szakaszban a középiskolában tanult, síkbeli vektorok algebrájának összefoglalását találja az olvasó, hangsúlyozva azokat a tulajdonságokat, amelyek a későbbiekben definiált absztrakt vektorterek értelmezésénél elengedhetetlenek.

A vektorok bevezetését az az észrevétel tette szükségessé, hogy bizonyos mennyiségek matematikai jellemzésére a számok nem elegendőek, például egy mozgó objektum viselkedésének leírásakor nemcsak az objektum sebességének nagysága, de mozgásának iránya is fontos jellemző. Hasonlóan, egy testre ható erő okozta elmozdulás nemcsak az erő nagyságának, hanem irányának is függvénye.

Az ilyen és hasonló mennyiségek jellemzésére használtuk a vektorokat, amelyeket irányított szakaszok reprezentáltak. Tekintettel arra, hogy például egy objektumra egyszerre több erő is hathat és azok összhatása dönti el az objektum mozgását, szükséges, hogy ez az összhatás kiszámítható legyen az összetevők ismeretében. Ezért célszerű a vektorok összeadását értelmezni. Persze az összeadást úgy kívántuk definiálni, hogy az egyes erőhatásokat reprezentáló vektorok összege alkalmas legyen az úgynevezett, eredő erő reprezentálására. Sok esetben, például gyorsabb mozgás elérése érdekében meg kell "sokszorozni" egy objektumra ható erőt. Ez a vektorokkal való reprezentáció nyelvén azt jelenti, hogy egy vektornak számmal való szorzatát is kellett definiálnunk. Ha az objektumra ható erőket vektorokkal kívánjuk reprezentálni, akkor tekintve, hogy az erő független az objektum térbeli helyétől, célszerő két vektort egyenlőnek tekinteni, amennyiben azok hossza is és iránya is egyenlő. Az alábbiakban a sík egy rögzített pontjából kiinduló úgynevezett helyvektorok halmazán értelmezett összeadás és skalárral való szorzás tulajdonságait foglaljuk össze. Annak érdekében, hogy ábráink szemléletesek legyenek, a vektorokat Descartes–féle derékszögű koordináta rendszerben helyeztük el, de hangsúlyozni szeretnénk, hogy sem a vektorok összeadásánál, sem azok skalárral való szorzásakor a koordináta rendszer felvétele nem szükséges, annak, a definiálandó műveletek szemszögéből nincs szerepe.

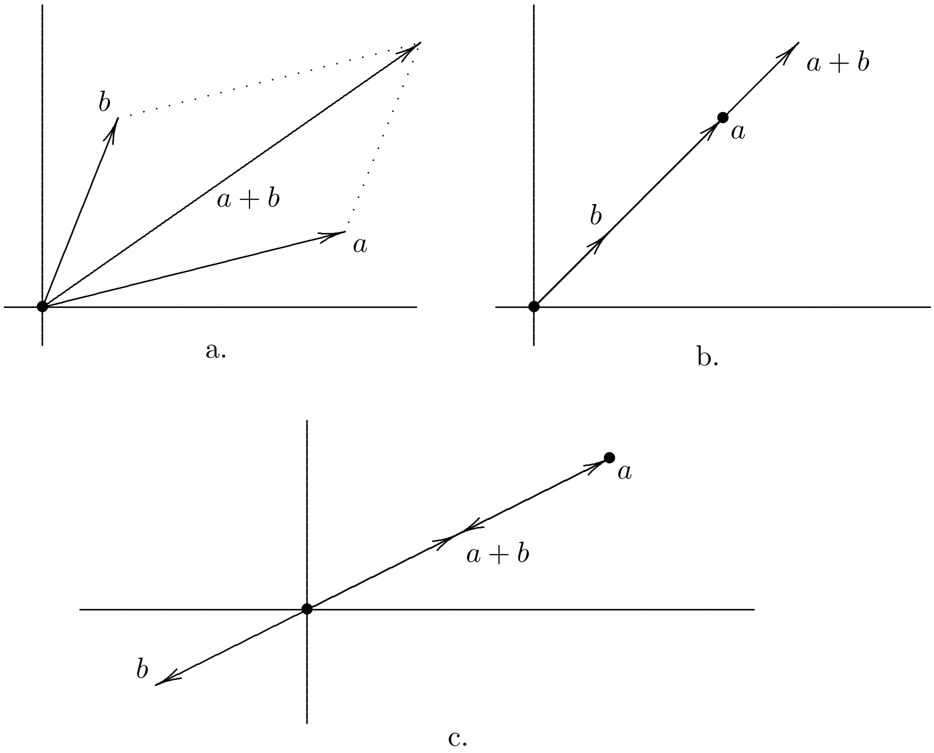
Két a és b vektor összeadása a paralelogramma– szabály alapján történik, az alábbi 1.1.a. ábrának megfelelően:

Természetesen értelmeznünk kell olyan vektorok összeadását is, amelyeknek azonos, vagy ellentétes az iránya. Ebben az esetben a második összeadandót az első végpontjába helyezzük és az első kezdőpontjából a második végpontjába mutató vektorral definiáljuk összegüket. (lásd az 1.1.b. és 1.1.c. ábrákat!) A vektorok összeadásának fenti definíciójából azonnal adódik, hogy az összeg független az összeadandók sorrendjétől, azaz a vektorok összeadása *kommutatív.* Adjunk most össze három vektort, az a-t, b-t és c-t. Ez kétféle sorrendben lehetséges, nevezetesen hozzáadhatjuk a-hoz a $(b + c)$ összeget, de az $(a + b)$-hez is hozzáadható a c vektor. Az 1.2.a. ábrán az első, az 1.2.b. ábrán pedig a második sorrendnek megfelelően képeztük három vektor összegét.

Azt tapasztaljuk, hogy ugyanaz a vektor adódik mindkét esetben. A vektorok összeadása tehát *asszociatív.* Vektoraink halmazában a zéró hosszúságú azzal a tulajdonsággal rendelkezik, hogy azt bármely a vektorhoz hozzáadhatjuk anélkül, hogy azt megváltoztatná. Tehát ugyanolyan szerepet játszik, mint a 0 a számok összeadásánál. Képezzük most egy tetszőleges a vektornak olyan $-a$-val jelölt vektorral való összegét, amelynek a hossza megegyezik a hosszával, de iránya ellentétes. Az 1.3 ábra mutat egy ilyen esetet.

Az összeg a zéró hosszúságú vektor.

Foglaljuk össze a vektorok V halmazán értelmezett összeadás fent illusztrált tulajdonságait. Tetszőleges a, b, és $c \in V$ esetén:



1.1. ábra: Vektorok összeadásának értelmezése

1. $a + b = b + a$ (a vektorok összeadása kommutatív),

2. $a + (b + c) = (a + b) + c$ (a vektorok összeadása asszociatív),

3. $(\exists\, 0 \in V) : 0 + a = a + 0 = a$ (létezik zéróvektor),

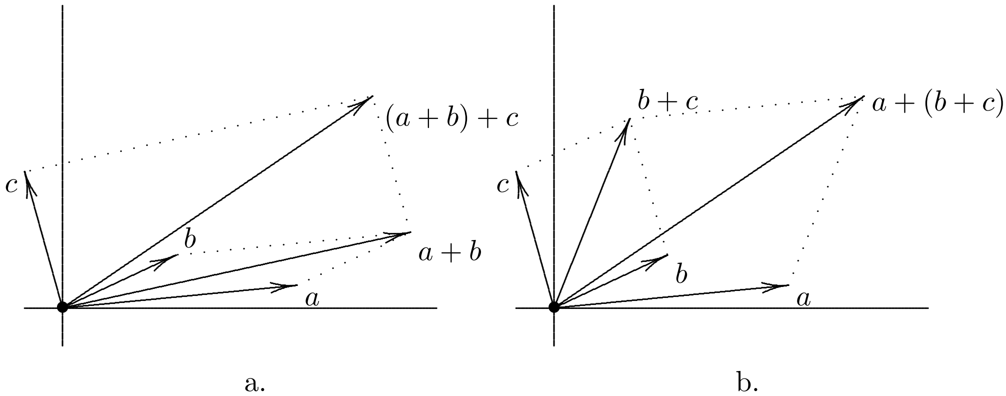
4. $(\forall\, a \in V) : (\exists\, (-a) \in V) : a + (-a) = (-a) + a = 0$ (minden vektornak van ellentettje).

Megállapíthatjuk tehát, hogy a síkbeli vektorok az összeadás művelettel *kommutatív csoportot* vagy más néven *Abel-csoportot* alkot. Definiáljuk egy vektornak számmal való szorzását az 1.4.a., illetve az 1.4.b. ábráknak megfelelően.

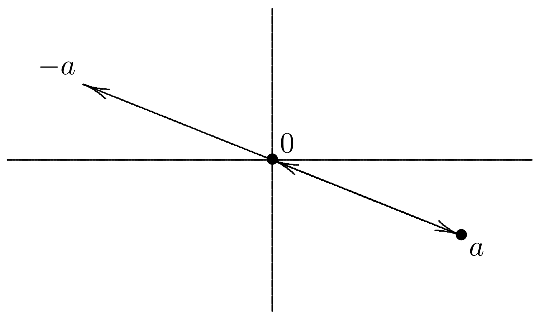
Tehát egy a vektor α-szorosa legyen olyan vektor, amelynek hossza a hosszának $|\alpha|$-szorosa iránya pedig a irányával megegyező, illetve azzal ellentétes, aszerint hogy α pozitív vagy negatív. Megvizsgáljuk, hogy a skalárokkal való, fent definiált szorzásnak milyen tulajdonságai vannak.

Az 1.5.a. és az 1.5.b. ábrák arról árulkodnak, hogy ugyanazt az eredményt kapjuk, ha két vektor összegét szorozzuk meg egy skalárral, mint amikor előbb az összeadandó vektorokat szorozzuk meg a számmal, s csak az így megváltoztatott vektorokat adjuk össze.

A skalárok összegével való szorzata egy a vektornak az 1.6.a. és 1.6.b. ábrák alapján ugyanazt a vektort eredményezi, mint annak a két vektornak az összege, amit az egyik, majd másik skalárral való szorzással kaptunk a-ból.



1.2. ábra: A vektorok összeadásának asszociativitása



1.3. ábra: Létezik nullvektor

Az 1.7.a. illetve 1.7.b. ábrák azt mutatják, hogy skalárok szorzatával úgy is lehet szorozni egy vektort, hogy előbb szorzunk az egyik skalárral, majd az így kapott vektort szorozzuk a másik skalárral.

A skalárokkal való szorzás definíciójából azonnal adódik az a tény, hogy bármely vektor 1-szerese a vektort nem változtatja meg.

Összefoglalva, ha a és $b \in V$ tetszőleges vektorok és α és β tetszőleges valós számok, akkor érvényesek az alábbi tulajdonságok:

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

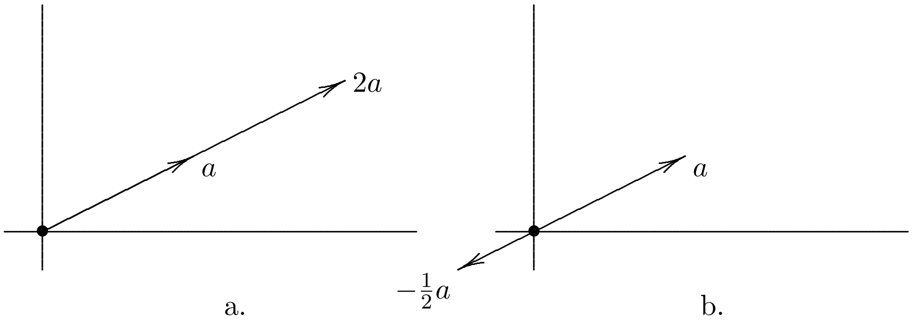
(b) $(\alpha + \beta)a = \alpha a + \beta a$,

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

(d) $1a = a$.

A vektorok skalárokkal való szorzása nem ugyanolyan művelet, mint azok összeadása, hiszen ebben az esetben egy másik halmaz, példánkban a valós számok halmazának elemei hatnak a vektorokra. Azt mondjuk, hogy a skalárok, mint *operátorok* hatnak a vektorokra. A valós számhalmaz a sík vektorainak úgynevezett *operátortartománya.*

A síkbeli helyvektorok V halmaza a definiált összeadással és valós skalárokkal való szorzással egy példát szolgáltat vektortérre, melyekkel e jegyzetben foglalkozni



1.4. ábra: Vektor skalárral való szorzása

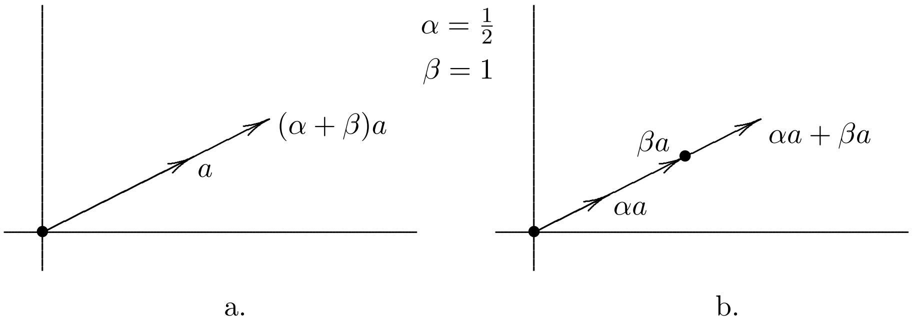
1.5. ábra: Összegvektor szorzása számmal

fogunk. A vektorterek számos helyen újra meg újra felbukkantak, és felbukkannak tanulmányaik során, persze más és más köntösben. Itt előbb, és már a középiskolában is, mint irányított szakaszok jelentkeztek, de valamely [a, b] intervallumon értelmezett valós függvények is vektorteret alkotnak a szokásos függvényösszeadással és a számmal való szokásos szorzással. Ami közös ezekben a vektorterekben az, hogy (1) az értelmezett összeadás művelet tulajdonságai azonosak, (2) valamilyen skalárhalmaz — hasonló a valós számok halmazához — elemei operátorokként hatnak a vektortér elemeire, és ugyanolyan tulajdonságokkal jellemezhető hatásuk, mint azt a fenti (a) — (d) pontokban láttuk.

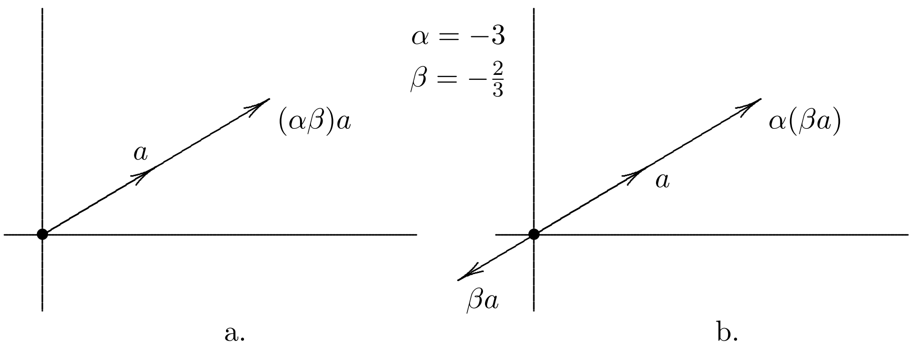
Persze, azt pontosan meg kell mondanunk, hogy milyen skalárhalmazt tekinthetünk a valós számok halmazához hasonlónak. Erre röviden azt válaszolhatjuk, hogy a skalárok struktúrájától azt kell megkövetelni, hogy test legyen.

1.4 A vektortér fogalma

Az előző szakaszban egy konkrét vektortérrel ismerkedtünk meg, a köznapi értelemben vett sík, úgynevezett helyvektorainak terével. Most megadjuk a vektortér absztrakt definícióját, hogy azután a minden vektortérre jellemző tulajdonságokat



1.6. ábra: Vektor szorzása számok összegével



1.7. ábra: Vektor szorzása számok szorzatával

ne kelljen minden konkrét vektortér esetén külön-külön igazolnunk, tehessük azt általánosan.

1.4.1 Definíció. *Legyen V egy halmaz, amelyben értelmezve van egy összeadás művelet az alább felsorolt tulajdonságokkal:*

(i) $(\forall\, x, y \in V) : x + y = y + x,$

(ii) $(\forall\, x, y, z \in V) : x + (y + z) = (x + y) + z,$

(iii) $(\exists\, 0 \in V) : (\forall\, x \in V) : 0 + x = x,$

(iv) $(\forall\, x \in V) : (\exists\, (-x) \in V) : x + (-x) = 0.$

Legyen $\mathbb{F}$ olyan halmaz, amelyen értelmezett egy összeadás és egy szorzás művelet, amelyekre nézve $\mathbb{F}$ testet alkot. Az $\mathbb{F}$ elemei legyenek operátorai V-nek, azaz $(\forall\, \alpha \in \mathbb{F})$-re

$$\exists\, x \longrightarrow \alpha x \quad (x \in V)$$

leképezés, a következő tulajdonságokkal:
$(\forall\, \alpha, \beta \in \mathbb{F})$ és $(\forall\, x, y \in V)$-re

(a) $\alpha(x + y) = \alpha x + \alpha y,$

(b) $(\alpha + \beta)x = \alpha x + \beta x$,

(c) $(\alpha\beta)x = \alpha(\beta x)$,

(d) $1x = x$, *ahol* $1 \in \mathbb{F}$ *az* $\mathbb{F}$ *test egységeleme.*

Akkor a V *halmazt az* $\mathbb{F}$ *test feletti vektortérnek nevezzük.*

A V halmaz elemeit vektoroknak, az $\mathbb{F}$ test elemeit skalároknak hívjuk. A fenti definícióban ugyanazt a $+$ szimbólumot használtuk mind a vektorok, mind a skalárok összeadásának jelölésére, de ez nem okozhat félreértést, hiszen a skalárok jelölésére görög, míg a vektorok jelölésére latin betűket használunk, így mindenütt nyilvánvaló, hogy vektorok vagy skalárok összeadásáról van e szó, egyedül a zéró vektor és a zéró skalár megkülönböztetéséről nem gondoskodik jelölésrendszerünk, mindkettőt a 0 szimbólum reprezentálja, de a környezet itt is az olvasó segítségére lesz annak kiderítésében, hogy a zéró skalárra vagy a zérusvektorra utal a 0 jel. A skalárok szorzatát egyszerűen a tényezők egymás mellé írásával jelöltük, csakúgy mint egy $\alpha \in \mathbb{F}$ operátor által az $x \in V$ vektorhoz rendelt $\alpha x \in V$ vektort. Egy operátor által indukált hozzárendelést egyszerűen *skalárral való szorzásnak* fogjuk hívni.

Vizsgálataink időnként valamely jól ismert számtest feletti vektorterekre irányulnak. Ilyenkor a következő terminológiát használjuk. Az $\mathbb{F}$ feletti V vektorteret racionális vektortérnek mondjuk, ha a skalárok csak racionális számok lehetnek, ha valós számok a skalárok, akkor V-t valós vektortérnek hívjuk, és ha komplex számok a vektorok operátorai, akkor V-t komplex vektortérnek nevezzük. Bár a vektortér definíciójában semmiféle kikötést nem tettünk az $\mathbb{F}$ test karakterisztikáját illetően, és az elmélet állításainak túlnyomó többsége tetszőleges karakterisztikájú test feletti vektorterekre is érvényes, de mi ebben a jegyzetben csak 0-*karakterisztikájú testek feletti vektorterekkel foglalkozunk.*

1.4.1 Példák vektorterekre

1. A bevezető szakaszban megismert, a sík egy rögzített pontjából, mint kezdőpontból kiinduló helyvektorok a paralelogramma–szabály szerinti összeadás művelettel és a megismert valós számokkal való szorzással, valós vektortér.

2. A mátrixaritmetikáról szóló pontban láttuk, hogy véve az összes $m \times n$ tipusú $\mathbb{F}$ test feletti mátrixok $\mathbb{F}_{m\times n}$-nel jelölt halmazát, az a mátrixok összeadásműveletével és az $\mathbb{F}$-beli skalárokkal való szorzással, vektorteret kapunk.

3. Tekintsük az összes t változójú, valós együtthatós polinomok $\mathbb{R}[t]$ halmazát. A
$$p(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_m t^m$$
m-edfokú és a
$$q(t) = \beta_0 + \beta_1 t + \cdots + \beta_n t^n$$
n-edfokú polinomok összege legyen
$$p(t) + q(t) \overset{\text{def}}{=} (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)t + \cdots + (\alpha_n + \beta_n)t^n,$$

ha $m = n$,

$$p(t) + q(t) \stackrel{\text{def}}{=} (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)t + \cdots + (\alpha_n + \beta_n)t^n + \cdots + \alpha_m t^m,$$

ha $m > n$ és

$$p(t) + q(t) \stackrel{\text{def}}{=} (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)t + \cdots + (\alpha_m + \beta_m)t^m + \cdots + \beta_n t^n$$

$n > m$ esetén.

A $p(t)$ polinom γ valós számszorosa pedig legyen

$$\gamma p(t) \stackrel{\text{def}}{=} (\gamma\alpha_0) + (\gamma\alpha_1)t + \cdots + (\gamma\alpha_m)t^m.$$

Tekintve, hogy a polinomok összeadása az azonos fokszámú tagok együtthatóinak összeadására van visszavezetve, azonnal adódik, hogy ez a művelet asszociatív és kommutatív, mivel a valós számok összeadása rendelkezik ezekkel a tulajdonságokkal. Az azonosan zéró polinom, tehát az, amelynek minden együtthatója nulla, zéróelem a polinomok fent definiált összeadására nézve. Egy tetszőleges $p(t)$ polinom ellentettje erre az összeadásra nézve a $-1p(t)$. A skalárokkal való szorzástól megkövetelt tulajdonságok teljesülése is azonnal adódik, ha figyelembe vesszük, hogy a valós együtthatóknak valós számokkal való szorzására vezettük vissza a polinomok skalárral való szorzását. Így a polinomok $\mathbb{R}[t]$ halmaza valós vektortér a definiált műveletekkel.

4. Legyen F az $[a, b]$ zárt intervallumon értelmezett valós függvények halmaza. Két $f, g \in F$ függvény összegét, illetve valós α számmal való szorzatát értelmezzük a szokásos módon, azaz legyen $\forall\, x \in [a, b]$-re

$$(f + g)(x) \stackrel{\text{def}}{=} f(x) + g(x)$$

és

$$(\alpha f)(x) \stackrel{\text{def}}{=} \alpha f(x).$$

Könnyen ellenőrizhető, hogy F valós vektortér az így definiált műveletekkel.

5. Legyen S az összes valós számsorozat, azaz az összes $f : \mathcal{N} \to \mathbb{R}$ függvények halmaza. Értelmezzük a számsorozatok összeadását és valós számmal való szorzását a szokásos módon, azaz bármely két $\{a_n\}$ és $\{b_n\}$ számsorozatra legyen

$$\{a_n\} + \{b_n\} \stackrel{\text{def}}{=} \{a_n + b_n\},$$

illetve bármely α valós számra és $\{a_n\} \in S$-re legyen

$$\alpha\{a_n\} \stackrel{\text{def}}{=} \{\alpha a_n\}.$$

Az így kapott struktúra megint valós vektortér.

6. Jelölje $\mathbb{R}_{n-1}[t]$ a legfeljebb $n-1$-edfokú valós együtthatós polinomok halmazát és e halmazbeli polinomokra ugyanúgy értelmezzük az összadást és a valós számmal való szorzást, mint azt az összes polinomok $\mathbb{R}[t]$ vektorterében. Azt tapasztalhatjuk, hogy $\mathbb{R}_{n-1}[t]$ is valós vektortér.

7. Az alább adott példa kicsit mesterkéltnek hat, mert tisztán matematikai konstrukció. A számítástechnikában valamennyire jártas olvasó azonban már találkozhatott olyan lineáris elrendezésű tömbökkel, amelyeknek valós számok az elemei. Az azonos méretű tömbök halmaza is vektortér megfelelő műveletekkel. Az előre bocsájtottakat pontosítandó jelölje $\mathbb{R}$, mint általában, a valós számok halmazát és legyen $\mathbb{R}^n$ a rendezett valós szám-n-esek halmaza. Definiáljunk összeadást $\mathbb{R}^n$-en a következőképpen: bármely

$$x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \quad \text{és} \quad y = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}\text{-ra legyen } \; x + y \stackrel{\text{def}}{=} \begin{bmatrix} \xi_1 + \eta_1 \\ \vdots \\ \xi_n + \eta_n \end{bmatrix}.$$

Ugyancsak értelmezzük egy tetszőleges

$$x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$$

szám-n-es $\alpha \in \mathbb{R}$ skalárral való szorzatát az

$$\alpha x \stackrel{\text{def}}{=} \begin{bmatrix} \alpha\xi_1 \\ \vdots \\ \alpha\xi_n \end{bmatrix}$$

egyenlőséggel. Nagyon könnyen igazolható, hogy $\mathbb{R}^n$ valós vektortér a fentiekben definiált műveletekkel, amelyet *n-dimenziós valós koordinátatérnek* nevezünk. Az elnevezés magyarázatát későbbre halasztjuk, egyelőre csak arra emlékeztetnénk az olvasót, hogy a sík vektoraihoz is rendeltünk koordinátákat a középiskolában, a koordinátatereknek hasonló reprezentációs szerepük lesz a vektorterek elméletében, mint a síkbeli helyvektorok koordinátáinak a középiskolai tanulmányaik során.

A figyelmes olvasó jogosan veti fel a kérdést, hogy milyen vektorteret kapunk, ha az $\mathbb{R}^n$ operátortartományaként csak a racionális számok testét vesszük. Természetesen racionális vektorteret, ami a fentiekben megadott tértől nagyon különbözik. Tehát egy vektortér megadása nemcsak a vektorok halmazának, de a skalárok halmazának megadását és a műveletek értelmezését is jelenti.

8. Tetszőleges $\mathbb{F}$ test esetén, hasonlóan értelmezve $\mathbb{F}^n$-beli n-esek összeadását és az $\mathbb{F}$-beli skalárokkal való szorzását, $\mathbb{F}^n$ vektortér lesz az $\mathbb{F}$ test felett, ez az úgynevezett *n-dimenziós $\mathbb{F}$-feletti koordinátatér.* Ebben az általános esetben is előfordulhat, hogy $\mathbb{F}^n$ operátortartományaként valamely $\mathbb{F}$-től különböző testet veszünk, de arra minden esetben fel fogjuk hívni az olvasó figyelmét, és ha külön nem specifikáljuk a skalárok testét, akkor az $\mathbb{F}^n$ jelöléssel mindig az n-dimenziós $\mathbb{F}$ feletti koordinátatérre utalunk.

9. Utolsó példaként megmutatjuk, hogy egy tetszőleges $\mathbb{F}$ feletti V vektortér birtokában hogyan lehet megkonstruálni egy másikat, az úgynevezett *duális vektorteret.*

1.4.2 Definíció. *Legyen V egy tetszőleges $\mathbb{F}$ test feletti vektortér és jelölje V^* az olyan $y^* : V \to \mathbb{F}$ leképezések, az úgynevezett lineáris függvények vagy más néven lineáris funkcionálok halmazát, amelyek eleget tesznek az alábbi feltételnek:*

$$\forall\, v, w \in V : \forall\, \alpha\beta \in \mathbb{F} : y^*(\alpha v + \beta w) = \alpha y^*(v) + \beta y^*(w)\,.$$

Értelmezzük most a V^* halmazon az összeadást a

$$\forall\, y_1^*, y_2^* \in V^* : \forall\, v \in V : (y_1^* + y_2^*)(v) \stackrel{\text{def}}{=} y_1^*(v) + y_2^*(v)$$

egyenlőséggel és az $\mathbb{F}$-beli skalárokkal való szorzást pedig a

$$\forall\, y^* \in V^* : \forall\, \alpha \in \mathbb{F} : \forall\, v \in V : (\alpha y^*)(v) \stackrel{\text{def}}{=} \alpha(y^*(v))$$

egyenlőséggel. A V^* vektortér ezekkel a műveletekkel $\mathbb{F}$ felett, amelyet a V vektortér duálisának nevezzük.

Tulajdonképpen igazolnunk kellene, hogy V^* valóban vektortér, de a részletes és formális bizonyítást az olvasóra bizzuk. Segítséget nyújthatnak ehhez a következő megjegyzések: Lineáris függvények összege is, egy lineáris függvény skalárszorosa is lineáris függvény. A lineáris függvények fenti összeadása kommutatív és asszociatív, hiszen a függvények összeadása vissza van vezetve a függvények értékeinek összeadására, és a függvényértékek az $\mathbb{F}$ test elemei. Az azonosan zéró függvény, tehát amely minden $v \in V$ vektorhoz a zéró skalárt rendeli, a lineáris függvények összeadására nézve zéróelem. Végül, egy y^* lineáris függvény ellentettje az a $-y^*$ függvény, amely bármely $v \in V$ vektornál a $-y^*(v)$ skalárt veszi fel értékül. Így tehát a lineáris függvények a definiált összeadásra nézve Abel-csoportot alkotnak. Az, hogy az $\mathbb{F}$ test elemeivel való szorzás ugyancsak eleget tesz a skalárokkal való szorzástól megkövetelt négy tulajdonságnak, megint csak azonnal adódik abból, hogy egy lineáris függvény skalárral való szorzása vissza van vezetve a függvény értékeinek skalárral való szorzására.

Fel kell hívjuk az olvasó figyelmét arra a tényre, hogy az $\mathbb{R}_{n-1}[t]$ vektortér és az $\mathbb{R}^n$ vektortér nagyon hasonlóak, legalábbis ami elemeik összeadását, illetve skalárral való szorzását illeti. Ugyanis, ha tetszőleges $p(t) = \alpha_0 + \alpha_1 t + \ldots + \alpha_{n-1} t^{n-1}$ polinomnak megfeleltetjük az együtthatókból felépített

$$a = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \end{bmatrix}$$

$\mathbb{R}^n$-beli szám n-est, akkor ez egy olyan egy–egyértelmű $\Phi : \mathbb{R}_{n-1}[t] \rightleftharpoons \mathbb{R}^n$ megfeleltetés, amely felcserélhető a vektortérbeli műveletekkel, azaz

$$\Phi(p(t) + q(t)) = \Phi(p(t)) + \Phi(q(t))$$

és

$$\Phi(\alpha p(t)) = \alpha \Phi(p(t))$$

teljesül. Azt mondhatjuk tehát, hogy a vektorterek végeredményben csak elemeik jelölésében térnek el egymástól.

Ez az észrevétel vezet el bennünket az izomorf vektorterek fogalmához.

1.4.3 Definíció. *Legyenek V és W ugyanazon $\mathbb{F}$ test feletti vektorterek és legyen*

$$\Phi : V \rightleftharpoons W$$

bijektív leképezés (egy-egyértelmű ráképezés), amely eleget tesz az alábbi feltételnek:

$$\forall\, x, y \in V : \forall\, \alpha, \beta \in \mathbb{F} : \Phi(\alpha x + \beta y) = \alpha\Phi(x) + \beta\Phi(y).$$

Akkor a V és W vektortereket izomorfoknak nevezzük. A $\Phi : V \rightleftharpoons W$ bijektív leképezést izomorfizmusnak hívjuk.

A valós együtthatós legfeljebb $n-1$-edfokú polinomok $\mathbb{R}_{n-1}[t]$ tere és az $\mathbb{R}^n$ koordinátatér tehát izomorfak. A bevezetőben vizsgált síkbeli helyvektorok vektortere és a kétdimenziós $\mathbb{R}^2$ koordinátatér ugyancsak izomorfak. Később még számos példát fogunk látni vektorterek izomorfiájára, többek között azt is be fogjuk bizonyítani, hogy az úgynevezett véges dimenziós vektorterek és duálisaik is izomorfak egymással.

Az eddigiek szerint egy vektortérben lehet vektorokat skalárral szorozni, ami vektort eredményez, és vektorokat össze lehet adni, aminek megint vektor az eredménye. Most ezen műveletekre támaszkodva értelmezzük a lineáris kombináció fogalmát. Ehhez szükség van skalároknak egy $\{\alpha_1, \dots, \alpha_n\}$ és vektoroknak egy $\mathcal{X} = \{x_1, \dots, x_n\}$ rendszerére. Azért használjuk itt a kicsit univerzális rendszer elnevezést a halmaz helyett, mert megengedett, hogy ugyanaz a skalár, illetve vektor több példánya is szerepeljen. Az $\mathcal{X}$ *vektorrendszer* $\{\alpha_i, i = 1, \dots, n\}$ skalárokkal képzett *lineáris kombinációja* az

$$\alpha_1 x_1 + \cdots + \alpha_n x_n$$

vektor. A rövidség kedvéért sokszor használni fogjuk a szummációs jelölést is, így például az előbbi lineáris kombinációt $\sum_{i=1}^{n} \alpha_i x_i$-vel is fogjuk jelölni. Az is előfordul, hogy a vektorok megkülönböztetésére egy I halmazból való indexeket használunk, ilyen esetben azok lineáris kombinációját $\sum_{i\in I} \alpha_i x_i$-nek írjuk. A lényeges az, hogy amikor lineáris kombinációról beszélünk, akkor egy olyan vektorra kell gondolnunk, amely előállítható véges sok vektor skalárszorosainak összegeként.

1. Gyakorlatok, feladatok

1. Mutassuk meg, hogy tetszőleges $\mathbb{F}$ test feletti V vektortérben

   (a) az αx vektor akkor és csak akkor a nullvektor, ha α az $\mathbb{F}$ test zéró eleme, vagy ha x a V zéró vektora.

   (b) az $x(\in V)$ vektor $-x$ ellentettje megkapható úgy, hogy az $\mathbb{F}$ test 1 egységelemének -1 ellentettjével szorozzuk az x vektort, azaz $-x = -1x$.

2. Legyen $C_{[a,b]}$ az $[a, b]$ intervallumon folytonos valós függvények halmaza, amelyet a függvények szokásos összeadásával és valós számmal való szokásos szorzásával teszünk struktúrává. Igazolja, hogy az így kapott struktúra valós vektortér.

3. Tekintsük a valós számok $\mathbb{R}$ halmazát a szokásos összeadással és valós számmal való szokásos szorzással. Mutassuk meg, hogy $\mathbb{R}$ valós vektortér a fenti műveletekkel.

4. Legyen $\mathbb{F}$ tetszőleges test. Mutassuk meg, hogy csakúgy, mint az előző feladatnál, $\mathbb{F}$ önmaga feletti vektortér.

5. Legyen $\mathbb{F}$ az ötelemű véges test. Soroljuk meg az $\mathbb{F}^2$ koordinátatér elemeit.

6. Legyenek V és W ugyanazon $\mathbb{F}$ test feletti vektorterek. Képezzük az $V \times W$ Descartes–szorzatot, és azon értelmezzük az összeadást a következőképpen:

$$\forall\,(v, w), (v', w') \in V \times W : (v, w) + (v', w') \overset{\text{def}}{=} (v + v', w + w').$$

Definiáljuk az $\alpha \in \mathbb{F}$ skalárral való szorzást is az

$$\forall\,(v, w) \in V \times W : \forall\,\alpha \in \mathbb{F} : \alpha(v, w) \overset{\text{def}}{=} (\alpha v, \alpha w)$$

egyenlőséggel. Az így kapott struktúrát jelöljük $V \otimes W$-nel. Mutassuk meg, hogy $V \otimes W$ vektortér az $\mathbb{F}$ test felett.

7. Legyen H tetszőleges nemüres halmaz, $\mathbb{F}$ test és V az összes olyan

$$f : H \longrightarrow \mathbb{F}$$

függvények halmaza, amelyek a H halmaz legfeljebb véges sok eleméhez rendelnek nemnulla skalárt. Értelmezzük V-beli függvények összeadását minden $f, g \in V$-re az

$$(f + g)(h) \overset{\text{def}}{=} f(h) + g(h) \quad (h \in H)$$

és $\mathbb{F}$-beli α skálrral való szorzását az

$$(\alpha f)(h) \overset{\text{def}}{=} \alpha f(h) \quad (h \in H)$$

egyenlőséggel. Mutassuk meg, hogy V $\mathbb{F}$ feletti vektortér.

8. + Jelölje $\mathbb{R}_+^{0n}$ a pozitív valós szám n-esek halmazát. Definiáljuk $\mathbb{R}_+^{0n}$-on az összeadást a következőképpen:

$$\forall\,x = [\xi_1, \ldots, \xi_n], y = [\eta_1 \ldots, \eta_n] \in \mathbb{R}_+^{0n} : x + y \overset{\text{def}}{=} [\xi_1 \cdot \eta_1, \ldots, \xi_n \cdot \eta_n],$$

ahol a jobboldalon a pozitív valós koordináták szokásos szorzata áll. És a valós α skalárral való szorzás legyen az

$$\alpha x \overset{\text{def}}{=} [(\xi_1)^\alpha, \ldots, (\xi_n)^\alpha]$$

egyenlőséggel megadva. Mutassuk meg, hogy $\mathbb{R}_+^{0n}$ valós vektortér ezekkel a műveletekkel.

1.5 Alterek

Az előzőekben láttuk, hogy a sík helyvektorai valós vektorteret alkotnak. De az egy egyenesre eső helyvektorok alkotta részhalmaz is vektortér. Így az általunk vizsgált síkbeli helyvektorok vektorterében vannak olyan részhalmazok, amelyek maguk is vektorterek.

A vektorterekre felsorolt példáink között szerepelt az összes polinomok $\mathbb{R}[t]$ tere és ugyancsak említettük a legfeljebb $n-1$-edfokú polinomok $\mathbb{R}_{n-1}[t]$ lineáris terét. Ez utóbbi térben a műveletek ugyanúgy voltak értelmezve, mint az összes polinomok terében, tehát az $\mathbb{R}_{n-1}[t]$ az $\mathbb{R}[t]$ vektortérnek olyan részhalmaza, ami maga is vektortér.

1.5.1 Definíció. *Legyen V vektortér az $\mathbb{F}$ test felett és legyen M olyan részhalmaza V-nek, amely maga is $\mathbb{F}$ feletti vektortér az eredeti térben értelmezett műveletekkel. Akkor M-et a V vektortér* alterének *nevezzük.*

Az altérbeli vektorok összeadása, és skalárral való szorzása ugyanúgy történik, mint az eredeti vektortérben, nem értelmezünk új műveleteket. Ezért a műveletek nyilvánvalóan rendelkeznek azokkal a tulajdonságokkal, amelyeket a vektortér definíciójában megköveteltünk. Ennek az észrevételnek egyszerű következménye a következő állítás.

1.5.2 Állítás. *Egy $\mathbb{F}$ test feletti V vektortér valamely nemüres M részhalmaza pontosan akkor altér, ha*

(i) $\forall\, v, w \in M : v + w \in M$,

(ii) $\forall\, v \in M : \forall\, \alpha \in \mathbb{F} : \alpha v \in M$ *teljesül.*

Bizonyítás. A szükségesség teljesen nyilvánvaló. Az elegendőség igazolásához csak azt kell megmutatnunk, hogy a zéróvektor és minden M-beli vektor ellentettje is benne van M-ben. Ez viszont abból adódik, hogy a zéró skalárnak bármely vektorral való szorzata a zéró vektort adja, és bármely vektor -1-gyel való szorzata annak ellentettjét eredményezi. □

Amikor egy vektortér valamely részhalmazáról azt kell eldöntenünk, hogy az altér-e, akkor leggyakrabban a következő állítás használható a kérdés megválaszolásához.

1.5.3 Állítás. *Egy $\mathbb{F}$ test feletti V vektortér valamely nemüres M részhalmaza pontosan akkor altér, ha zárt a két tagú lineáris kombináció képzésre nézve, azaz*

$$\forall\, v, w \in M : \forall\, \alpha, \beta \in \mathbb{F} : \alpha v + \beta w \in M.$$

Bizonyítás. Tekintve, hogy az előző tétel igaz, elegendő belátnunk azt, hogy ez az állítás az előzővel ekvivalens. Valóban, ha M zárt a skalárral való szorzásra, akkor $\forall\, v, w \in M : \forall\, \alpha, \beta \in \mathbb{F} : \alpha v, \beta w \in M$ és a vektorok összeadására vonatkozó zártsága miatt, akkor $\alpha v + \beta w \in M$.

Fordítva, a lineáris kombinációra vonatkozó zártságból az összeadásra, illetve a skalárral való szorzásra vonatkozó zártság azért nyílvánvaló, mert ezek speciális lineáris kombinációk. Ugyanis

$$\alpha v + \beta w = \begin{cases} v + w & \text{ha} \quad \alpha = \beta = 1, \\ \alpha v & \text{ha} \quad \beta = 0. \end{cases}$$

□

Megjegyezzük, hogy teljes indukcióval könnyen igazolható, hogy ha egy vektortér egy részhalmaza bármely két vektorának minden lineáris kombinációját tartalmazza, akkor bármely véges sok vektorának összes lineáris kombinációját is tartalmazza.

Az alterek között azt, amelynek egyetlen eleme a zérus vektor, azaz a $\{0\}$ alteret, és az egész vektorteret, mint önmaga alterét nem valódi altereknek mondjuk, a többi alteret pedig valódi altereknek. Egy altér mindig tartalmazza legalább a nullvektort, így alterek közös része biztosan nemüres, sőt igaz a következő állítás.

1.5.4 Állítás. *Egy vektortér altereinek metszete is altér.*

Bizonyítás. Ha az alterek metszetéből kiválasztunk tetszőlegesen két vektort, akkor azok mindegyik altérnek elemei, ezért az (1.5.3) állítás szerint, azok bármely lineáris kombinációja is mindegyik altérnek eleme, így megint az (1.5.3) állítás felhasználásával adódik, hogy a metszet altér. □

Az (1.5.4) állítás igaz volta lehetővé teszi, hogy egy V vektortér tetszőleges A részhalmaza segítségével generáljunk alteret.

1.5.5 Definíció. *Legyen A tetszőleges részhalmaza a V vektortérnek. Az A által generált* $\mathrm{V}(A)$ *altér V összes A-t tartalmazó altereinek közös része.*

Az A részhalmazt V *generátorrendszerének* fogjuk nevezni, ha $\mathrm{V}(A) = V$ teljesül, tehát ha az általa generált altér az egész vektortér.

Az A vektorhalmazhoz másképpen is rendelhető altér.

1.5.6 Definíció. *Legyen* $\mathrm{lin}\,(A)$ *az A-beli vektorok összes lineáris kombinációját tartalmazó vektorok halmaza, üres A halmaz esetén pedig, megállapodás alapján az egyelemű — a nullvektort tartalmazó — halmaz.* $\mathrm{lin}\,(A)$*-t az A lineáris burkának nevezzük.*

Megjegyzés: Hangsúlyoznunk kell, hogy amennyiben A egy végtelen vektorhalmaz, akkor $\mathrm{lin}\,(A)$ képzésekor az A összes véges részhalmazának összes lineáris kombinációját kell vennünk, hiszen csak véges vektorrendszerre értelmeztük a lineáris kombináció fogalmát.

Nem nehéz belátnunk, hogy $\mathrm{lin}\,(A)$ is altér. Az (1.5.3) állítás szerint ehhez elég azt megmutatni, hogy $\mathrm{lin}\,(A)$-beli vektorok lineáris kombinációja is $\mathrm{lin}\,(A)$-ban van. Valóban ha

$$a = \sum_{i \in I} \alpha_i a_i, \quad \text{és} \quad a' = \sum_{j \in J} \beta_j a'_j$$

tetszőleges $\mathrm{lin}\,(A)$-beli vektorok továbbá δ és γ tetszőleges skalárok, akkor

$$\delta a + \gamma a' = \sum_{i \in I} \delta \alpha_i a_i + \sum_{j \in J} \gamma \beta_j a'_j$$

is $\mathrm{lin}\,(A)$-beli vektor, hiszen A vektorainak lineáris kombinációja.

Egy vektortér valamely A részhalmaza által generált altérnek definíció szerinti megadása nehezen megfogható. Megtalálni egy A részhalmazt tartalmazó összes alteret, majd azoknak a metszetét képezni, nem mutat rá, hogy a keletkező alteret milyen vektorok alkotják. Ezért hasznos a következő tétel, amely kapcsolatot teremt egy vektorhalmazhoz a fentiek szerint rendelt két altér között.

1.5.7 Tétel. *Legyen A a V vektortér tetszőleges részhalmaza, és jelölje* $\mathrm{lin}\,(A)$ *az A lineáris burkát, és* $\mathrm{V}(A)$ *pedig az A által generált alteret. Akkor* $\mathrm{lin}\,(A) = \mathrm{V}(A)$.

Bizonyítás. A $\mathrm{lin}\,(A) = \mathrm{V}(A)$ igazolása végett vegyük először észre, hogy $\mathrm{V}(A) \subseteq \mathrm{lin}\,(A)$ hiszen $\mathrm{lin}\,(A)$ maga is egy A-t tartalmazó altér és $\mathrm{V}(A)$ az A-t tartalmazó alterek metszete. Másrészt $\mathrm{lin}\,(A) \subseteq \mathrm{V}(A)$ nyilván teljesül, hiszen $\mathrm{V}(A)$ altér lévén, zárt a lineáris kombinációra, és így $\mathrm{lin}\,(A)$ minden elemét tartalmazza. □

Az (1.5.7) tételt kihasználva a továbbiakban egy vektortér valamely A részhalmaza által generált alteret $\mathrm{lin}\,(A)$-val fogjuk jelölni. Egy vektortér tetszőleges vektorrendszerének is képezhetjük a lineáris burkát, nem lényeges, hogy minden vektornak csak egy példánya szerepeljen a vektorrendszerben, és ez lehetővé teszi, hogy egy vektorrendszert is nevezhetünk ezentúl generátorrendszernek, amennyiben annak lineáris burka az egész tér.

Az (1.5.4) állítás szerint egy vektortér altereinek közös része is altér. Alterek egyesítéséről nem állíthatjuk ugyanezt, de az alterek egyesítése által generált altereknek van egy figyelemre méltó tulajdonsága.

1.5.8 Tétel. *Legyenek X és Y egy V vektortér alterei.*
(1) *Akkor az egyesítésük által generált* $\mathrm{lin}\,(X \cup Y)$ *altér minden vektora egy X-beli és egy Y-beli vektor összege.*
(2) *Ha $X \cap Y$ a zérus altér, akkor a* $\mathrm{lin}\,(X \cup Y)$ *altér minden vektora egyértelműen áll elő egy X-beli és egy Y-beli vektor összegeként.*

Bizonyítás. (1) A $\mathrm{lin}\,(X \cup Y)$ elemei az $X \cup Y$ halmaz elemeinek lineáris kombinációi, azaz
$$\alpha_1 x_1 + \ldots + \alpha_s x_s + \beta_1 y_1 + \ldots + \beta_t y_t$$
alakúak. Mivel X és Y alterek,
$$\alpha_1 x_1 + \ldots + \alpha_s x_s \in X \quad \text{és} \quad \beta_1 y_1 + \ldots + \beta_t y_t \in Y$$
és a tétel első állítását ezzel igazoltuk.

A (2)-es állítás bizonyítása végett tegyük fel, hogy valamely $v \in \mathrm{lin}\,(X \cup Y)$ vektorra
$$v = x + y \quad \text{és} \quad v = x' + y' \;\; (x, x' \in X,\; y, y' \in Y)$$
teljesül. Akkor az $x + y = x' + y'$ egyenlőségből következik, hogy
$$x - x' = y' - y \in X \cap Y,$$
hiszen $x - x'$ nyilván X-beli és $y' - y$ pedig Y-beli vektor. Így, ha X-nek és Y-nek a zérusvektor az egyetlen közös eleme, akkor
$$x - x' = y' - y = 0, \quad \text{tehát} \quad x = x' \quad \text{és} \quad y = y'$$

amint azt állítottuk. □

Értelmezzük ezek után az alterek direktösszegének fogalmát.

1.5.9 Definíció. *Azt mondjuk, hogy a V vektortér az X és az Y altereinek direktösszege, jelölése $V = X \oplus Y$, ha $X \cap Y$ a zérusaltér, és $V = \operatorname{lin}(X \cup Y)$.*

Az (1.5.8) tételből azonnal következik a

1.5.10 Következmény. *A V vektortér az X és Y altereinek direktösszege akkor és csak akkor, ha az $X \cap Y$ a zérusaltér és minden $v \in V$ vektor előállítható egy X-beli x vektor és egy Y-beli y vektor összegeként.*

1 Példa. *Tekintsük most az $\mathbb{R}^n$ valós koordinátatér összes olyan vektorainak az L halmazát, amelyek komponenseinek összege zérus, tehát*

$$L = \{x = [\xi_1, \ldots, \xi_n] \mid \sum_{i=1}^{n} \xi_i = 0\}.$$

Megmutatjuk, hogy L altér.

Ehhez az (1.5.3) állítás szerint elegendő a lineáris kombinációra vonatkozó zártságot igazolni. Legyenek ezért

$$x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \quad \text{és} \quad y = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}$$

tetszőleges L-beli vektorok és α és β teszőleges valós számok. Az $\alpha x + \beta y$ vektor komponenseinek az összege

$$\sum_{i=1}^{n} (\alpha \xi_i + \beta \eta_i) = \alpha \sum_{i=1}^{n} \xi_i + \beta \sum_{i=1}^{n} \eta_i = \alpha 0 + \beta 0 = 0.$$

Ezért $\alpha x + \beta y \in L$ és így L valóban altér. □

A következő példa kedvéért bevezetünk egy új fogalmat.

1.5.11 Definíció. *Legyen M az $\mathbb{F}$ test feletti V vektortér valamely részhalmaza (nem feltétlenül altér), és legyen M° a V^* duális vektortér azon részhalmaza, amely pontosan azokat az y^* lineáris függvényeket tartalmazza, amelyekre*

$$\forall\, x \in M : y^*(x) = 0$$

teljesül. Az M° halmazt az M annullátorának nevezzük.

2 Példa. *Mutasuk meg, hogy M° altere a duális V^* vektortérnek.*

Megoldás: Mindenekelőtt megjegyezzük, hogy M° nemüres, hiszen az azonosan zéró lineáris függvény minden vektorhoz a zéró skalárt rendeli, így M° biztosan tartalmazza az azonosan zéró lineáris függvényt. Az (1.5.3) állítás szerint azt kell még megmutatnunk, hogy M°-beli lineáris függvények tetszőleges lineáris kombinációja is M°-ben van. Ha $y_1^*, y_2^* \in M^\circ$ továbbá α és β tetszőleges $\mathbb{F}$-beli skalárok, akkor bármely $x \in M$-re

$$(\alpha y_1^* + \beta y_2^*)(x) = \alpha y_1^*(x) + \beta y_2^*(x) = \alpha 0 + \beta 0 = 0,$$

igazolva, hogy $\alpha y_1^* + \beta y_2^* \in M^\circ$. □

2. Gyakorlatok, feladatok

1. Igazolja, hogy amennyiben Y altere a V vektortér X alterének, akkor Y altere V-nek is.

2. Mutassuk meg, hogy ha A és B olyan részhalmazai a V vektortérnek, hogy $A \subseteq B$ teljesül, akkor $\mathrm{lin}\,(A)$ altere $\mathrm{lin}\,(B)$-nek.

3. Legyen

$$L = \{x \in \mathbb{R}^n \mid x = [\alpha, \alpha+\delta, \ldots \alpha + (n-1)\delta]; \alpha, \delta \in \mathbb{R}\},$$

   tehát L elmei azok az n-esek, amelyeknek egymást követő elemei számtani sorozatot alkotnak. Altér–e L?

4. Legyen L olyan $x \in \mathbb{R}^n$ vektorok halmaza, amelyeknek komponensei egy mértani sorozat egymás utáni elemei. Altér-e L?

5. Tekintsük az $[a, b]$ intervallumon Riemann-integrálható függvények vektorterének azon I részhalmazát, amely pontosan azokat az f függvényeket tartalmazza, melyekre $\int_a^b f(x)dx = 0$ teljesül. Mutassuk meg, hogy I altér.

6. Igazolja, hogy a valós Cauchy-sorozatok alterét alkotják az összes valós számsorozatok vektorterének. Igaz vajon ugyanez az állítás a divergens valós számsorozatokra is?

7. Ha V és W az $\mathbb{F}$ test feletti vektorterek, akkor az az első szakasz 6. feladatában látott módon értelmezett $V \otimes W$ is $\mathbb{F}$ feletti vektortér. Mutassuk meg, hogy vannak $V \otimes W$-nek olyan $\bar{V}$ és $\bar{W}$ alterei, amelyek izomorfak V-vel, illetve W-vel, és $V \otimes W$ az $\bar{V}$ és $\bar{W}$ altereinek direktösszege.

1.6 Lineáris függetlenség és összefüggőség

A címben jelzett lineáris függetlenség, illetve lineáris összefüggőség vektorrendszerekre vonatkozó fogalmak, és a lineáris algebra talán legalapvetőbb fogalmai. Olyannyira, hogy megértésük nélkül nem építhető tovább a vektorterek elmélete. Ezért a következő részben kicsit talán a szükségesnél is több magyarázattal fog találkozni az olvasó.

A középiskolában vektoralgebrai tanulmányaik során megállapították, hogy ha adott a síkban két tetszőleges, nemzéró és nem egy egyenesre eső vektor, akkor ezek lineáris kombinációjaként a nullvektor csak úgy kapható, ha mindkét vektort a zéró skalárral szorozzuk, majd az így kapott vektorokat adjuk össze. Teljesen hasonlóan, ha tekintünk három nem egy síkba eső térbeli helyvektort, ezek lineáris kombinációja csak abban az esetben egyenlő a nullvektorral, ha a lineáris kombinációban szereplő skalár együtthatók mindegyike nulla. Ezt a tulajdonságot általánosítva jutunk a lineárisan független vektorrendszer alábbi fogalmához.

1.6.1 Definíció. *Egy vektortér vektorainak egy $\mathcal{X} = \{x_1, \ldots, x_n\}$ rendszerét lineárisan függetlennek nevezzük, ha*

$$\sum_{i=1}^{n} \alpha_i x_i = 0 \Longrightarrow \alpha_1 = \ldots = \alpha_n = 0\,.$$

Szavakkal is megfogalmazva ugyanezt, azt mondhatjuk, hogy egy vektorrendszer pontosan akkor lineárisan független, ha vektorainak a lineáris kombinációja csak úgy lehet a zéróvektor, ha a lineáris kombinációban szereplő skalárok mindegyike a zéró skalár. Azt a lineáris kombinációt, amelyben minden skalár együttható zérus, *triviális lineáris kombinációnak* mondjuk. Természetesen bármely vektorrendszer triviális lineáris kombinációja zéróvektor, de vannak olyan vektorrendszerek is — például magát a zéróvektort is tartalmazó vektorrendszerek — amelyek nemcsak triviális lineáris kombinációval tudják a zéróvektort előállítani.

A lineárisan nem független vektorrendszert *lineárisan összefüggőnek* vagy röviden lineárisan függőnek hívjuk. Ez tehát azt jelenti, hogy egy $\{y_1, \ldots, y_m\}$ vektorrendszer pontosan akkor lineárisan összefüggő, ha *léteznek* olyan $\beta_1, \ldots \beta_m$ skalárok, hogy legalább egy közülük nem nulla, mégis $\sum_{i=1}^{m} \beta_i y_i = 0$.

Felvetődik a kérdés, hogy lineárisan független, vagy összefüggő az üres vektorrendszer, amelynek egyetlen vektor sem eleme. Tekintve, hogy lineárisan összefüggő csak akkor lehetne, ha vektorainak nemtriviális lineáris kombinációjaként előállítható lenne a zéróvektor, és ez az üres vektorrendszerre nem teljesül, az üres vektorrendszer lineárisan független.

Tekintsük a valós polinomok $\mathbb{R}[t]$ vektorterében a $p_1(t) = t$, $p_2(t) = t - 1$, $p_3(t) = t + 1$ vektorokat (polinomokat). Ezek lineárisan függő rendszert alkotnak, hiszen a $p_2(t) + p_3(t) - 2p_1(t) \equiv 0$. Megjegyzendő ugyanakkor, hogy $\mathbb{R}[t]$-ben vannak tetszőleges elemszámú lineárisan független vektorrendszerek is, mert tetszőleges n pozitív egész mellett, az $\{1, t, \ldots, t^n\}$ vektorrendszer (polinomrendszer) lineárisan független, hiszen ezen polinomok lineáris kombinációja csak úgy lehet az azonosan zéró polinom, azaz

$$\alpha_0 1 + \alpha_1 t + \cdots + \alpha_n t^n \equiv 0,$$

ha

$$\alpha_0 = \alpha_1 = \ldots = \alpha_n = 0.$$

3 Példa. *Mutassuk meg, hogy ha az $\{x_1, \ldots, x_n\}$ vektorrendszer lineárisan független, akkor az $\{x_1, x_2 - x_1, \ldots, x_n - x_{n-1}\}$ vektorrendszer is lineárisan független!*

A definíció értelmében a vektorrendszerről úgy láthatjuk be, hogy lineárisan független, ha sikerül megmutatni, hogy vektorai lineáris kombinációja csak úgy lehet a nullvektor, ha minden skalár együttható nulla. Legyen ezért

$$\alpha_1 x_1 + \alpha_2(x_2 - x_1) + \cdots + \alpha_n(x_n - x_{n-1}) = 0,$$

ahol az $\{\alpha_i \ \ (i = 1, \ldots, n)\}$ együtthatók egyelőre ismeretlenek. Ha ezekről az együtthatókról ki tudjuk deríteni, hogy mindegyike csak nulla lehet, akkor a vektorrendszer lineárisan független. Ez tehát a feladatunk. A vektoregyenlet, baloldalának átrendezése után

$$(\alpha_1 - \alpha_2)x_1 + \cdots + (\alpha_{n-1} - \alpha_n)x_{n-1} + \alpha_n x_n = 0,$$

amiből az $\{x_1, \ldots, x_n\}$ vektorrendszer függetlensége miatt következik, hogy

$$\alpha_1 - \alpha_2 = \ldots = \alpha_{n-1} - \alpha_n = \alpha_n = 0.$$

De ha $\alpha_n = 0$ és $\alpha_{n-1} - \alpha_n = 0$, akkor $\alpha_{n-1} = 0$ és hasonlóan lépésről lépésre visszafelé haladva kapjuk, hogy $\alpha_1 = \alpha_2 = \ldots \alpha_{n-1} = \alpha_n = 0$. Ezzel megmutattuk, hogy a lineáris kombináció a triviális kellett legyen. □

A lineáris függetlenség, illetve a lineáris összefüggőség vektorrendszerek tulajdonsága, mégis sokszor fogjuk mondani vektorokra, hogy azok lineárisan függetlenek, vagy lineárisan függők. Természetesen ezen azt értjük, hogy az általuk alkotott vektorrendszer rendelkezik a mondott tulajdonsággal.

Az alábbiakban felsorolunk néhány nagyon egyszerű állítást, igazolásukat az olvasóra bizzuk.

1. a zéróvektort tartalmazó vektorrendszer lineárisan összefüggő,
2. ha egy vektorrendszerben egy vektor legalább két példánya benne van, akkor a rendszer lineárisan összefüggő,
3. egyetlen nemzéró vektort tartalmazó vektorrendszer lineárisan független,
4. lineárisan független vektorrendszer minden nemüres részrendszere is lineárisan független.

A következő tétel lineárisan összefüggő vektorrendszerek egy nagyon fontos tulajdonságát írja le.

1.6.2 Tétel. *Egy $\mathcal{X} = \{x_1, \ldots, x_n\}$ vektorrendszer pontosan akkor lineárisan összefüggő, ha valamelyik vektora kifejezhető a többi vektora lineáris kombinációjaként.*

Bizonyítás. Ha $\mathcal{X}$ lineárisan összefüggő, akkor van vektorainak olyan nemtriviális lineáris kombinációja ami zéróvektor. Legyen a zéróvektornak egy ilyen nemtriviális előállítása az

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0.$$

Legyen mondjuk az α_i nemzéró együttható. Akkor az

$$x_i = -\frac{\alpha_1}{\alpha_i}x_1 - \cdots - \frac{\alpha_{i-1}}{\alpha_i}x_{i-1} - \frac{\alpha_{i+1}}{\alpha_i}x_{i+1} - \cdots - \frac{\alpha_n}{\alpha_i}x_n$$

számolás mutatja, hogy a feltétel szükséges.

Fordítva, ha

$$x_i = \beta_1 x_1 + \cdots + \beta_{i-1}x_{i-1} + \beta_{i+1}x_{i+1} + \cdots + \beta_n x_n,$$

akkor

$$\beta_1 x_1 + \cdots + \beta_{i-1}x_{i-1} - x_i + \beta_{i+1}x_{i+1} + \cdots + \beta_n x_n = 0$$

egy nemtriviális előállítása a zéróvektornak — hiszen az x_i együtthatója nem nulla — tehát a vektorrendszer lineárisan összefüggő. □

A fentiekben bizonyított tétel kicsit élesíthető abban az esetben, ha a vektorrendszer nem tartalmazza a zéróvektort.

1.6.3 Tétel. *Egy a zéróvektort nem tartalmazó $\mathcal{X} = \{x_1, \ldots, x_n\}$ vektorrendszer pontosan akkor lineárisan összefüggő, ha létezik olyan i $(2 \leq i \leq n)$ index, hogy x_i kifejezhető az $x_1, \ldots, x_{i-1}$ vektorok lineáris kombinációjaként.*

Bizonyítás. Ha az $\mathcal{X}$ vektorrendszer lineárisan összefüggő, akkor létezik vektorainak olyan nemtriviális lineáris kombinációja, ami a nullvektort adja, mondjuk:

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$$

Tegyük fel, hogy i a legnagyobb indexű nemzéró együttható a lineáris kombinációban, azaz

$$\alpha_1 x_1 + \cdots + \alpha_i x_i = 0$$

is teljesül és $\alpha_i \neq 0$. Mivel egyetlen nemzéró vektorból álló vektorrendszer nyilván lineárisan független, $i \geq 2$. Akkor a vektoregyenlet átrendezésével kapjuk, hogy

$$x_i = -\frac{\alpha_1}{\alpha_i} x_1 - \ldots - \frac{\alpha_{i-1}}{\alpha_i} x_{i-1},$$

amint azt állítottuk.

Fordítva, ha

$$x_i = \beta_1 x_1 + \cdots + \beta_{i-1} x_{i-1}$$

valamely i $(2 \leq i \leq n)$-re, akkor

$$\beta_1 x_1 + \cdots + \beta_{i-1} x_{i-1} - x_i + 0 x_{i+1} + \cdots + 0 x_n = 0$$

a zéróvektor nemtriviális előállítása a vektorrendszer vektorainak lineáris kombinációjaként, igazolva annak lineáris összefüggőségét. □

Az a definíció nyilvánvaló következménye, hogy lineárisan összefüggő vektorrendszer vektorainak lineáris kombinációjaként a zéróvektor többféleképpen is előállítható, hiszen ha $\sum_{i=1}^m \beta_i y_i = 0$ egy nemtriviális előállítás és α egy tetszőleges nemzéró skalár, akkor $\alpha \sum_{i=1}^m \beta_i y_i = \sum_{i=1}^n \alpha\beta_i y_i = 0$ is nemtriviális előállítás lesz. Ez nemcsak a zéróvektorra igaz, hanem az is teljesül, hogy amennyiben egy a vektor előállítható egy lineárisan összefüggő vektorrendszer vektorainak lineáris kombinációjaként, akkor ez nemcsak egyféleképpen lehetséges. Valóban az $a = \sum_{i=1}^m \gamma_i y_i$ előállítás mellett az $a = \sum_{i=1}^m (\gamma_i + \alpha\beta_i) y_i$ előállítások is lehetségesek.

A következő tételben azt bizonyítjuk, hogy a lineárisan független vektorrendszerek esetében nemcsak a zéróvektor, de minden olyan vektor, amely előállítható a vektorrendszer vektorainak lineáris kombinációjaként, egyértelműen állítható elő.

1.6.4 Tétel. *Az $\mathcal{X} = \{x_1, \ldots, x_n\}$ vektorrendszer akkor és csak akkor lineárisan független, ha bármely olyan a vektor, amely benne van $\mathcal{X}$ lineáris burkában, egyértelműen állítható elő $\mathcal{X}$ vektorai lineáris kombinációjaként.*

Bizonyítás. Először feltesszük, hogy

$$a = \sum_{i=1}^{n} \alpha_i x_i \quad \text{és} \quad a = \sum_{i=1}^{n} \beta_i x_i$$

az a vektor előállításai. Akkor képezve a vektoregyenletek különbségét, kapjuk, hogy

$$0 = \sum_{i=1}^{n} (\alpha_i - \beta_i) x_i$$

és ebből mivel a feltétel szerint $\mathcal{X}$ lineárisan független

$$\alpha_1 - \beta_1 = \ldots = \alpha_n - \beta_n = 0\,,$$

tehát

$$\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n\,,$$

igazolva, hogy az a előállítása egyértelmű.

Az elegendőség nyilvánvaló, hiszen ha a zéróvektor, ami biztosan kifejezhető $\mathcal{X}$ vektorai triviális lineáris kombinációjaként, másképpen nem állítható elő, akkor ez, definíció szerint $\mathcal{X}$ lineáris függetlenségét jelenti. □

3. Gyakorlatok, feladatok

1. Legyen $\mathbb{F}^3$ az $\mathbb{F}$ test feletti koordinátatér. A $\mathbb{F}^3$-beli

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix}, \quad y = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix}, \quad z = \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix}$$

vektorrendszer lineáris függetlenségének vizsgálatát végezzük a következő módon. Tegyük fel, hogy α, β és γ olyan $\mathbb{F}$-beli skalárok, hogy $\alpha x + \beta y + \gamma z = 0$, ami a vektorok komponenseire vonatkozóan azt jelenti, hogy

$$\begin{aligned} \alpha\xi_1 + \beta\eta_1 + \gamma\zeta_1 &= 0, \\ \alpha\xi_2 + \beta\eta_2 + \gamma\zeta_2 &= 0, \\ \alpha\xi_3 + \beta\eta_3 + \gamma\zeta_3 &= 0. \end{aligned}$$

Ez tehát adott x, y és z vektorok mellett, egy három ismeretlenes egyenletrendszer α, β és γ-ra nézve. Az egyenletrendszert megoldhatjuk a középiskolában tanult módszerek valamelyikével. Ha azt találjuk, hogy az egyenletrendszernek az $\alpha = \beta = \gamma = 0$ az egyetlen megoldása, akkor az x, y, z vektorrendszer lineárisan független. Ha van más megoldás is, akkor a vektorrendszer lineárisan összefüggő.

Felhasználva a fentiekben adott módszert, igazolja, hogy az $\mathbb{R}^3$ térben az

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

vektorok rendszere lineárisan összefüggő, de közülük bármelyik három lineárisan független rendszert alkot!

2. Tételezzük fel, hogy valamely vektortérben az x, y és z vektorok lineárisan független rendszert alkotnak. Döntse el, hogy az $x+y$, $y+z$ és a $z+x$ vektorokból álló rendszer lineárisan összefüggő, vagy független!

3. Tekintsük a valós számok $\mathbb{R}$ halmazát, mint racionális vektorteret, azaz minden valós szám vektorként szerepel, amelyek összeadása a szokásos módon történik, de skalárszorzóként csak racionális számot engedünk meg. Mutassa meg, hogy az $\mathbb{R}$-beli 1 és ξ vektorok akkor és csak akkor alkotnak lineárisan független rendszert, ha ξ irracionális szám.

4. A $\mathbb{C}^2$ komplex vektortérben az
$$x = \begin{bmatrix} -1 \\ i \end{bmatrix} \quad \text{és} \quad y = \begin{bmatrix} i \\ 1 \end{bmatrix}$$
vektorok lineárisan összefüggő rendszert alkotnak, mert $x - iy = 0$. Igazolja, hogy ha $\mathbb{C}^2$-t valós vektortérként kezeljük, akkor ugyanezek az x, y vektorok lineárisan független rendszert alkotnak.

1.7 Vektortér dimenziója és bázisa

Az alterekről szóló részben említettük, hogy ha egy vektortér valamely részhalmaza által generált altér az egész tér, akkor azt a részhalmazt generátorrendszernek nevezzük. Ezt az (1.5.7) tétel bizonyítása után azzal egészítettük ki, hogy ha egy vektorrendszer lineáris burka az egész tér, akkor azt a vektorrendszert is generátorrendszernek nevezzük, függetlenül attól, hogy az halmaz, vagy esetleg egy vektorának több példánya is szerepel a rendszerben. Természetesen egy vektortérnek mindig van generátorrendszere, hiszen az egész vektortér önmaga generátorrendszere, de nagyon sok különböző valódi generátorrendszere is lehet. Az azonban már jellemző a vektortérre, hogy van-e véges sok vektort tartalmazó generátorrendszere, vagy mindegyik végtelen számosságú. Ennek megfelelően azt mondjuk, hogy egy vektortér *véges dimenziós,* ha van véges generátorrendszere; *végtelen dimenziós,* ha minden generátorrendszere végtelen sok vektort tartalmaz. Mi ebben a jegyzetben véges dimenziós terekkel kapcsolatos eredményeket közlünk, de az analízis tanulmányaink során megismert vektorterek között sok végtelen dimenziós volt.

1.7.1 Definíció. *A V végesen generálható vektorteret n-dimenziósnak mondjuk, ha van n vektort tartalmazó generátorrendszere, de bármely n-nél kevesebb vektort tartalmazó vektorrendszere már nem generátorrendszere V-nek. A V vektortér dimenzióját* $\dim V$*-vel jelöljük.*

A térbeli rögzített kezdőpontú, úgynevezett helyvektorok tere véges dimenziós, hiszen minden ilyen vektor felbontható három, nem egy síkba eső vektor irányába mutató vektor összegére, ami pontosan azt jelenti, hogy mindegyik megkapható három nem egy síkba eső vektor lineáris kombinációjaként. Igy ez a vektortér generálható három elemű generátorrendszerrel. Mivel háromnál kevesebb helyvektor nyilvánvalóan nem elegendő e vektortér generálásához, ezért az 3-dimenziós.

A valós együtthatós polinomok $\mathbb{R}[t]$ vektortere ezzel szemben végtelen dimenziós, hiszen egyetlen polinom sem állítható elő fokszámánál alacsonyabb fokú polinomok lineáris kombinációjaként, ezért minden n nemnegatív egészre kell legyen

n-edfokú polinom $\mathbb{R}[t]$ bármely generátorrendszerében. Nem nehéz belátni, hogy az $\{1, t, t^2, \ldots, t^n, \ldots\}$ végtelen sok polinomból álló rendszer generátorrendszere $\mathbb{R}[t]$-nek.

Az analízis tantárgy tanulásakor megismert legtöbb vektortér ugyancsak végtelen dimenziós, például az $[a, b]$ intervallumon folytonos függvények $C_{[a,b]}$ tere is.

A generátorrendszerek olyan szerepet játszanak a vektorterek építményében, mint például egy televízió összeállításánál az alkatrészek. A készülék összeállításánál használt kondenzátorok, illetve ellenállások némelyike azonban nem feltétlenül szükséges, azok helyettesíthetők más kapacitású, illetve ellenállású egységek soros és/vagy párhuzamos kapcsolásával. Persze nem hagyható el minden kondenzátor és minden ellenállás, mert azok nélkül TV készülék nem építhető.

Egy V vektortér valamely generátorrendszere is tartalmazhat nélkülözhető vektorokat, némelyik eleme esetleg elhagyható, amennyiben a megmaradt vektorok lineáris kombinációjaként az előállítható. Ennek megfelelően bevezetjük a következő elnevezést. Egy V vektortér valamely $\mathcal{X}$ generátorrendszerét *minimális generátorrendszernek* nevezzük, ha bármely vektorának elhagyásával kapott részrendszer már nem generátorrendszere V-nek. A minimális generátorrendszer elnevezés helyett gyakoribb a bázis szó használata. Tekintve, hogy véges dimenziós konkrét vektorterek vektorrendszereinek lineáris függetlenségét, illetve, hogy egy adott vektor kifejezhető-e a vektorrendszer elemeinek lineáris kombinációjaként, nem nehéz ellenőrizni, a bázis fogalmát a következő ekvivalens definícióval adjuk meg.

1.7.2 Definíció. *A véges dimenziós V vektortér egy $\mathcal{X}$ vektorrendszerét bázisnak vagy koordinátarendszernek nevezzük, ha*

(i) *lineárisan független rendszer,*

(ii) *V-nek generátorrendszere.*

Felmerülhet a kérdés, hogy mi a bázisa a zéró vektortérnek, hiszen annak egyetlen eleme a zéróvektor, és a zéróvektort tartalmazó vektorrendszerek lineárisan összefüggőek. Mivel az üres vektorhalmaz nyilván minimális generátorrendszere a zéró vektortérnek, a zéró vektortér bázisa az üres vektorrendszer. Ez összhangban van korábbi megállapodásunkkal miszerint az üres halmaz lineáris burka a zéró vektortér. A további vizsgálataink szempontjából a zéró vektortér érdektelen. Ezért, hogy állításaink megfogalmazása ne legyen feleslegesen körülményes, a továbbiakban feltételezzük, hogy a vizsgált vektorterek különböznek a zéró vektortértől.

Megjegyezzük, hogy a definíció (ii) feltétele azt jelenti, hogy a tér minden vektora kifejezhető az $\mathcal{X}$ vektorainak lineáris kombinációjaként, hiszen az (1.5.7) tétel értelmében, $\mathcal{X}$ pontosan akkor generátorrendszere V-nek, ha $\operatorname{lin}(\mathcal{X}) = V$. Másrészt az (i) feltétel miatt az (1.6.4) tétel felhasználásával az is adódik, hogy a tér minden vektora egyértelműen állítható elő a bázisvektorok lineáris kombinációjaként. Igy annak igazolásához, hogy véges dimenziós vektorterek esetén a minimális generátorrendszer fogalma és az általunk adott bázis definíció valóban ekvivalens csupán az alábbi állítást kell bizonyítanunk.

1.7.3 Állítás. *Egy véges dimenziós V vektortér valamely $\mathcal{X}$ generátorrendszere akkor és csak akkor minimális, ha lineárisan független.*

Bizonyítás. A szükségesség indirekt érveléssel azonnal adódik. Ha ugyanis $\mathcal{X} = \{x_1, \ldots, x_n\}$ minimális generátorrendszer, de lineárisan összefüggő, akkor valamelyik vektora, mondjuk x_i kifejezhető a többi vektora valamilyen

$$x_i = \alpha_1 x_1 + \cdots + \alpha_{i-1} x_{i-1} + \alpha_{i+1} x_{i+1} + \cdots + \alpha_n x_n \tag{1.1}$$

lineáris kombinációjaként. Akkor viszont V bármely

$$y = \beta_1 x_1 + \cdots + \beta_i x_i + \cdots + \beta_n x_n \tag{1.2}$$

vektora kifejezhető a $\mathcal{X} \setminus \{x_i\}$ vektorrendszer elemeinek lineáris kombinációjaként is, csupán az (1.2) egyenletben x_i helyébe az (1.1) egyenlet jobboldalát kell behelyettesíteni.

$$\begin{aligned} y = \beta_1 x_1 + \cdots + \beta_{i-1} x_{i-1} + \\ +\beta_i(\alpha_1 x_1 + \cdots + \alpha_{i-1} x_{i-1} + \alpha_{i+1} x_{i+1} + \cdots + \alpha_n x_n) + \\ +\beta_{i+1} x_{i+1} + \cdots + \beta_n x_n = \end{aligned}$$

$$\begin{aligned} = (\beta_1 + \beta_i \alpha_1) x_1 + \cdots + (\beta_{i-1} + \beta_i \alpha_{i-1}) x_{i-1} + \\ +(\beta_{i+1} + \beta_i \alpha_{i+1}) x_{i+1} + \cdots + (\beta_n + \beta_i \alpha_n) x_n \end{aligned}$$

Tehát az x_i vektor elhagyható anélkül, hogy a maradék rendszer megszűnne generátorrendszer lenni, ellentmondva az $\mathcal{X}$ generátorrendszer minimalitásának.

Az elegendőség talán még egyszerűbben kapható, ha ugyanis az $\mathcal{X}$ generátorrendszer lineárisan független, akkor egyik vektora sem fejezhető ki a többi vektora lineáris kombinációjaként, következésképpen, bármelyik vektorának elhagyásával a maradék rendszer már nem generátorrendszer, bizonyítva, hogy $\mathcal{X}$ minimális. □

Az előzőekben láttuk, hogy ha egy véges dimenziós vektortér egy generátorrendszeréből elhagyunk olyan vektort, amely előállítható a megmaradtak lineáris kombinációjaként, akkor a maradék vektorrendszer is generátorrendszer. Ha pedig már lineárisan független, akkor bázis. Tehát igaz az alábbi következmény:

1.7.4 Következmény. *Egy véges dimenziós vektortér minden generátorrendszere tartalmaz bázist.*

A véges dimenziós vektorterek dimenziója újrafogalmazható. Azt mondhatjuk, hogy egy V vektortér n-dimenziós, ha van n vektort tartalmazó bázisa, de — és ez egyelőre nem nyilvánvaló — nincs n-nél kevesebb vektort tartalmazó bázisa. A következő tételben éppen azt kívánjuk igazolni, hogy a tér dimenziója egy tetszőleges bázisa vektorainak számával egyenlő.

1.7.5 Tétel. *Tetszőleges véges dimenziós vektortér bármely két bázisában ugyanannyi vektor van.*

Bizonyítás. Tekintsünk két tetszőleges bázist, az egyik legyen az $\mathcal{X} = \{x_1, \ldots, x_n\}$, és a másik pedig az $\mathcal{Y} = \{y_1, \ldots, y_m\}$ vektorrendszer. Az $\mathcal{S}_1 = \{y_1, x_1, \ldots, x_n\}$ vektorrendszer nyilvánvalóan generátorrendszer, de nem lineárisan független rendszer, hiszen az y_1 a $\mathcal{X}$-beli vektorok lineáris kombinációja. Akkor, az (1.6.3) tétel szerint létezik olyan i $(1 \leq i \leq n)$ index, hogy x_i lineáris kombinációja az $y_1, x_1, \ldots x_{i-1}$ vektoroknak. (Az i index most azért lehet esetleg 1 is, mert a $\mathcal{S}_1$ vektorrendszerben x_1 már a második elem.) De akkor az $\mathcal{S}_1' = \mathcal{S}_1 \setminus \{x_i\}$

vektorrendszer is generátorrendszer. Folytassuk ezt a vektorcserét tovább, vegyük az $\mathcal{S}_2 = \mathcal{S}_1' \bigcup \{y_2\} = \{y_1, y_2, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n\}$ lineárisan összefüggő (mert $\mathcal{S}_1'$ generátorrendszer) vektorrendszert. Megint az (1.6.3) tétel alapján létezik olyan j $(1 \leq j \leq n, j \neq i)$ index, hogy x_j az y_1, y_2 és a j-nél kisebb indexű $\mathcal{S}_2$-beli x_k vektorok lineáris kombinációja. $\mathcal{S}_2' = \mathcal{S}_2 \setminus \{x_j\}$ tehát továbbra is generátorrendszer és az $\mathcal{S}_3 = \mathcal{S}_2' \bigcup \{y_3\}$ generátorrendszer pedig lineárisan összefüggő. Hasonlóan folytatva $\mathcal{X}$-beli vektorok $\mathcal{Y}$-beli vektorokkal való kicserélését végül eljutunk az $\mathcal{S}_m'$ generátorrendszerhez, ami már $\mathcal{Y}$ minden elemét tartalmazza, miközben pontosan m darab $\mathcal{X}$-beli vektort hagytunk el, tehát $\mathcal{X}$ elemeinek a száma nem lehetett kisebb m-nél. Az, hogy a vektorcserék során mindig $\mathcal{X}$-beli vektort kellett elhagynunk abból következik, hogy $\mathcal{Y}$ lineárisan független rendszer, így az eljárás bármelyik lépésénél az $\mathcal{S}_i$ generátorrendszert az $\mathcal{Y}$-beli vektorok mellett még szereplő $\mathcal{X}$-beli vektor(ok) tették lineárisan összefüggővé. Megismételve ezt a gondolatmenetet, csak most $\mathcal{X}$ és $\mathcal{Y}$ szerepét felcserélve, kapjuk, hogy $\mathcal{Y}$-nak sem lehet kevesebb eleme, mint $\mathcal{X}$-nek, azaz mindkét bázis ugyanannyi elemet kell tartalmazzon. □

Most már azt is állíthatjuk, hogy egy véges dimenziós vektortér bármely bázisában a vektorok száma a tér dimenziója.

Ha a bizonyítást újra végiggondoljuk látjuk, hogy kicsit általánosabb eredményt igazoltunk, és ezt következményként meg is fogalmazzuk.

1.7.6 Következmény. *Ha egy vektortérnek $\mathcal{X} = \{x_1, \ldots, x_n\}$ tetszőleges generátorrendszere és $\mathcal{Y} = \{y_1, \ldots, y_m\}$ pedig tetszőleges lineárisan független vektorrendszere, akkor $m \leq n$.*

Azt látjuk tehát, hogy egy lineárisan független vektorrendszer elemeinek a száma kisebb, vagy egyenlő, mint a vektortér dimenziója és egyenlőség pontosan akkor teljesül, ha a vektorrendszer egyúttal generátorrendszer is, azaz bázis. Felmerül a kérdés, hogy vajon nem lehetne-e egy tetszőleges lineárisan független vektorrendszert bázissá kiegészíteni? A válasz igenlő, amit az alábbi állításban fogalmaztunk meg.

1.7.7 Állítás. *Egy véges dimenziós V vektortérnek bármely lineárisan független $\mathcal{Y} = \{y_1, \ldots, y_m\}$ vektorrendszere kiegészíthető bázissá.*

Bizonyítás. Az $\mathcal{Y}$ vektorrendszer a V térnek valamilyen alterét generálja. Mivel ez az altér $\mathcal{Y}$ lineáris burka, ezért ebben $\mathcal{Y}$ bázis. Ha a $\operatorname{lin}(\mathcal{Y})$ valódi altér, akkor a $V \setminus \operatorname{lin}(\mathcal{Y})$ nemüres részhalmaz bármely x_1 vektora lineárisan független $\mathcal{Y}$ vektoraitól. Ha most az $\mathcal{Y}' = \{y_1, \ldots, y_m, x_1\}$ lineárisan független vektorrendszer már generátorrendszer is, akkor készen vagyunk, ellenkező esetben van V-nek olyan eleme, amely nincs benne $\mathcal{Y}'$ lineáris burkában és azzal bővíthetjük $\mathcal{Y}'$-t. Ez az eljárás folytatható mindaddig, amíg végül a kapott lineárisan független vektorrendszer már az egész V vektorteret generálja, tehát annak bázisa. V véges dimenziós volta biztosítja, hogy bővítési eljárásunk véges lépésben befejeződik. □

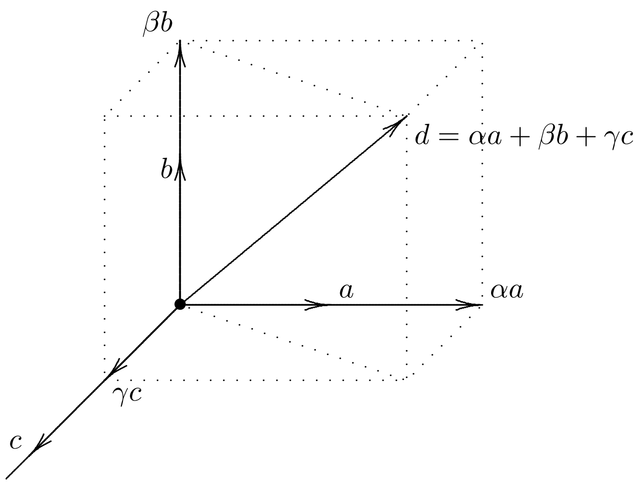
Egy lineárisan független vektorrendszer, újabb vektorok hozzávételével kibővíthető generátorrendszerré úgy, hogy lineáris függetlenségét megőrzi. Ettől kezdve azonban, már bármely vektor hozzávétele a vektorrendszert lineárisan összefüggővé teszi, azaz a tér bázisa *maximális, lineárisan független vektorrendszer.* A tér egy bázisa tehát megkapható úgy, hogy kiindulva egy generátorrendszeréből, a "felesleges" vektorokat elhagyva, azt minimalizáljuk, vagy egy lineárisan független vektorrendszeréhez új vektorokat hozzávéve addig bővítjük, amíg az a lineáris függetlenség megőrzése mellett lehetséges. Ezért igazak az alábbi tételben megfogalmazott állítások.

1.7.8 Tétel. *(a) Egy n-dimenziós V vektortérben bármely n+1 elemet tartalmazó vektorrendszer lineárisan összefüggő.*
(b) Egy n-dimenziós V vektortérnek egy n elemű vektorrendszere pontosan akkor bázis, ha lineárisan független.
(c) Egy n-dimenziós V vektortérnek egy n elemű vektorrendszere pontosan akkor bázisa, ha az generátorrendszer.

Mielőtt példákat adunk vektorterek dimenziójának meghatározására, megadjuk a vektorrendszerek rangjának fogalmát, amellyel sok más lineáris algebrával foglakozó könyvben találkozhat az olvasó.

Legyen $\mathcal{A} = \{a_1, \dots, a_m\}$ egy V vektortér valamely vektorrendszere. Ha létezik $\mathcal{A}$-nak r elemű lineárisan független részrendszere, de minden $r+1$ vektort tartalmazó részrendszere már lineárisan összefüggő, akkor az $\mathcal{A}$ *rangja* a nemnegatív r szám. Az $\mathcal{A}$ vektorrendszer rangját $\rho(\mathcal{A})$-val jelöljük. Nagyon könnyű igazolni, hogy $\rho(\mathcal{A})$ éppen az $\mathcal{A}$ vektorrendszer által generált $\operatorname{lin}(\mathcal{A})$ altér dimenziójával egyenlő, ezért vektorrendszerek rangjára vonatkozó állításokat külön nem fogalmazunk meg ebben a jegyzetben.

A térbeli, rögzített kezdőpontú, helyvektorok terében válasszunk három, páronként merőleges vektort, amit az alábbi 1.8 ábrán a-val, b-vel és c-vel jelöltünk, és vegyünk egy tetszőleges, d-vel jelölt negyediket.



1.8. ábra: Bázis a 3-dimenziós térben

Amint az látható, a d vektort elő tudtuk állítani az a, b és c vektorok lineáris kombinációjaként. Az a, b, c hármas tehát e térnek generátorrendszere. Persze az $\{a, b, c\}$ vektorrendszer lineárisan független is, azaz bázis. Tulajdonképpen az a, b, c vektorokkal a jól ismert térbeli Descartes–féle koordináta rendszert vettük fel, annak tengelyeit az a, b, c vektorok skalárszorosai alkotják. Ezért mondhatjuk, hogy maga az a, b, c vektorrendszer a koordináta rendszer.

4 Példa. *Megmutatjuk, hogy az $\mathbb{R}^n$ valós koordinátatér n-dimenziós.*

Ennek a kijelentésnek az igazolására megadjuk $\mathbb{R}^n$ egy n-elemű bázisát. Legyen

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

az az n-es, amelyben pontosan az i-dik komponens az 1-es és a többi 0, amit a továbbiakban i-dik egységvektornak nevezünk, és legyen $\mathcal{E} = \{e_1, \ldots, e_n\}$ az egységvektorok halmaza. Tekintve, hogy

$$\sum_{i=1}^{n} \alpha_i e_i = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Longleftrightarrow \alpha_1 = \ldots = \alpha_n = 0,$$

$\mathcal{E}$ független vektorrendszer, másrészt egy tetszőleges

$$x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{R}^n$$

vektor előállítható az $\mathcal{E}$-beli vektorok

$$x = \sum_{i=1}^{n} \xi_i e_i$$

lineáris kombinációjaként, igazolva, hogy $\mathcal{E}$ generátorrendszer is. Ezzel $\mathcal{E}$ bázis voltát igazoltuk, és mert n eleme van, azt is, hogy $\mathbb{R}^n$ n-dimenziós. □

A fenti igazolás szó szerint átvihető tetszőleges $\mathbb{F}$ test feletti $\mathbb{F}^n$ koordinátatérre, természetesen az $\mathbb{F}$ test egység– illetve zéróelemét használva a bázisvektorok konstruálásakor. Tehát állíthatjuk, hogy a tetszőleges $\mathbb{F}$ test feletti $\mathbb{F}^n$ koordinátatér n-dimenziós.

5 Példa. *Megmutatjuk, hogy ha a V vektortérnek M és N két véges dimenziós altere, akkor* $\dim(\operatorname{lin}(M \cup N)) = \dim M + \dim N - \dim(M \cap N)$.

Először is azt jegyezzük meg, hogy mivel az M és N alterek egyesítése által generált altér minden vektora egy M-beli és egy N-beli vektor összege,(lásd az (1.5.8) tételt) ezen altér dimenziója is biztosan véges lesz. Minthogy az $M \cap N$ altér egy $\mathcal{X}$ bázisa mind M-nek mind N-nek lineárisan független vektorrendszere, azt ki lehet egészíteni M egy $\mathcal{X}_M$ és N egy $\mathcal{X}_N$ bázisává. Ezen bázisok egyesítésével keletkezett halmaz, ami a közös elemeknek persze csak egy példányát tartalmazza, bázisa $\operatorname{lin}(M \cup N)$-nek és $\dim M + \dim N - \dim(M \cap N)$ elemű, és ez az, amit meg kellett mutatnunk. □

Fel kell hívjuk az olvasó figyelmét egy fontos következményre, nevezetesen arra, hogy amennyiben a V vektortér M és N altereinek direktösszege, azaz $V = M \oplus N$, akkor $\dim V = \dim M + \dim N$, hiszen ilyenkor az $M \cap N$ a zérus altér.

4. Gyakorlatok, feladatok

1. Mutassa meg, hogy ha A a V vektortér tetszőleges részhalmaza és v egy tetszőleges V-beli vektor, akkor a $\mathrm{lin}\,(A)$ és $\mathrm{lin}\,(A\cup\{v\})$ alterek dimenziója pontosan akkor egyenlő, ha v kifejezhető A vektorainak lineáris kombinációjaként!

2. Legyen $L = \{x \in \mathbb{R}^n \mid x = [\alpha, \alpha+\delta, \dots \alpha+(n-1)\delta], \alpha, \delta \in \mathbb{R}\}$, tehát L elemei azok az n-esek, amelyeknek egymást követő elemei számtani sorozatot alkotnak. Mint azt megmutatták az előző részt követő feladatok megoldása során, L altere $\mathbb{R}^n$-nek. Határozza meg L egy bázisát és ennek alapján a dimenzióját!

3. Tekintsük most az $\mathbb{R}^n$ valós koordinátatér összes olyan vektorainak az M halmazát, amelyek komponenseinek összege zérus, tehát
$$M = \{x = [\xi_1, \dots, \xi_n] \mid \sum_{i=1}^{n} \xi_i = 0\}.$$
Mint azt megmutattuk az előző részben, M altér. Adja meg M egy bázisát és ennek alapján állapítsa meg a dimenzióját!

4. A fentiekben igazoltuk, hogy a $\mathbb{C}^n$ komplex koordinátatér n-dimenziós a komplex számok $\mathbb{C}$ teste felett. Mennyi lesz a $\mathbb{C}^n$ tér, mint valós vektortér dimenziója, azaz, ha csak valós skalárokkal való szorzásra szorítkozunk?

5. Igazolja, hogy a legfeljebb $(n-1)$-edfokú valós polinomok terében azok a polinomok, amelyeknek az α zérushelye, alteret alkotnak! Állapítsa meg ezen altér dimenzióját és adja meg egy bázisát!

6. Mutassa meg, hogy ha V n-dimenziós vektortér és M V-nek r-dimenziós altere $(r \le n)$, akkor van olyan $n-r$-dimenziós N altere is, hogy V az M és N altereinek direkt összege!

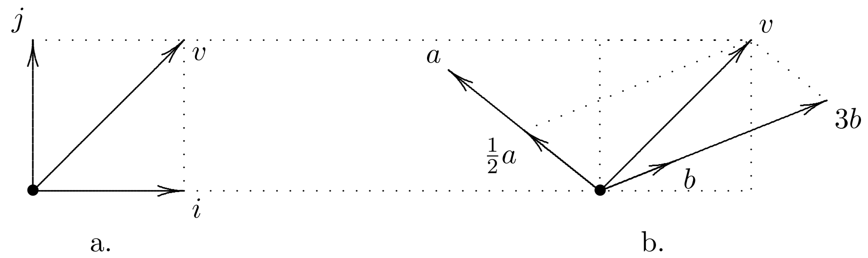
7. Az előző feladat felhasználásával igazolja, hogy egy n-dimenziós V vektortér előállítható n darab 1-dimenziós altere direkt összegeként!

1.8 Koordináta reprezentáció

Ha V egy n-dimenziós $\mathbb{F}$ test feletti vektortér, akkor bármely $\mathcal{X} = \{x_1, \dots, x_n\}$ bázisa segítségével a tér minden vektora előállítható, mint a bázisvektorok lineáris kombinációja, és ez az előállítás az (1.6.4 tétel szerint egyértelmű is. Ha tehát $v(\in V)$ egy tetszőleges vektora a térnek, akkor léteznek egyértelműen meghatározott $\varepsilon_1, \dots, \varepsilon_n(\in \mathbb{F})$ skalárok, úgy, hogy
$$v = \varepsilon_1 x_1 + \cdots + \varepsilon_n x_n.$$
Ezeket a $\varepsilon_1, \dots, \varepsilon_n$ skalárokat a v vektor $\mathcal{X}$ bázisra vonatkozó *koordinátái*nak nevezzük. Magát az (oszlopba) rendezett skalár-n-est a v vektor $\mathcal{X}$ bázisra vonatkozó *koordináta vektorának* hívjuk és $\mathbf{v}_{\mathcal{X}}$-el jelöljük. Tehát
$$\mathbf{v}_{\mathcal{X}} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

a v vektor $\mathcal{X}$ bázisra vonatkozó koordináta vektora. Felhívjuk az olvasó figyelmét arra, hogy ha szövegen belül kívánjuk megadni valamely vektor koordináta vektorát, akkor helykímélés céljából sorba rendezett n-esekkel reprezentáljuk azokat. skalár-n-essel

Remélve, hogy a koordinata reprezentáció bázistól való függőségének megértését segíti, az 1.9 ábra a síkbeli helyvektorok terében ugyanazon v vektor két különböző, a már középiskolából ismert $\mathcal{I} = \{i,\ j\}$ és egy másik, $\mathcal{A} = \{a,\ b\}$ bázis vektorainak lineáris kombinációiként van előállítva. Ennek megfelelően felírtuk v mindkét bázisra vonatkozó koordináta vektorát.



1.9. ábra: Koordináta reprezentáció különböző bázisokban

Az 1.9 ábráról leolvasható, hogy

$$\mathbf{v}_{\mathcal{I}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{míg} \quad \mathbf{v}_{\mathcal{A}} = \begin{bmatrix} 1/2 \\ 3 \end{bmatrix}.$$

Hangsúlyoznunk kell a különbséget a $v \in V$ vektor és a v vektor valamely bázisra vonatkozó koordináta vektora között, ami az $\mathbb{F}$ feletti $\mathbb{F}^n$ koordinátatér eleme, nem pedig V-beli. Ezt kívánjuk elérni azzal, hogy a koordináta vektort vastagon szedett betűvel jelöljük, tehát például a v vektor koordináta vektorát $\mathbf{v}$-vel, amelynek indexeként az alapul vett bázist is feltüntetjük. Ez persze sokszor bonyolulttá teszi a koordináta vektorok jelölését, ezért állapodjunk meg abban, hogy ha a szövegkörnyezetből kiderül, hogy melyik bázisra vonatkozó koordinátákról van szó, akkor az egyszerűség kedvéért a bázisra utaló indexet elhagyjuk. De nagyon fontos, hogy az olvasó megértse, hogy a vektortér vektorainak a koordinátái bázisfüggőek, és ugyanazon vektornak két különböző bázisra vonatkozó koordinátái különbözőek. Ami pedig egy vektor koordináta vektorát illeti, az még a bázis vektorainak rendezésétől is függ. Különösen bonyolulttá válik a helyzet abban az esetben, amikor magának az $\mathbb{F}^n$ koordinátatér vektorainak valamely bázisra vonatkozó koordináta vektorairól van szó, hiszen ekkor formailag nincs különbség a tér elemei és a koordináta vektorok között. Tartalmilag azonban igenis lényeges az eltérés, hiszen $\mathbb{F}^n$ egy mesterséges matematikai konstrukcióval képzett vektortér, nevezetesen $\mathbb{F}$-beli skalárok oszlopba rendezett n-eseinek halmaza, alkalmas operációkkal ellátva, míg egy elemének valamely bázisra vonatkozó koordináta vektora azt mutatja meg, hogy a kérdéses elemet a bázisvektorok milyen $\mathbb{F}$-beli skalárokkal képzett lineáris kombinációja állítja elő. Ezért egy tetszőleges $v \in \mathbb{F}^n$ skalár-n-es lehet egy más $w \in \mathbb{F}^n$ n-es valamely bázisra vonatkozó $\mathbf{w}$ koordináta vektorával egyenlő.

A következő tétel azt állítja, hogy lényegében minden vektortér olyan, mint egy koordinátatér, csak az elemek jelölésében van eltérés.

1.8.1 Tétel. *Ha V az $\mathbb{F}$ test feletti n-dimenziós vektortér, akkor izomorf az $\mathbb{F}^n$ koordinátatérrel.*

Bizonyítás. Legyen $\mathcal{X} = \{x_1, \ldots, x_n\}$ egy bázisa V-nek. Értelmezzük azt a

$$\Phi : V \longrightarrow \mathbb{F}^n$$

leképezést, amely minden

$$v = \varepsilon_1 x_1 + \cdots + \varepsilon_n x_n$$

V-beli vektorhoz hozzárendeli, annak $\mathcal{X}$ bázisra vonatkozó

$$\mathbf{v}_{\mathcal{X}} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

koordináta vektorát. Azt fogjuk megmutatni, hogy Φ izomorfizmus. A Φ leképezés egyértelműsége abból következik, hogy rögzített bázis mellett a koordináták egyértelműen meghatározottak. Ha két v, $w \in V$ vektornak ugyanaz a képe, akkor a

$$v - w = 0x_1 + \cdots + 0x_n = 0$$

számolás mutatja, hogy $v = w$, igazolva, hogy Φ egy–egyértelmű is. Továbbá minden $\mathbb{F}^n$-beli $[\alpha_1, \ldots, \alpha_n]$ vektor Φ képe valamely V-beli vektornak, nevezetesen az $\alpha_1 x_1 + \cdots + \alpha_n x_n$ vektornak, mutatva ezzel, hogy Φ ráképezés.

Ha $w = \omega_1 x_1 + \cdots + \omega_n x_n$ egy tetszőleges másik V-beli vektor és δ, $\gamma \in \mathbb{F}$ tetszőleges skalárok, akkor könnyű számolással kapjuk, hogy

$$\delta v + \gamma w = (\delta\varepsilon_1 + \gamma\omega_1)x_1 + \cdots + (\delta\varepsilon_n + \gamma\omega_n)x_n.$$

Így tehát azt kapjuk, hogy

$$\boldsymbol{\delta v} + \boldsymbol{\gamma w} = \begin{bmatrix} \delta\varepsilon_1 + \gamma\omega_1 \\ \vdots \\ \delta\varepsilon_n + \gamma\omega_n \end{bmatrix} = \delta \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} + \gamma \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix} = \delta\mathbf{v} + \gamma\mathbf{w},$$

ami éppen annak az igazolása, hogy

$$\Phi(\delta v + \gamma w) = \delta\Phi(v) + \gamma\Phi(w).$$

Ezzel bizonyítottuk, hogy Φ rendelkezik a műveletekkel való felcserélhetőség tulajdonságával is, tehát izomorfizmus. $\square$

Az (1.8.1) tétel bizonyítása rámutat arra, hogy egy $\mathbb{F}$ feletti V vektortér minden $\mathcal{X} = \{x_1, \ldots, x_n\}$ bázisa meghatároz egy $\Phi : V \rightleftharpoons \mathbb{F}^n$ izomorf leképezést. A Φ izomorf lképezést úgy is értelmezhettünk volna, hogy a bázisvektorokhoz az $\mathbb{F}^n$

$$\Phi(x_i) = e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad (i = 1, \ldots, n),$$

úgynevezett egységvektorait rendeljük, majd minden más V-beli v vektor képét annak a követelménynek a figyelembevételével határozzuk meg, hogy lineáris kombináció izomorf képe meg kell egyezzen a képvektorok ugyanazon skalárokkal képzett lineáris kombinációjával. Tehát ha

$$v = \varepsilon_1 x_1 + \cdots + \varepsilon_n x_n,$$

akkor

$$\Phi(v) = \varepsilon_1 \Phi(x_1) + \cdots + \varepsilon_n \Phi(x_n)$$

kell legyen. Ezek alapján egy $\mathbb{F}$ feletti V vektortér v elemének egy $\mathcal{X} = \{x_1, \ldots, x_n\}$ bázisra vonatkozó koordináta vektorát úgy tekinthetjük, mint a $\Phi : V \rightleftharpoons \mathbb{F}^n$ izomorfizmus által meghatározott $\Phi(v)$ képét.

Könnyen igazolható, hogy ha Θ a V vektortérnek a W vektortérre való izomorfizmusa, és $\{x_1, \ldots, x_n\}$ a V vektortér lineárisan független vektorrendszere vagy bázisa, akkor $\{\Theta(x_1), \ldots, \Theta(x_n)\}$ a W vektortérnek ugyancsak lineárisan független vektorrendszere, illetve bázisa.

Az (1.8.1) tétel felhasználásával nem nehéz igazolnunk az alábbi állítást.

1.8.2 Állítás. *Ha az $\mathbb{F}$ test feletti V és W vektorterek dimenziója egyenlő, akkor izomorfak.*

Bizonyítás. Legyen $\dim V = \dim W = n$. Akkor az (1.8.1) tétel szerint léteznek

$$\Phi : V \rightleftharpoons \mathbb{F}^n$$

és

$$\Psi : W \rightleftharpoons \mathbb{F}^n$$

izomorf leképezés. Tekintsük a

$$(\Psi^{-1} \cdot \Phi) : V \rightleftharpoons W$$

szorzatleképezést, amely minden $v \in V$ vektornak a $\Psi^{-1}(\Phi(v))$ vektort felelteti meg. Ez könnyen igazolhatóan V-nek W-re való izomorf leképezése. □

Az (1.7.7) állítás nyilvánvaló következménye, hogy egy n-dimenziós vektortérnek van m-dimenziós altere minden $m(\leq n)$ nemnegatív egész esetén. Ebből a tényből és az előző állításból azonnal adódik:

1.8.3 Következmény. *Ha az $\mathbb{F}$ test feletti V vektortér n-dimenziós a W vektortér pedig m-dimenziós és $m \leq n$, akkor V-nek van W-vel izomorf altere.*

1.8.4 Tétel. *Ha az $\mathbb{F}$ test feletti V vektortér n-dimenziós, akkor a duális V^* vektortér is n-dimenziós, következésképpen minden véges dimenziós vektortér izomorf a duálisával.*

Bizonyítás. Mivel $\dim V = n$, V-nek van n-elemű bázisa. Legyen $\mathcal{X} = \{x_1, \ldots, x_n\}$ egy ilyen bázis. Legyenek $\{x_1^*, \ldots, x_n^*\}$ olyan a V téren értelmezett függvények, melyekre

$$x_i^*(x_j) = \delta_{ij}, \quad (i, j = 1, \ldots, n) \text{ ahol } \delta_{ij} = \begin{cases} 1 & \text{ha} \quad i = j, \\ 0 & \text{ha} \quad i \neq j \end{cases}$$

az úgynevezett *Kronecker-szimbólum,* és tetszőleges

$$v = \varepsilon_1 x_1 + \ldots + \varepsilon_n x_n \tag{1.3}$$

vektorra, legyen

$$x_i^*(v) = \varepsilon_1 x_i^*(x_1) + \cdots + \varepsilon_n x_i^*(x_n) = \varepsilon_i \ \ (i = 1, \ldots, n). \tag{1.4}$$

Könnyen látszik, hogy az x_i^* függvény lineáris, és az $\mathcal{X}^* = \{x_1^*, \ldots, x_n^*\}$ különböző lineáris függvények rendszere V^*-ban. Megmutatjuk, hogy $\mathcal{X}^*$ bázisa V^*-nak. Először lássuk be, hogy az $\mathcal{X}^*$ függvényrendszer lineárisan független. Ha az

$$\xi_1 x_1^* + \cdots + \xi_n x_n^* = 0\,,$$

tehát minden V-beli vektornak a 0 skalárt felelteti meg, akkor az $\mathcal{X}$ bázis vektorait helyettesítve, kapjuk, hogy minden $i(= 1, \ldots, n)$-re

$$0 = (\sum_{j=1}^{n} \xi_j x_j^*) x_i = \sum_{j=1}^{n} \xi_j x_j^*(x_i) = \xi_i\,,$$

ami éppen $\mathcal{X}^*$ lineáris függetlenségét igazolja.

Legyen most y^* tetszőleges eleme V^*-nak. Tegyük fel, hogy

$$y^*(x_i) = \eta_i\,.$$

Megmutatjuk, hogy akkor

$$y^* = \eta_1 x_1^* + \cdots + \eta_n x_n^*\,, \tag{1.5}$$

ami azt fogja igazolni, hogy $\mathcal{X}^*$ generátorrendszere is V^*-nak.

Az y^* lineáris függvény az (1.3) egyenlőséggel adott tetszőleges $v \in V$ vektorhoz az

$$y^*(v) = y^*(\sum_{i=1}^{n} \varepsilon_i x_i) = \sum_{i=1}^{n} \varepsilon_i \eta_i$$

skalárt rendeli. Az (1.5) egyenlőség jobboldalán lévő lineáris függvény értéke a v helyen az (1.4) egyenlet szerint

$$(\sum_{j=1}^{n} \eta_j x_j^*)(v) = (\sum_{j=1}^{n} \eta_j x_j^*)(\sum_{i=1}^{n} \varepsilon_i x_i) =$$

$$\sum_{j=1}^{n} \eta_j (\sum_{i=1}^{n} \varepsilon_i \delta_{ij}) = \sum_{i=1}^{n} \eta_i \varepsilon_i \,,$$

tehát ugyanaz a skalár. Akkor az (1.5) egyenlőség teljesül, és ezzel megmutattuk, hogy $\mathcal{X}^*$ nemcsak lineárisan független, de generátorrendszere is V^*-nak, azaz bázis. $\square$

1.8.5 Definíció. *Legyen $\mathcal{X} = \{x_1, \dots, x_n\}$ bázisa V vektortérnek és legyenek $\mathcal{X}^* = \{x_1^*, \dots, x_n^*\}$ a V téren értelmezett olyan függvények, melyekre*

$$x_i^*(x_j) = \delta_{ij}, \quad (i, j = 1, \dots, n) \ \text{ahol} \ \delta_{ij} = \begin{cases} 1 & ha \quad i = j, \\ 0 & ha \quad i \neq j. \end{cases}$$

Akkor $\mathcal{X}^*$ a duális tér bázisa (lásd az előző tételt), amelyet a V vektortér $\mathcal{X}$ bázisához rendelt duális bázisnak nevezünk.

Egy véges dimenziós vektortér és duálisa közötti kapcsolatot az előző tétel jól jellemzi. Az alábbi állítás altereik kapcsolatát írja le.

1.8.6 Állítás. *Legyen az n-dimenziós V vektortérnek M r-dimenziós altere. Megmutatjuk, hogy M annullátora $(n-r)$-dimenziós altere V^*-nak.*

Bizonyítás. Emlékeztetjük az olvasót arra, hogy M annullátorán azoknak a V^*-beli y^* lineáris függvényeknek a halmazát értjük, amelyekre

$$y^*(x) = 0 \quad \text{minden } x \in M\text{-re}$$

teljesül. Az M r-dimenziós lévén, van r elemű bázisa. Legyen $\mathcal{X} = \{x_1, \dots, x_r\}$ M-nek egy ilyen bázisa, és egészítsük ezt ki az $\{x_{r+1}, \dots, x_n\}$ vektorok hozzávételével a V vektortér egy bázisává, ami az (1.7.7) állítás szerint lehetséges, hiszen V alterének egy bázisa nyilván lineárisan független vektorrendszere V-nek. Legyen $\mathcal{X}^* = \{x_1^*, \dots, x_r^*, x_{r+1}^*, \dots, x_n^*\}$ a duális bázis. Megmutatjuk, hogy $\{x_{r+1}^*, \dots, x_n^*\}$ bázisa M annullátorának, M°-nek. A duális bázis értelmezése alapján nyilvánvaló, hogy $x_{r+1}^*, \dots, x_n^*$ lineáris függvények mindegyike eleme M°-nek, és lineárisan független rendszer, hiszen egy bázis részrendszere. Másrészt, ha y^* tetszőleges eleme M°-nek, akkor

$$y^*(x_i) = 0 \quad (i = 1, \dots, r)\text{–re},$$

és

$$y^*(x_j) = \beta_j \quad (j = r+1, \dots, n)\text{-re}.$$

Így y^*-nak $\mathcal{X}^*$-beli függvények lineáris kombinációjaként való (1.5) egyenlőség szerinti előállításában

$$y^* = 0 \cdot x_1^* + \dots + 0 \cdot x_r^* + \beta_{r+1} x^{r+1} + \dots + \beta_n x_n^*$$

(lásd az előző tétel bizonyítását), csak az $x_{r+1}^*, \dots, x_n^*$ függvények együtthatói zérustól különbözőek, igazolva ezzel, hogy $\{x_{r+1}^*, \dots, x_n^*\}$ generátorrendszere is, és így bázisa M°-nek. Akkor viszont M° valóban $n-r$-dimenziós. $\square$

1.8.1 Elemi bázistranszformáció

Azt láttuk, hogy tetszőleges $\mathbb{F}$ test feletti n-dimenziós V vektortér izomorf az $\mathbb{F}^n$ koordinátatérrel. Azok után, hogy rögzítettük V egy $\mathcal{X}$ bázisát, a vektortér mindegyik v eleméhez egyértelműen hozzárendelhető egy $\mathbf{v}_{\mathcal{X}}$ oszlopba rendezett skalár n-es, a v vektornak az $\mathcal{X}$ bázisra vonatkozó, úgynevezett koordináta vektora, amelynek i-dik eleme éppen a v vektornak az i-dik bázisvektorra vonatkozó koordinátája. Ez a koordináta vektor az $\mathbb{F}^n$ koordinátatér egy eleme, a v vektor izomorf képe, ahol az izomorf leképezést az $\mathcal{X}$ bázis határozza meg, előírva, hogy az $\mathcal{X}$-beli x_i bázisvektor képe legyen az $e_i = [0, \dots, 0, 1, 0, \dots, \dots, 0]^*$, úgynevezett i-edik egységvektor, minden $(i = 1, \dots, n)$-re. Minthogy a v vektor koordinátái bázistól függőek, más bázis más izomorf leképezést határoz meg, következésképpen, más bázis esetén ugyanazon v vektornak mások lesznek a koordinátái. Az alábbiakban azt kívánjuk kideríteni, hogy a bázis megváltoztatásakor, a tér vektorainak koordinátái hogyan változnak. Itt most csak azzal az esettel foglalkozunk, amikor a bázist csak egyetlen vektor kicserélésével változtatjuk meg.

Legyen ezért $\mathcal{X} = \{x_1, \dots, x_n\}$ az eredeti bázis és azt tételezzük fel, hogy y a térnek egy nemzéró vektora. Egy olyan új bázisban akarjuk meghatározni a vektortér vektorainak koordinátáit, amely $\mathcal{X}$-től csak annyiban különbözik, hogy valamelyik vektora helyett az y vektor lesz az új bázisvektor.

y csakúgy, mint a tér bármelyik vektora, kifejezhető $\mathcal{X}$ vektorainak lineáris kombinációjaként,

$$y = \eta_1 x_1 + \cdots + \eta_i x_i + \cdots + \eta_n x_n,$$

ahol az y koordinátái között van nullától különböző, mondjuk $\eta_i \neq 0$, hiszen feltevésünk szerint y nemzéró vektor, és nyilvánvalóan csak nemzéró vektor lehet bázisnak eleme. Az η_i koordinátát *generáló elemnek* nevezzük. A $\eta_i \neq 0$ volta teszi lehetővé, hogy az x_i vektor kifejezhető az y és $\mathcal{X}$ többi vektorának lineáris kombinációjaként

$$x_i = -\frac{\eta_1}{\eta_i} x_1 - \cdots - \frac{\eta_{i-1}}{\eta_i} x_{i-1} + \frac{1}{\eta_i} y - \frac{\eta_{i+1}}{\eta_i} x_{i+1} - \cdots - \frac{\eta_n}{\eta_i} x_n, \tag{1.6}$$

amiből következik, hogy az $\mathcal{X}' = \{x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n\}$ vektorrendszer is bázis. Ezt igazolandó, először lássuk be, hogy $\mathcal{X}'$ generátorrendszer. Mivel $\mathcal{X}$ bázis volt és az elhagyott x_i vektor $\mathcal{X}'$ vektorainak lineáris kombinációja, ha egy V-beli v vektor $\mathcal{X}$ vektorainak lineáris kombinációjaként való előállításában az x_i vektort helyettesítjük az (1.6) egyenlet jobboldalával, akkor $\mathcal{X}'$ vektorainak lineáris kombinációjaként való kifejezését kapjuk. Másrészt $\mathcal{X}'$ lineárisan független vektorrendszer is, mert n elemű generátorrendszer és ha lineárisan összefüggő lenne, akkor a térnek lenne n-nél kevesebb vektort tartalmazó bázisa, ellentmondva az (1.7.5) tételnek. Tekintsük most a vektortér egy tetszőleges v elemét. Ha ez az $\mathcal{X}$ bázis vektorainak

$$v = \varepsilon_1 x_1 + \cdots + \varepsilon_i x_i + \cdots + \varepsilon_n x_n$$

lineáris kombinációja, akkor az x_i vektort helyettesítve az (1.6) vektoregyenlet jobboldalával, kapjuk, hogy

$$v \;=\; \varepsilon_1 x_1 + \cdots + \varepsilon_i \Big(-\frac{\eta_1}{\eta_i} x_1 - \cdots - \frac{\eta_{i-1}}{\eta_i} x_{i-1} + \frac{1}{\eta_i} y -$$

$$- \quad \frac{\eta_{i+1}}{\eta_i} x_{i+1} - \cdots - \frac{\eta_n}{\eta_i} x_n) + \cdots + \varepsilon_n x_n, \tag{1.7}$$

ami a bázisvektorok együtthatóinak rendezése után

$$v = (\varepsilon_1 - \eta_1 \frac{\varepsilon_i}{\eta_i}) x_1 + \cdots + \frac{\varepsilon_i}{\eta_i} y + \cdots (\varepsilon_n - \eta_n \frac{\xi_i}{\eta_i}) x_n \,. \tag{1.8}$$

Összefoglalva a fenti számítás eredményét, ha a v vektor $\mathcal{X}$ bázisra vonatkozó koordináta vektora $\mathbf{v}_{\mathcal{X}}$ és az $\mathbf{y}_{\mathcal{X}}$ koordináta vektorú nemzéró y vektorral kicseréljük az $\mathcal{X}$ bázis x_i elemét, hogy megkapjuk az $\mathcal{X}'$ új bázist, akkor $\mathbf{v}_{\mathcal{X}'}$ ugyanazon v vektornak az új $\mathcal{X}'$ bázisra vonatkozó koordináta vektora. Tehát ha

$$\mathbf{v}_{\mathcal{X}} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{i-1} \\ \varepsilon_i \\ \varepsilon_{i+1} \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad \text{és} \quad \mathbf{y}_{\mathcal{X}} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_{i-1} \\ \eta_i \\ \eta_{i+1} \\ \vdots \\ \eta_n \end{bmatrix} \quad \text{akkor} \quad \mathbf{v}_{\mathcal{X}'} = \begin{bmatrix} \varepsilon_1 - \eta_1 \frac{\varepsilon_i}{\eta_i} \\ \vdots \\ \varepsilon_{i-1} - \eta_{i-1} \frac{\varepsilon_i}{\eta_i} \\ \frac{\varepsilon_i}{\eta_i} \\ \varepsilon_{i+1} - \eta_{i+1} \frac{\varepsilon_i}{\eta_i} \\ \vdots \\ \varepsilon_n - \eta_n \frac{\varepsilon_i}{\eta_i} \end{bmatrix}.$$

Látható, hogy legelőször célszerű a bázisba éppen bevont új bázisvektorra vonatkozó koordináta meghatározása, hiszen ez minden más koordináta kiszámításában szerepet kap. Ezt a bázisból elhagyandó vektorra vonatkozó koordinátának a generáló elemmel való osztásával kapjuk. Ezt a hányadost δ-val jelölve azt mondhatjuk, hogy minden más új bázisra vonatkozó koordináta megkapható, ha a régi bázisra vonatkozó koordinátából kivonjuk az új bázisvektor megfelelő koordinátájának δ-szorosát.

Hangsúlyoznunk kell, hogy az y vektort olyan x_i bázisvektor helyett tudtuk bázisba vonni, amelyre vonatkozó koordinátája nem nulla, hiszen ez tette lehetővé, hogy a bázisból elhagyandó x_i vektort előállítsuk az y és a megmaradt $\mathcal{X}$-beli vektorok lineáris kombinációjaként az (1.6) egyenletnek megfelelően.

6 Példa. *Tegyük fel, hogy a 4-dimenziós valós V vektortérben az $\mathcal{X} = \{x_1, \ldots, x_4\}$ vektorrendszer egy bázis és $y = 2x_1 + x_2 - 3x_3 + x_4$. Az x_1 vektort ki akarjuk cserélni az y vektorral és meg kell határozni a $v = x_1 + 2x_2 + x_3 - 3x_4$ vektor koordináta vektorát az új $\mathcal{X}' = \{y, x_2, \ldots, x_4\}$ bázisban.*

A számításokat a

	y	**v**
x_1	$\boxed{2}$	1
x_2	1	2
x_3	-3	1
x_4	1	-3

táblázat módosításával végezzük, amelynek baloldali oszlopában az $\mathcal{X}$ bázis elemeit tüntettük fel, második oszlopában az y vektor e bázisra vonatkozó koordinátáit, ezt a második oszlop fejlécében az $\mathbf{y}$ feltüntetésével is jelöltük, a harmadik oszlop pedig a v $\mathcal{X}$ bázisra vonatkozó koordináta vektorát tartalmazza, ugyancsak az $\mathbf{v}$ jellel fejlécezve. Minthogy a táblázat baloldali oszlopa az $\mathcal{X}$ bázis vektorait mutatja, ebből egyértelmű, hogy a további oszlopok melyik bázisra vonatkozó koordináta vektorok. Ezért hagyhattuk el a bázisra utaló indexet a koordináta vektorok jeléből. Az y vektor első koordinátája be van keretezve, amellyel azt kívántuk jelezni, hogy az y vektort az első bázisvektor helyére akarjuk a bázisba bevonni. Ezt a koordinátát választjuk generáló elemnek. Újfent hangsúlyozzuk, hogy a generáló elem mindig zérustól különböző kell legyen, hiszen, amint azt az előzőekben láttuk, csak olyan bázisvektor cserélhető ki egy új vektorra, amely kifejezhető az új vektor és az eredeti bázis megmaradó vektorainak lineáris kombinációjaként. A következő táblázatban már a számítási eljárás van feltüntetve,

	$\mathbf{v}$		
y	$1/2$	$=$	$1/2$
x_2	$2 - 1 \cdot \frac{1}{2}$	$=$	$3/2$
x_3	$1 - (-3) \cdot \frac{1}{2}$	$=$	$5/2$
x_4	$-3 - 1 \cdot \frac{1}{2}$	$=$	$-7/2$

Természetesen a $\mathbf{v}$ koordináta vektort feladatok megoldásakor azonnal az

$$\begin{bmatrix} 1/2 \\ 3/2 \\ 5/2 \\ -7/2 \end{bmatrix}$$

alakban írjuk a táblázatba, itt csupán azt kívántuk megmutatni, hogy az új bázisra vonatkozó koordináták hogyan számíthatók ki. Azt állíthatjuk az új koordináta vektor ismeretében, hogy a v vektor az új bázisvektorok

$$v = \frac{1}{2}y + \frac{3}{2}x_2 + \frac{5}{2}x_3 - \frac{7}{2}x_4$$

lineáris kombinációja. Más szavakkal megfogalmazva ugyanezt, amíg az eredeti $\mathcal{X}$ bázis által meghatározott izomorfizmus a v vektorhoz $\mathbb{R}^4$-nek $[1, 2, 1, -3]$ elemét rendeli, addig az $\{y, x_2, x_3, x_4\}$ bázis által definiált másik izomorf leképezés ugyanezen v vektort az $[1/2, 3/2, 5/2, -7/2]$ szám-4-esbe viszi. □

Számos lineáris algebrai probléma megoldásához lehet használni az elemi bázistranszformáció módszerét, hogy vektorok koordináta vektorát egy alkalmas bázisra vonatkozóan meghatározzuk. Persze általában elemi bázistranszformációk sorozatával jutunk csak az alkalmas bázishoz. A következő részben néhány alkalmazási lehetőséget mutatunk be.

1.8.2 Az elemi bázistranszformáció néhány alkalmazása

Vektorrendszerek lineáris függetlenségének, illetve összefüggőségének vizsgálata

Az a feladat, hogy valamely V vektortér egy $\mathcal{Y} = \{y_1, \dots, y_n\}$ vektorrendszeréről el kell dönteni, hogy lineárisan független, vagy lineárisan összefüggő. Amennyiben ismerjük az $\mathcal{Y}$ vektorainak V valamely $\mathcal{X}$ bázisára vonatkozó koordináta vektorait, akkor elemi bázistranszformációk sorozatával az $\mathcal{X}$ bázis vektorait az $\mathcal{Y}$ vektorrendszer vektoraival cseréljük ki. Ha az $\mathcal{Y}$ vektorrendszer minden vektora bevonható a bázisba, kicserélve az $\mathcal{X}$-beli vektorokat, akkor egy bázis részrendszere lévén, lineárisan független. Ha azonban $\mathcal{Y}$ valamelyik y_i vektora nem cserélhető ki $\mathcal{X}$-beli vektorral, mert ezekre vonatkozó koordinátái mind nullák, akkor y_i vagy a zéró vektor, vagy az előzőleg már a bázisba bevont $\mathcal{Y}$-beli vektorok lineáris kombinációja, és akkor $\mathcal{Y}$ lineárisan összefüggő.

7 Példa. *Legyen a 4–dimenziós valós V vektortér egy bázisa $\mathcal{X} = \{x_1, \ldots, x_4\}$ és az $\mathcal{Y} = \{y_1, y_2, y_3\}$ vektorrendszer vektorainak ezen bázisra vonatkozó koordinátái legyenek*

$$\mathbf{y_1} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{y_2} = \begin{bmatrix} -1 \\ 2 \\ -2 \\ 3 \end{bmatrix} \quad \textit{és} \quad \mathbf{y_3} = \begin{bmatrix} 3 \\ -4 \\ 6 \\ -3 \end{bmatrix}$$

Állapítsuk meg, hogy $\mathcal{Y}$ lineárisan összefüggő, vagy lineárisan független vektorrendszer!

A számításokat az alábbi táblázatokban végeztük

	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$
x_1	$\boxed{1}$	-1	3
x_2	-1	2	-4
x_3	2	-2	6
x_4	0	3	-3

	$\mathbf{y}_2$	$\mathbf{y}_3$
y_1	-1	3
x_2	$\boxed{1}$	-1
x_3	0	0
x_4	3	-3

	$\mathbf{y}_3$
y_1	2
y_2	-1
x_3	0
x_4	0

Az utolsó táblázatból kiolvasható, hogy az y_3 vektor az előző lépések során a bázisba bevont y_1 és y_2 vektorok lineáris kombinációja, nevezetesen $y_3 = 2y_1 - 1y_2$, következésképpen az $\mathcal{Y}$ vektorrendszer lineárisan összefüggő. □

Kompatibilitás vizsgálat

Mindenekelőtt meg kell magyarázzuk, hogy mit kell azon érteni, hogy egy V vektortér valamely v vektora kompatibilis valamely $\mathcal{Y} = \{y_1, \ldots, y_n\}$ vektorrendszerre vonatkozóan. Az $\mathcal{Y}$ vektorrendszer lineáris burka, mint az már ismert, a V térnek egy altere, amelynek elemei éppen $\mathcal{Y}$ vektorainak lineáris kombinációi. A v vektort az $\mathcal{Y}$ vektorrendszerre vonatkozóan *kompatibilis*nek nevezzük, ha v benne van az $\mathcal{Y}$ lineáris burkában, azaz, ha v az $y_1, \ldots, y_n$ vektorok lineáris kombinációja.

8 Példa. *Állapítsuk meg, hogy az $\mathbb{R}_3[t]$ vektortér (a legfeljebb 3–adfokú valós együtthatós polinomok tere) $p(t) = -t + 3t^2 + 2t^3$ vektora kompatibilis–e a $q_1(t) = 1 - t$, $q_2(t) = 1 - t^2$, $q_3(t) = 1 + t^3$ vektorok rendszerére vonatkozóan!*

Nem nehéz belátni, hogy az $\mathbb{R}_3[t]$ vektortérnek az $\mathcal{A} = \{1,\ t,\ t^2,\ t^3\}$ vektorrendszere (polinomrendszere) bázis, így számításainkat végezhetjük az erre a bázisra vonatkozó koordináta vektorokkal. Persze az $\mathcal{A}$ bázis helyett $\mathbb{R}_3[t]$ bármely más bázisára vonatkozó koordináta vektorokkal is dolgozhatnánk, csak akkor a

$q_1(t), q_2(t), q_3(t)$ és $p(t)$ polinomok koordináta vektorainak meghatározása külön feladatot jelentene. Számításainkat az alábbi táblázatok mutatják:

	$\mathbf{q}_1(t)$	$\mathbf{q}_2(t)$	$\mathbf{q}_3(t)$	$\mathbf{p}(t)$
1	$\boxed{1}$	1	1	0
t	-1	0	0	-1
t^2	0	-1	0	3
t^3	0	0	1	2

$\longrightarrow$

	$\mathbf{q}_2(t)$	$\mathbf{q}_3(t)$	$\mathbf{p}(t)$
$q_1(t)$	1	1	0
t	$\boxed{1}$	1	-1
t^2	-1	0	3
t^3	0	1	2

$\longrightarrow$

	$\mathbf{q}_3(t)$	$\mathbf{p}(t)$
$q_1(t)$	0	1
$q_2(t)$	1	-1
t^2	1	2
t^3	$\boxed{1}$	2

$\longrightarrow$

	$\mathbf{p}(t)$
$q_1(t)$	1
$q_2(t)$	-3
t^2	0
$q_3(t)$	2

A legutolsó táblázatból kiolvasható, hogy a

$$p(t) = q_1(t) - 3q_2(t) + 2q_3(t),$$

és valóban $-t+3t^2+2t^3 = (1-t)-3(1-t^2)+2(1+t^3)$, ami tehát azt mutatja, hogy a $p(t)$ vektor kompatibilis a $\{q_1(t),\ q_2(t),\ q_3(t)\}$ vektorrendszerre vonatkozóan.
□

Következő példánkban megmutatjuk, hogy a lineáris egyenletrendszerek kompatibilitási problémaként is kezelhetők. Persze a lineáris egyenletrendszerek általános tárgyalására majd még később visszatérünk az alkalmazásokról szóló fejezetben, itt most csak azt szeretnénk érzékeltetni, hogy tulajdonképpen, már rendelkezik az olvasó azzal a technikával, amely egy lineáris egyenletrendszer megoldásához szükséges.

9 Példa. *A kislányom, Anna nagy pénzgyűjtő. Természetesen a papírpénzek mellett számos fém pénzérmét is összegyűjtött. A teljes kollekciója 48 db érmét tartalmaz, és értéke 160 Ft. Gyűjteményében legtöbb az 1 Ft–osok száma. 2 Ft–os viszont 10–zel kevesebb van, mint forintos. A 10 Ft–osok, 5 Ft–osok és 2 Ft–osok száma összesen is 5–tel kevesebb, mint a gyűjteményben lévő 1 Ft–osok összértéke. Azt tudjuk még, hogy 1–gyel több 10 Ft–osa van mint 20 Ft–osa. Állapítsa meg, hogy Anna pénzérme kollekciója milyen összetételű!*

A látszólag komplikált feladathoz egy nagyon is egyszerű matematikai modell rendelhető, amelynek megoldása valóban gyerekjáték. Legyen ugyanis az 1 Ft–osok, 2 Ft–osok, 5 Ft–osok, 10 Ft–osok és 20 Ft–osok egyelőre ismeretlen száma rendre: $\xi_1,\ \xi_2,\ \xi_3,\ \xi_4$, illetve ξ_5 darab. Akkor ezek az ismeretlenek a fenti információk szerint eleget tesznek a következő egyenleteknek.

$$\begin{aligned}
\xi_1 + \xi_2 + \xi_3 + \xi_4 + \xi_5 &= 48 \\
\xi_1 + 2\xi_2 + 5\xi_3 + 10\xi_4 + 20\xi_5 &= 160 \\
\xi_1 - \xi_2 &= 10 \\
\xi_1 - \xi_2 - \xi_3 - \xi_4 &= 5 \\
\xi_4 - \xi_5 &= 1
\end{aligned}$$

Ha az egyes ismeretlenek együtthatóit egy–egy oszlopvektorba gyűjtjük, csakúgy, mint az egyenletek jobboldalán lévő konstansokat, akkor a fenti egyenletrendszer a

$$\xi_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + \xi_2 \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \\ 0 \end{bmatrix} + \xi_3 \begin{bmatrix} 1 \\ 5 \\ 0 \\ -1 \\ 0 \end{bmatrix} + \xi_4 \begin{bmatrix} 1 \\ 10 \\ 0 \\ -1 \\ 1 \end{bmatrix} + \xi_5 \begin{bmatrix} 1 \\ 20 \\ 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 48 \\ 160 \\ 10 \\ 5 \\ 1 \end{bmatrix}$$

alakot ölti. Ezt a vektoregyenletet interpretálhatjuk úgy, hogy adott 5-komponensű oszlopvektorok olyan lineáris kombinációi keresendők, amelyek egy adott 5-komponensű oszlopvektort eredményeznek. A lineáris kombinációban szereplő együtthatókat kell meghatároznunk. Ilyen feladatot már oldottunk meg, ugyanis ha megvizsgáljuk, hogy az egyenletrendszer jobboldalán szereplő konstansok oszlopvektora kompatibilis-e az együtthatók oszlopai alkotta vektorrendszerre vonatkozóan, akkor, amennyiben igenlő a válasz rögtön megkapjuk, hogy a vektorrendszer vektorainak milyen lineáris kombinációja a konstansok oszlopvektora, nemleges válasz esetén pedig azt állíthatjuk, hogy az egyenletrendszernek nincs megoldása. Az a jogos kérdés merülhet fel, hogy az itt szereplő oszlopvektorok melyik vektortér vektorainak és azon vektortér melyik bázisára vonatkozó koordináta vektoraiként kezelendők. Az 5-dimenziós valós $\mathbb{R}^5$ térben az $\mathcal{E} = \{e_i = [0, \ldots, 0, 1, 0, \ldots, 0] \mid i = 1, \ldots, 5\}$, ahol tehát e_i az i-edik egységvektor, olyan bázis, amelyre vonatkozóan minden $\mathbb{R}^5$-beli vektor koordinátái és komponensei egyenlők. Célszerű tehát ezt a vektorteret választani, és ebben az egyenletrendszert kompatibilitási problémának tekinteni. Alább a számítások menetének érzékeltetésére feltüntettük az induló, a közbülső és megoldást szolgáltató elemi bázistranszformációs táblázatokat, amelyben $\mathbf{a}_1, \ldots, \mathbf{a}_5$-tel fejléceztük a $\xi_1, \ldots, \xi_5$ ismeretlenek együtthatóiból képzett oszlopokat, a konstansok oszlopát pedig $\mathbf{b}$ jelöli. Minden táblázatban bekereteztük a generáló elemet és persze a generáló elemeket igyekeztünk úgy megválasztani, hogy a számítások minél könnyebben legyenek elvégezhetők.

	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_4$	$\mathbf{a}_5$	$\mathbf{b}$
e_1	1	1	1	1	1	48
e_2	1	2	5	10	20	160
e_3	$\boxed{1}$	−1	0	0	0	10
e_4	1	−1	−1	−1	0	5
e_5	0	0	0	1	−1	1

$\longrightarrow$

	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_4$	$\mathbf{a}_5$	$\mathbf{b}$
e_1	2	1	1	1	38
e_2	3	5	10	20	150
a_1	−1	0	0	0	10
e_4	0	−1	$\boxed{-1}$	0	−5
e_5	0	0	1	−1	1

$\longrightarrow$

	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{a}_5$	$\mathbf{b}$
e_1	2	0	1	33
e_2	3	−5	20	100
a_1	−1	0	0	10
a_4	0	1	0	5
e_5	0	−1	$\boxed{-1}$	−4

$\longrightarrow$

	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{b}$
e_1	2	$\boxed{-1}$	29
e_2	3	−25	20
a_1	−1	0	10
a_4	0	1	5
a_5	0	1	4

$$\longrightarrow \begin{array}{c|c|c} & \mathbf{a}_2 & \mathbf{b} \\ \hline a_3 & -2 & -29 \\ e_2 & \boxed{\text{-47}} & -705 \\ a_1 & -1 & 10 \\ a_4 & 2 & 34 \\ a_5 & 2 & 33 \\ \hline \end{array} \longrightarrow \begin{array}{c|c} & \mathbf{b} \\ \hline a_3 & 1 \\ a_2 & 15 \\ a_1 & 25 \\ a_4 & 4 \\ a_5 & 3 \\ \hline \end{array}$$

A lineáris egyenletrendszernek egyetlen megoldása van, mert a jobboldalon szereplő konstansok b oszlopvektora az együtthatók oszlopvektorainak egyértelmű

$$25 \cdot a_1 + 15 \cdot a_2 + 1 \cdot a_3 + 4 \cdot a_4 + 3 \cdot a_5$$

lineáris kombinációja. Ezek alapján, Anna gyűjteménye 25 db 1 Ft–ost, 15 db 2 Ft–ost, 1 db 5 Ft–ost, 4 db 10 Ft–ost és 3 db 20 Ft–ost tartalmaz. □

2. Fejezet

Lineáris leképezések, transzformációk

A lineáris leképezések és transzformációk a vektorterek elméletének, mind az alkalmazások, mind matematikai szempontból, legérdekesebb és legfontosabb fejezete. Innen ered az oly sok alkalmazásban szerephez jutó mátrixaritmetika is.

2.1 A lineáris leképezések elemi tulajdonságai

Bevezetésként egy nagyon egyszerű problémát fogalmazunk meg, annak érzékeltetésére, hogy a lineáris leképezések fogalma a mindennapi élet feladatainak modellezése során, természetesen absztrakcióval keletkezett.

Egy üzem m különböző erőforrás felhasználásával n-féle terméket gyárt. Ismeretes, hogy a j-edik termék egy darabjának elkészítéséhez az i-edik erőforrásból α_{ij} egységnyire van szükség. Megállapítandó, hogy különböző termelési programok megvalósításához az egyes erőforrásokból mennyi szükséges. A probléma matematikai modellezése a következőképpen történhet. Minden termelési programhoz hozzárendelhető egy-egy $t \in \mathbb{R}^n$ vektor, amelynek j-edik komponense a j-edik termékből gyártandó mennyiség. Ugyancsak minden termelési programhoz tartozik egy erőforrás felhasználási $s \in \mathbb{R}^m$ vektor, amelynek i-edik komponense pedig az i-edik erőforrásból felhasznált mennyiséget mutatja. A különböző lehetséges t tervek és megvalósításukhoz szükséges s erőforrásvektorok között, legalábbis ha további mellékfeltételeket nem veszünk figyelembe, lineáris a kapcsolat,

$$s_i = \sum_{j=1}^{n} \alpha_{ij} t_j \quad (i = 1, \ldots, m),$$

amin azt értjük, hogy ha a t_1 terv megvalósításához s_1, a t_2 program megvalósításához pedig s_2 erőforrásfelhasználás tartozik, akkor az egyesített program azaz a $t_1 + t_2$ megvalósítása $s_1 + s_2$ erőforrás felhasználásával lehetséges. Ugyancsak, valamely t program β-szorosának megvalósítása β-szor annyi erőforrás felhasználásával lehetséges, mint a t terv teljesítése. Hasonló problémák matematikai absztrakciója vezet a vektorterek lineáris leképezéseinek, illetve transzformációinak a fogalmához.

2.1.1 Definíció. *Legyenek V és W ugyanazon $\mathbb{F}$ test feletti vektorterek. Egy $A : V \to W$ leképezést lineáris leképezésnek nevezünk, ha*

$$\forall x, y \in V : \forall \alpha, \beta \in \mathbb{F} : A(\alpha x + \beta y) = \alpha A(x) + \beta A(y)$$

teljesül. Ha $V = W$, akkor az $A : V \to V$ lineáris leképezést lineáris transzformációnak nevezzük.

A V vektorteret *tárgyvektortérnek*, míg a W-t *képvektortérnek* nevezzük. A definíció alapján mondhatjuk, hogy a lineáris transzformációk speciális, egy vektortérnek önmagába való lineáris leképezései, ezért ahol ez lehetséges, állításainkat lineáris leképezésekre bizonyítjuk.

2.1.1 Példák lineáris leképezésekre és transzformációkra

1. Ha $\Phi : V \rightleftharpoons W$ izomorf leképezés, akkor az izomorf leképezés definíciója alapján, eleget tesz a lineáris leképezésektől megkövetelt feltételeknek. Hangsúlyoznunk kell, hogy fordítva nem igaz, nem minden lineáris leképezés izomorfizmus.

2. Egy $\mathbb{F}$ test feletti V vektortér lineáris transzformációjára egyszerű példa az a leképezés, amit tetszőleges $\alpha \in \mathbb{F}$ skalár indukál az

   $$x \to \alpha x \quad (x \in V)$$

   megfeleltetéssel. Ez a vektortér axiómáiból következik.

3. Tekintsünk a 3-dimenziós valós térben, — amin értsük a térbeli, rögzített kezdőpontú helyvektorok terét — egy tetszőleges, origón átmenő egyenest és forgassunk el minden vektort ezen egyenes körül valamilyen rögzített ϕ szöggel. Ezt úgy kell végrehajtani, hogy a helyvektor minden pontját elforgatjuk az egyenes körül a rögzített ϕ szöggel. Ez a leképezés is lineáris transzformáció. (A figyelmes olvasónak igaza van, amikor megjegyzi, hogy nem tudja mit értsen ponton és mit egyenesen. Kérjük, hogy pontos definiálásukig gondoljon a köznapi értelemben használt térbeli pontra és egyenesre!)

4. Vegyük most a 3-dimenziós valós tér egy origón átmenő síkját és tükrözzünk minden vektort erre a síkra. Ez a leképezés is lineáris transzformáció. Az a leképezés, ami minden vektorhoz egy origón átmenő síkon való vetületét rendeli, ugyancsak lineáris transzformáció.

   Mindkét előző példában feltűnő lehetett, hogy a tér vektorait origón átmenő egyenes körül forgattuk, illetve origón átmenő síkra tükröztük. Ennek az az egyszerű oka, hogy a tér zéróeleme, az origó, fix kell maradjon bármely lineáris transzformációnál, illetve a zéró vektor képe zéró vektor minden lineáris leképezésnél.

5. Tekintsük most a valós együtthatós polinomok $\mathbb{R}[t]$ terét és minden $p(t) \in \mathbb{R}[t]$-re legyen $\frac{\mathrm{d}}{\mathrm{d\,t}}p(t)$ a deriváltja $p(t)$-nek. Analízisből jól tudjuk, hogy a $\mathrm{d}/\mathrm{d\,t}$ deriválási operáció is lineáris transzformáció, hiszen

   $$\frac{\mathrm{d}}{\mathrm{d\,t}}(\alpha f + \beta g) = \alpha \frac{\mathrm{d}}{\mathrm{d\,t}} f + \beta \frac{\mathrm{d}}{\mathrm{d\,t}} g\,.$$

Ebből következik, hogy igaz az az általánosabb állítás is, hogy tetszőleges [a ,b] intervallumon differenciálható függvények $D_{[a,b]}$ vektorterének a deriválás lineáris transzformációja.

6. Legyen a valós együtthatós polinomok $\mathbb{R}[t]$ vektorterének S az az önmagába való leképezése, amely minden $p(t) \in \mathbb{R}[t]$ polinomhoz, annak $\int_0^t p(x)\,dx$ integrál függvényét rendeli. Akkor S lineáris transzformációja $\mathbb{R}[t]$-nek, amint azt analízis tanulmányainkból ugyancsak jól tudjuk.

7. Legyen $I_{[a,b]}$ az [a, b] intervallumon Riemann–integrálható valós függvények vektortere és $\mathbb{R}$ a valós számok 1-dimenziós valós vektortere. Az az $\int : I_{[a,b]} \to \mathbb{R}$ hozzárendelés, amely minden $f \in I_{[a,b]}$ függvényhez annak $\int_a^b f(x)\,dx$ integrálját rendeli lineáris leképezés.

   Valóban, ha $f, g \in I_{[a,b]}$ és α, β tetszőleges valós számok, akkor
$$\int_a^b (\alpha f(x) + \beta g(x))\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx\,.$$

8. Egy vektortér duálisának, azaz a lineáris függvények vektorterének minden eleme, a lineáris függvények definíciója alapján, ugyancsak lineáris leképezés.

Megmutatjuk, hogy általában hogyan lehet lineáris leképezéseket definiálni.

2.1.2 Állítás. *Legyenek ezért V és W tetszőleges, ugyanazon $\mathbb{F}$ test feletti vektorterek és $\mathcal{X} = \{x_1, \ldots, x_n\}$ V-nek tetszőleges bázisa. Értelmezzük az $A : V \to W$ leképezést a következőképpen:*

Legyenek $A(x_1), \ldots, A(x_n)$ tetszőleges elemek W-ben és bármely $v \in V$ vektor $A(v)$ képét a következő eljárással határozzuk meg:

(a) *Előállítjuk v-t $\mathcal{X}$ vektorainak lineáris kombinációjaként:*
$$v = \varepsilon_1 x_1 + \ldots + \varepsilon_n x_n,$$

(b) *és az $A(v)$ képelemet az $A(x_1), \ldots, A(x_n) \in W$ vektorok ugyanazon együtthatókkal képzett lineáris kombinációjaként adjuk meg, azaz*
$$A(v) \overset{\text{def}}{=} \varepsilon_1 A(x_1) + \ldots + \varepsilon_n A(x_n).$$

Akkor az így értelmezett A leképezés V-nek W-be való lineáris leképezése.

Bizonyítás. Az értelmezett A hozzárendelés egyértelmű, mert minden $v \in V$ vektornak az $\mathcal{X}$ bázisra vonatkozó koordinátái egyértelműen meghatározottak. Így csak azt kell még megmutatnunk, hogy az A leképezés lineáris. Legyen $u \in V$ egy tetszőleges másik vektor, és tegyük fel, hogy $u = \varphi_1 x_1 + \cdots + \varphi_n x_n$. Tetszőleges α és β skalárokra,
$$\alpha u + \beta v = \sum_{i=1}^{n} (\alpha\varphi_i + \beta\varepsilon_i) x_i\,,$$

így ehhez a vektorhoz az A leképezés az

$$A(\alpha u+\beta v)=\sum_{i=1}^{n}(\alpha\varphi_i+\beta\varepsilon_i)A(x_i)$$

vektort rendeli. Az $A(u)=\varphi_1A(x_1)+\dots+\varphi_nA(x_n)$ és az $A(v)=\varepsilon_1A(x_1)+\ldots+\varepsilon_nA(x_n)$ vektorok α és β skalárokkal képzett lineáris kombinációja

$$\alpha A(u)+\beta A(v)=\sum_{i=1}^{n}(\alpha\varphi_i+\beta\varepsilon_i)A(x_i)\,,$$

tehát teljesül a

$$A(\alpha u+\beta v)=\alpha A(u)+\beta A(v)$$

linearitási feltétel. □

Vegyük észre, hogy minden lineáris leképezés ilyen, abban az értelemben, hogy a bázisvektorok képei már egyértelműen meghatározzák a leképezést.

2.1.2 Lineáris leképezések magtere és képtere

2.1.3 Definíció. *Minden $A:V\to W$ lineáris leképezéshez tartozik két halmaz, a* $\ker(A)$*-val jelölt magtér, és az* $\operatorname{im}(A)$*-val jelölt képtér. Ezeket formálisan a következőképpen adhatjuk meg:*

- $\ker(A)=\{x\in V\mid A(x)=0\}$
- $\operatorname{im}(A)=\{x'\in W\mid \exists\, x\in V: A(x)=x'\}$

Szavakkal megfogalmazva ugyanezt: az $A:V\to W$ lineáris leképezés magtere, a V vektortér azon elemeinek halmaza, amelyek a W vektortér zéróvektorára a képződnek, a képtér pedig a W azon elemeinek a halmaza, amelyek hozzá vannak rendelve V-beli vektorokhoz képekként. Fontosnak tartjuk hangsúlyozni, hogy a képtér általában nem azonos a W vektortérrel, hanem annak csak részhalmaza.

A következő állítás alátámasztja, hogy a mag*tér* és kép*tér* elnevezések indokoltak.

2.1.4 Állítás. *Ha $A:V\to W$ lineáris leképezés, akkor* $\ker(A)$ *altere V-nek és* $\operatorname{im}(A)$ *altere W-nek.*

Bizonyítás. Mindkét állítást a (1.5.3) állításra támaszkodva úgy bizonyítjuk, hogy megmutatjuk, hogy mind $\ker(A)$, mind $\operatorname{im}(A)$ zárt a lineáris kombináció képzésére nézve. Legyenek $u,v\in\ker(A)$ tetszőleges vektorok és $\alpha,\beta\in\mathbb{F}$ tetszőleges skalárok. Akkor a

$$A(\alpha u+\beta v)=\alpha A(u)+\beta A(v)=\alpha 0+\beta 0=0$$

számolás mutatja, hogy $\alpha u+\beta v\in\ker(A)$.

Ha, $s',t'\in\operatorname{im}(A)$ tetszőleges vektorok, akkor vannak olyan $s,t\in V$ vektorok, hogy $A(s)=s'$ és $A(t)=t'$. De akkor bármilyen $\alpha,\beta\in\mathbb{F}$ skalárok mellett

$$\alpha s'+\beta t'=\alpha A(s)+\beta A(t)=A(\alpha s+\beta t),$$

ami éppen azt mutatja, hogy $\alpha s' + \beta t'$ is képe valamilyen, nevezetesen az $\alpha s + \beta t \in V$ vektornak. Igy $\alpha s' + \beta t' \in \mathrm{im}\,(A)$, tehát $\mathrm{im}\,(A)$ is zárt a lineáris kombináció képzésre, következésképpen altér. □

Az előző állításból azonnal adódik, hogy egy V vektortér valamely A lineáris transzformációjának magtere is és képtere is altere V-nek.

2.1.5 Definíció. *Az A lineáris leképezés $\rho(A)$ rangján képterének dimenzióját, $\nu(A)$ defektusán, pedig magterének dimenzióját értjük.*

A következő tételben kapcsolatot teremtünk V, az úgynevezett tárgyvektortér dimenziója, a lineáris leképezés rangja és defektusa között.

2.1.6 Tétel. *Ha $A : V \to W$ lineáris leképezés, és V véges dimenziós, akkor $\dim V = \rho(A) + \nu(A)$.*

Bizonyítás. Mivel $\nu(A) = \dim \ker(A)$ és $\rho(A) = \dim \mathrm{im}\,(A)$, elegendő azt megmutatni, hogy ha $\ker(A)$ egy bázisa az $\{a_1, \ldots, a_r\}$ vektorrendszer és $\mathrm{im}\,(A)$ egy bázisa az $\{A(b_1), \ldots, A(b_s)\}$ vektorrendszer, akkor az $\{a_1, \ldots, a_r, b_1, \ldots, b_s\}$ vektorrendszer bázisa V-nek. E vektorrendszer lineáris függetlenségét bizonyítandó, legyen

$$\alpha_1 a_1 + \cdots + \alpha_r a_r + \beta_1 b_1 + \cdots + \beta_s b_s = 0. \tag{2.1}$$

Véve az egyenlőség mindkét oldalán lévő vektornak az A leképezés által meghatározott képét, figyelembe véve, hogy $a_i \in \ker(A)$ $\quad (i = 1, \ldots r)$-re, és, hogy a zéró vektor képe minden lineáris leképezés mellett a zéró vektor, kapjuk, hogy

$$\beta_1 A(b_1) + \cdots \beta_s A(b_s) = 0\,.$$

De ez csak úgy lehet, ha $\beta_1 = \ldots = \beta_s = 0$, mert hiszen $\{A(b_1), \ldots, A(b_s)\}$ bázisa, így lineárisan független vektorrendszere $\mathrm{im}\,(A)$-nak. Akkor viszont a (2.1) egyenlet már az egyszerűbb

$$\alpha_1 a_1 + \cdots + \alpha_r a_r = 0$$

alakú. Ebből már az is következik, hogy $\alpha_1 = \ldots = \alpha_r = 0$, mert az $\{a_1, \ldots, a_r\}$ bázisa, így lineárisan független vektorrendszere $\ker(A)$-nak.

Meg kell még mutatni azt is, hogy az $\{a_1, \ldots, a_r, b_1, \ldots, b_s\}$ vektorrendszer generálja V-et. Legyen ezért $v \in V$ egy tetszőleges vektor. Tekintve, hogy $A(v) \in \mathrm{im}\,(A)$, kapjuk, hogy

$$A(v) = \delta_1 A(b_1) + \cdots \delta_s A(b_s),$$

amiből, átrendezéssel, és kihasználva, hogy A lineáris leképezés az

$$A(v - (\delta_1 b_1 + \cdots + \delta_s b_s)) = 0$$

egyenlőséghez jutunk. Ebből az következik, hogy

$$v - (\delta_1 b_1 + \cdots + \delta_s b_s) \in \ker(A)$$

és így

$$v - (\delta_1 b_1 + \cdots + \delta_s b_s) = \gamma_1 a_1 + \cdots + \gamma_r a_r\,.$$

Innen az egyenlőség átrendezése után kapjuk, hogy

$$v = \gamma_1 a_1 + \cdots + \gamma_r a_r + \delta_1 b_1 + \cdots + \delta_s b_s\,,$$

tehát az $\{a_1, \ldots, a_r, b_1, \ldots, b_s\}$ vektorrendszer generátorrendszere is V-nek. □

A fenti bizonyításból kiderül, hogy ha az A a véges dimenziós V vektortérnek valamely W vektortérbe való lineáris leképezése, akkor van V-nek az A képterével izomorf altere. Ezért azt gondolhatnánk, hogy egy V vektortér megkapható bármely lineáris transzformációja magterének és képterének direktösszegeként. Ez azonban távolról sem igaz. Egyszerű ellenpélda erre a legfeljebb n-edfokú valós polinomok $\mathbb{R}_n[t]$ terének az a lineáris transzformációja, amely minden polinomhoz annak deriváltját rendeli. Ennek a transzformációnak a konstans polinomok alkotják a magterét, míg a képtere a legfeljebb $(n-1)$-edfokú polinomok tere. Minthogy egyetlen n-edfokú polinom sem kapható meg egy konstans polinom és egy legfeljebb $(n-1)$-edfokú polinom összegeként, a képtér és a magtér direktösszege nem egyenlő $\mathbb{R}_n[t]$-vel. Megmutatható azonban, hogy ha az A lineáris transzformáció magterének és képterének a nullvektoron kívül nincs közös eleme, akkor $V = \ker(A) \oplus \operatorname{im}(A)$.

2.2 Műveletek lineáris leképezésekkel

Ebben a részben értelmezzük lineáris leképezések összeadását és skalárral való szorzását, megmutatjuk, hogy maguk a lineáris leképezések is vektorteret alkotnak a fenti műveletekkel.

2.2.1 Lineáris leképezések összeadása és szorzása skalárral

Legyen V és W két, ugyanazon $\mathbb{F}$ test feletti vektortér és legyen $L(V, W)$ az összes V-nek W-be való lineáris leképezéseinek a halmaza. Értelmezzük két $A, B \in L(V, W)$ lineáris leképezés összegét az

$$\forall x \in V : (A + B)(x) \stackrel{\text{def}}{=} A(x) + B(x)$$

definiáló egyenlőséggel, és legyen tetszőleges $A \in L(V, W)$ lineáris leképezés $\alpha \in \mathbb{F}$ skalárral való szorzata az

$$\forall x \in V : (\alpha A)(x) \stackrel{\text{def}}{=} \alpha A(x)$$

egyenlőséggel adott.

2.2.1 Tétel. *A lineáris leképezések $L(V, W)$ halmaza $\mathbb{F}$ feletti vektortér a fent definiált összeadással és skalárral való szorzással.*

Bizonyítás. Először azt kell megmutatni, hogy a definiált összeadásra és skalárral való szorzásra vonatkozóan $L(V, W)$ valóban zárt, azaz két lineáris leképezés összege is és egy lineáris leképezés skalárral való szorzata is lineáris leképezés. Ha A és $B \in L(V, W)$, akkor az nyilvánvaló, hogy összegük is V-ből W-be való leképezés, csak azt kell tehát igazolnunk, hogy vektorok lineáris kombinációját a képvektorok ugyanazon skalárokkal képzett lineáris kombinációjába viszi. Ezt viszont a következő számolás verifikálja:

$$\forall v, w \in V : \forall \alpha, \beta \in \mathbb{F} : (A + B)(\alpha v + \beta w) =$$

$$= A(\alpha v + \beta w) + B(\alpha v + \beta w) = \alpha A(v) + \beta A(w) + \alpha B(v) + \beta B(w) =$$

$$= \alpha(A(v) + B(v)) + \beta(A(w) + B(w)) = \alpha(A + B)(v) + \beta(A + B)(w)$$

Teljesen hasonlóan ha $\gamma \in \mathbb{F}$ és $A \in L(V, W)$, akkor nyilván γA is V-ből W-be való leképezés, de linearitását még igazolnunk kell. Ezt az alábbi számolás mutatja:

$$\forall v, w \in V : \forall \alpha, \beta \in \mathbb{F} : (\gamma A)(\alpha v + \beta w) =$$

$$= \gamma A(\alpha v + \beta w) = \gamma(\alpha A(v) + \beta A(w)) = (\gamma\alpha)A(v) + (\gamma\beta)A(w) =$$

$$= \alpha(\gamma A(v)) + \beta(\gamma A(w)) = \alpha(\gamma A)(v) + \beta(\gamma A)(w)$$

Igazolnunk kell még, hogy a lineáris leképezések az összeadásra nézve kommutatív–csoportot alkotnak, illetve, hogy a skalárokkal való szorzás is eleget tesz a vektortér definíciójában megkövetelt négy tulajdonságnak.

Legyenek ezért $A, B, C \in L(V, W)$ tetszőleges lineáris leképezések és v bármelyik vektora V-nek. Akkor kihasználva, hogy $A(v)$, $B(v)$ és $C(v)$ W-beli vektorok, illetve a leképezések összeadásának definícióját, kapjuk, hogy

(a) $$[A + (B + C)](v) = A(v) + (B + C)(v) = A(v) + (B(v) + C(v)) =$$

$$= (A(v) + B(v)) + C(v) = (A + B)(v) + C(v) = [(A + B) + C](v)$$

(b) $$(A + B)(v) = A(v) + B(v) = B(v) + A(v) = (B + A)(v)$$

(c)Legyen $O \in L(V, W)$ az a leképezés, amelyre $\forall v \in V : O(v) = 0$ teljesül. Ez nyilvánvalóan lineáris leképezés és a leképezések összeadására nézve zéróelem.
(d) Tetszőleges $A \in L(V, W)$-re a $(-1)A \in L(V, W)$ pedig, ahol -1 az $\mathbb{F}$ test egységelemének ellentettje, az A leképezés additív inverze.

A skalárral való szorzás tulajdonságait ellenőrizendő legyenek $\alpha, \beta \in \mathbb{F}$ tetszőleges skalárok. Akkor bármely $v \in V$-re

(1) $$[\alpha(A + B)](v) = \alpha[(A + B)(v)] = \alpha[A(v) + B(v)] =$$

$$= \alpha A(v) + \alpha B(v) = (\alpha A)(v) + (\alpha B)(v) = [(\alpha A) + (\alpha B)](v)$$

(2) $$[(\alpha + \beta)A](v) = (\alpha + \beta)A(v) = \alpha A(v) + \beta A(v) = (\alpha A + \beta A)(v)$$

(3) $$[(\alpha\beta)A](v) = (\alpha\beta)A(v) = \alpha[\beta A(v)] = \alpha[(\beta A)(v)] = [\alpha(\beta A)](v)$$

(4) Végül, ha 1 az $\mathbb{F}$ test egységeleme, akkor $(1A)(v) = 1A(v) = A(v)$, amivel igazoltuk a skalárral való szorzástól elvárt négy tulajdonságot. □

2.2.2 Lineáris leképezések szorzása

Még egy lineáris leképezéseken értelmezett művelettel foglalkozunk ebben a részben, amit szorzásnak fogunk nevezni, bár a kompozíció elnevezés talán szerencsésebb lenne, mert a valós analízisben a függvények kompozíciójának megfelelően definiáljuk a leképezések szorzatát.

Legyen a $B \in L(V, W)$ és $A \in L(W, Z)$ lineáris leképezések szorzata az az AB-vel jelzett és V-ből Z-be képező hozzárendelés, amely a

$$\forall v \in V : (AB)(v) \stackrel{\text{def}}{=} A(B(v))$$

egyenlőséggel adott, tehát a $v \in V$ vektort előbb a B leképezés W-be viszi, majd a $B(v)$ képet az A leképezés a Z-be. A lineáris leképezések szorzata is lineáris leképezés. Ennek igazolására, legyenek $v, w \in V$ és $\alpha, \beta \in \mathbb{F}$ tetszőleges vektorok, illetve skalárok, akkor

$$(AB)(\alpha v + \beta w) = A(B(\alpha v + \beta w)) = A(\alpha B(v) + \beta B(w)) =$$

$$= \alpha A(B(v)) + \beta A(B(v)) = \alpha(AB)(v) + \beta(AB)(w),$$

és ezt kellett megmutatnunk.

A lineáris transzformációk fontosságát hangsúlyozandó, megjegyezzük, hogy egy V vektortér $L(V)$-vel jelölt, lineáris transzformációinak a halmaza, amely persze a (2.2.1) tétel értelmében maga is vektortér, zárt a fenti szorzás műveletre is, azaz bármely két $A, B \in L(V)$ lineáris transzformáció AB szorzata is $L(V)$-ben van. Ez lehetővé teszi lineáris transzformációk nemnegatív egész kitevős hatványainak értelmezését a következő rekurzív definícióval:

$$A \in L(V) : A^0 \stackrel{\text{def}}{=} I \qquad \text{és} \qquad A^n \stackrel{\text{def}}{=} A^{n-1}A \quad \text{ha} \quad n > 0,$$

ahol I a V vektortér identikus leképezését jelöli, azt, amely minden vektornak önmagát felelteti meg. I természetesen lineáris transzformáció és tetszőleges $A \in L(V)$-re $IA = AI = A$. Amint azt általánosabb esetben, nevezetesen tetszőleges függvények kompozíciójánál láttuk, a függvények kompozíciója nem kommutatív, de asszociatív. Így nem meglepő, hogy a lineáris transzformációk szorzása sem kommutatív, de asszociatív és az összeadásra vonatkozóan disztributív.

2.2.3 Lineáris transzformációk inverze

A függvények tanulmányozásakor azt láttuk, hogy amennyiben a függvény egy–egyértelmű megfeleltetés az értelmezési tartomány és az érték készlet elemei között, akkor és csak akkor nyilik lehetőségünk a függvény inverzének értelmezésére. Persze egy tetszőleges $A \in L(V, W)$ lineáris leképezés akkor és csak akkor egy–egyértelmű megfeleltetés V és W elemei között, ha izomorfizmus, amikor is a V és W vektorterek között nincs lényegi különbség, legalábbis lineáris algebrai szempontból.

2.2.2 Definíció. *Az $\mathbb{F}$ test feletti V vektortér egy $A \in L(V)$ lineáris transzformációját invertálhatónak mondjuk, ha kielégíti az alábbi két feltételt:*

(i) *ha $v, w \in V$-re $A(v) = A(w)$, akkor $v = w$,*

(ii) *minden* $v' \in V$*-hez létezik olyan* $v \in V$*, amelyre* $A(v) = v'$.

Az (i) feltétel azt fejezi ki, hogy egy invertálható lineáris transzformációnak injektívnek kell lennie, míg az (ii) feltétel azt írja elő, hogy a V vektortér minden eleme elő kell álljon képként, azaz a leképezés szürjektív kell legyen. Ezeket a feltételeket úgy is megfogalmazhattuk volna, hogy a $\ker(A)$, magtérnek a zérus altérnek kell lennie, és az $\operatorname{im}(A)$, képtér meg kell egyezzen V-vel. Valóban, ha $\ker(A)$ a zérus altér, akkor

$$v, w \in V : A(v) = A(w) \Longrightarrow A(v) - A(w) = A(v - w) = 0 \Longrightarrow$$

$$\Longrightarrow v - w \in \ker(A) \Longrightarrow v - w = 0 \leftrightarrow v = w.$$

Fordítva pedig, mivel $A(0) = 0$, bármely lineáris leképezés esetén, azonnal adódik, hogy az (i) feltételnek eleget tevő lineáris transzformációra $\ker(A) = \{0\}$ teljesül. Az (ii) feltétel $\operatorname{im}(A) = V$-vel való ekvivalenciája a képtér értelmezéséből azonnal adódik. A két feltétel véges dimenziós terek esetében egymással is ekvivalens.

2.2.3 Állítás. *Ha* V *véges dimenziós vektortér, akkor az* $A \in L(V)$ *lineáris transzformációra* $\ker(A) = \{0\}$ *akkor és csak akkor, ha* $\operatorname{im}(A) = V$.

Bizonyítás. A (2.1.6) tétel miatt, ha $\ker(A) = \{0\}$, akkor $\dim V = \rho(A) = \dim \operatorname{im}(A)$, és lineáris transzformációkról lévén szó, ez azt jelenti, hogy $\operatorname{im}(A)$ altere V-nek nem lehet valódi altér, az egész V vektortérrel kell megegyezzen. Fordítva, ha $\operatorname{im}(A) = V$, akkor megint a (2.1.6) tétel alapján kapjuk, hogy $\nu(A) = \dim \ker(A) = 0$, és ez éppen azt jelenti, hogy $\ker(A)$ a zérus altér. □

A fent igazolt állítás végtelen dimenziós terek lineáris transzformációira nem igaz. Egyszerű ellenpélda a valós együtthatós polinomok $\mathbb{R}[t]$ vektorterének az a D lineáris transzformációja, amely minden polinomhoz, annak deriváltját rendeli. Nyilvánvaló, hogy D magtere 1–dimenziós, hiszen minden konstans polinom deriváltja a zéró polinom, azaz $\ker(D)$ nem a zérus altér. Ugyanakkor bármely

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

polinom előáll nem is egy polinom deriváltjaként. Valóban tetszőleges $\beta \in \mathbb{R}$ mellett a

$$q(t) = \beta + \alpha_0 t + \frac{\alpha_1}{2} t^2 + \ldots + \frac{\alpha_n}{n+1} t^{n+1}$$

polinom deriváltja egyenlő $p(t)$-vel, amivel igazoltuk, hogy $\operatorname{im}(D) = \mathbb{R}[t]$ teljesül.

Az invertálható lineáris transzformációkat *reguláris* lineáris transzformációknak is szokás nevezni, míg a nem invertálható transzformációkat *szingulárisaknak* mondjuk. A V vektortér reguláris lineáris transzformációinak halmazát $R(V)$-vel fogjuk jelölni.

2.2.4 Definíció. *Ha* $A \in R(V)$ *invertálható lineáris transzformáció, akkor minden* $v' \in V$*-hez létzik egy és csak egy* $v \in V$ *vektor, amelyre* $A(v) = v'$*. Definiáljuk azt az* A^{-1}*-gyel jelölt leképezést, amely ehhez a* v' *vektorhoz éppen* v*-t rendeli, és nevezzük ezt az* A *lineáris transzformáció* inverzének.

Könnyen igazolhatjuk, hogy A^{-1} is lineáris transzformáció. Valóban tetszőleges $v', w' \in V$ vektorokhoz léteznek olyan egyértelműen meghatározott $v = A^{-1}(v')$ és $w = A^{-1}(w')$ vektorok V-ben, amelyekre $A(v) = v'$ és $A(w) = w'$. Akkor A linearitása folytán, tetszőleges $\alpha, \beta \in \mathbb{F}$ skalárok mellett

$$A(\alpha v + \beta w) = \alpha A(v) + \beta A(w) = \alpha v' + \beta w',$$

következésképpen

$$A^{-1}(\alpha v' + \beta w') = \alpha v + \beta w = \alpha A^{-1}(v') + \beta A^{-1}(w').$$

Példák invertálható lineáris transzformációkra

1. A V vektortér identikus leképezése, I nyilvánvalóan invertálható lineáris transzformáció, és $I^{-1} = I$.

2. A sík helyvektorainak terében az a transzformáció, amely minden vektorhoz valamely, az origón átmenő egyenesre vonatkozó tükörképét rendeli ugyancsak invertálható, és inverze önmaga.

3. Szintén invertálható a 3-dimenziós tér helyvektorainak valamely origón átmenő egyenes körüli, adott α szöggel való elforgatása és inverze a $2\pi - \alpha$ szöggel való elforgatás.

2.2.5 Állítás. *Ha V véges dimenziós vektortér és A lineáris transzformációja V-nek, akkor A invertálhatóságának szükséges és elegendő feltétele, hogy V egy tetszőleges $\mathcal{X} = \{x_1, \ldots, x_n\}$ bázisát bázisba transzformálja, azaz, ha az $\{A(x_1), \ldots, A(x_n)\}$ vektorrendszer is bázisa V-nek.*

Bizonyítás. A (2.2.3) állítás szerint, véges dimenziós V vektortér esetén egy A lineáris transzformáció pontosan akkor invertálható, ha $\operatorname{im}(A) = V$ teljesül. Ugyanakkor $\operatorname{im}(A)$-t generálja az $\{A(x_1), \ldots, A(x_n)\}$ vektorrendszer. □

Bármely $A \in R(V)$-re A^{-1} is invertálható, továbbá

$$\left(A^{-1}\right)^{-1} = A$$

és

$$A \cdot A^{-1} = A^{-1} \cdot A = I. \tag{2.2}$$

A (2.2) egyenletek akár a lineáris transzformációk invertálhatóságának definiálására is alkalmasak, mert igaz a következő:

2.2.6 Tétel. *Ha $A, B, C \in L(V)$ lineáris transzformációkra teljesül, hogy*

$$A \cdot B = C \cdot A = I,$$

akkor A invertálható és $A^{-1} = B = C$.

Bizonyítás. Az invertálhatóság (i) feltétele könnyen kapható, mert ha $v, w \in V$-re $A(v) = A(w)$, akkor

$$v = I(v) = (C \cdot A)(v) = C(A(v)) = C(A(w)) = (C \cdot A)(w) = I(w) = w,$$

és nem nehezebb a (ii) feltétel teljesülésének igazolása sem, hiszen bármely $v \in V$-re a $B(v) \in V$ vektor olyan, hogy

$$A(B(v)) = (A \cdot B)(v) = I(v) = v.$$

Ez biztosítja A^{-1} létezését. Másrészt tetszőleges $v \in V$-re, kihasználva a leképezések szorzásának asszociativitását kapjuk, hogy

$$C(v) = C(I(v)) = C(A \cdot B)(v) = (C \cdot A)B(v) = I(B(v)) = B(v),$$

tehát $B = C$. A (2.2) egyenletek biztosítják, hogy A^{-1} átveheti B, vagy C szerepét az előbbi számítás során és kapjuk, hogy $A^{-1} = C$, illetve $A^{-1} = B$. □

A (2.2.6) tétel bizonyításából kiderül, hogy az $A \in L(V)$ lineáris transzformáció injektivitását biztosítja olyan $C \in L(V)$ létezése, amelyre $C \cdot A = I$, míg a szürjektivitás feltétele, olyan $B \in L(V)$ létezése amelyre $A \cdot B = I$ áll fenn. Tekintettel arra, hogy véges dimenziós V vektortér esetén a lineáris transzformációk invertálhatóságának (i) és (ii) feltételei ekvivalensek, adódik a következő:

2.2.7 Következmény. *Ha V véges dimenziós vektortérnek A lineáris transzformációja, akkor az alábbi állítások egymással ekvivalensek:*

(i) *A invertálható,*

(ii) *létezik olyan $B \in L(V)$, hogy $A \cdot B = I$,*

(iii) *létezik olyan $C \in L(V)$, hogy $C \cdot A = I$.*

Végtelen dimenziós vektorterek esetén, mint állítottuk, mindkét egyenletnek kell legyen megoldása. Mutatja ezt a valós együtthatós polinomok $\mathbb{R}[t]$ terében a deriválási D illetve integrálfüggvény képző S lineáris transzformációk példája. Ezekre $DS = I$, de egyik transzformáció sem reguláris.

A lineáris transzformációk invertálhatóságáról befejezésül bebizonyítjuk a következő tételt:

2.2.8 Tétel. *Ha A és B invertálható lineáris transzformációi a V vektortérnek, akkor AB is invertálható és $(AB)^{-1} = B^{-1}A^{-1}$.*

Bizonyítás. A (2.2.6) tétel szerint elegendő megmutatni, hogy AB felcserélhető a $B^{-1}A^{-1}$ lineáris transzformációval és szorzatuk az identikus transzformáció. Valóban, a lineáris transzformációk szorzásának asszociativitását kihasználva kapjuk, hogy

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

és

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I\,.$$

□

A tétel következménye, hogy ha A invertálható lineáris transzformációja V-nek, akkor A^n is invertálható és $(A^n)^{-1} = \left(A^{-1}\right)^n$.

1. Gyakorlatok, feladatok

1. Értelmezzük a V vektortér egy önmagába való T leképezését a következőképpen: rögzítsük V-nek egy v_0 elemét és minden $v \in V$-re legyen $T(v) \stackrel{\text{def}}{=} v_0 + v$, azaz v-nek v_0-lal való *eltoltja.* Állapítsa meg, hogy lineáris transzformáció-e a T.

2. Legyen $A : \mathbb{F}^n \to \mathbb{F}^{n-1}$ az a leképezés, amelyre

$$A(v) = A\left(\begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_{n-1} \\ \xi_n \end{bmatrix}\right) = \begin{bmatrix} \xi_1 - \xi_2 \\ \xi_2 - \xi_3 \\ \vdots \\ \xi_{n-1} - \xi_n \end{bmatrix}$$

minden $v \in \mathbb{F}^n$-re.

(a) Mutassa meg, hogy A lineáris leképezés!

(b) Adja meg $\ker(A)$ egy bázisát!

(c) Határozza meg $\operatorname{im}(A)$ dimenzióját!

3. Legyen $A \in L(V, W)$ és M egy altere V-nek. Jelölje $A(M)$ az M-beli vektorok képeinek halmazát. Mutassa meg, hogy $A(M)$ altere $\operatorname{im}(A)$-nak.

4. Bizonyítsa be, hogy ha egy lineáris transzformációnak van legalább két különböző jobbinverze, akkor nincs balinverze, és hasonlóan, ha van legalább két különböző balinverze, akkor nincs jobbinverze.

5. Igazolja, hogy egy lineáris transzformációnak pontosan akkor van inverze, ha egyetlen egy jobbinverze van.

6. Terjesszük ki a jobbinverz, illetve balinverz fogalmakat lineáris leképezésekre. Legyen $A \in L(V, W)$ és $B, C \in L(W, V)$, ahol V és W nem feltétlenül izomorf vektorterek. Azt mondjuk, hogy B jobbinverze A-nak, iletve C balinverze A-nak, ha AB az identikus transzformációja W-nek, illetve CA az identikus transzformációja V-nek. Mutassa meg, hogy ha $\dim V \neq \dim W$, akkor A-nak nem lehet jobbinverze is, meg balinverze is!

7. Jelölje $O \in L(V)$ a V vektortér azon lineáris transzformációját, amely minden vektorhoz a zéróvektort rendeli. Mutassa meg, hogy $A, B \in L(V)$-re $AB = O$ akkor és csak akkor teljesül, ha $\operatorname{im}(B) \subseteq \ker(A)$!

2.2.4 Faktorterek**

Ha A a V vektortér egy lineáris leképezése valamely W vektortérbe, akkor A indukál egy $\equiv_A \subseteq V \times V$ relációt a következőképpen:

$$v_1 \equiv_A v_2 \iff A(v_1) = A(v_2)\,.$$

Az $\equiv_A$ nyilvánvalóan ekvivalencia reláció, de azontúl még rendelkezik azzal a tulajdonsággal is, hogy bármely $v \in V$ és α skalár mellett, ha

$$v_1 \equiv_A v_2 \text{ akkor } v_1 + v \equiv_A v_2 + v \quad \text{és} \quad \alpha v_1 \equiv_A \alpha v_2$$

is teljesül. Ezzel a tulajdonsággal rendelkező ekvivalencia relációkat *kongruencia relációknak* hívjuk. Mint ismeretes, az ekvivalencia relációk mindig meghatároznak egy osztályozást, azaz esetünkben a V vektortér részhalmazainak olyan $\{V_i\,, i \in I\}$ rendszerét, hogy a különböző részhalmazok diszjunktak és $\bigcup_{i\in I} V_i = V$. A kongruencia relációk még azt is lehetővé teszik, hogy az osztályok halmazát vektortérré tegyük az alábbi összeadással és skalárral való szorzással:

$$V_i + V_j \stackrel{\text{def}}{=} V_k \quad \text{ha} \quad v_i + v_j \in V_k, \text{ ahol } v_i \in V_i \quad \text{és} \quad v_j \in V_j\,,$$

és bármely α skalár esetén

$$\alpha V_i \stackrel{\text{def}}{=} V_j \quad \text{ha} \quad \alpha v_i \in V_j \text{ ahol } v_i \in V_i\,.$$

Talán első pillanatra nem nyilvánvaló, hogy a megadott műveletek jóldefiniáltak. Kihasználva, hogy az osztályozást kongruencia reláció indukálta, azt kapjuk, hogy ha $v_i, v_i' \in V_i$ és $v_j, v_j' \in V_j$, akkor

$$v_i \equiv_A v_i' \Rightarrow v_i + v_j \equiv_A v_i' + v_j$$

másrészt

$$v_j \equiv_A v_j' \Rightarrow v_i' + v_j \equiv_A v_i' + v_j'$$

és kihasználva az ekvivalencia reláció tranzitív tulajdonságát adódik, hogy

$$v_i + v_j \equiv_A v_i' + v'j\,,$$

ami azt mutatja, hogy egy osztály bármelyik elemének egy másik osztály bármelyik elemével való összege ugyanabban az osztályban van. Hasonlóan bármely α skalárra, ha v_i és v_i' egy osztálybeli vektorok, akkor αv_i és $\alpha v_i'$ is egy osztályban vannak. A $\equiv_A$ kongruencia reláció szerinti osztályok fentiekben definiált vektorterét V faktorterének nevezzük és $V/\equiv_A$-val jelöljük. A $\equiv_A$ kongruencia reláció értelmezéséből azonnal következik, hogy $V/\equiv_A$ izomorf az A lineáris leképezés $\operatorname{im}(A)$ képterével. Bármely osztály két vektorának a különbsége a zéróvektort tartalmazó osztályban, azaz az A lineáris leképezés magterében van, amiből következik, hogy a $V_i\,, i \in I$ osztály akármelyik vektora megkapható, ha egy rögzített v_{i0} vektorához hozzáadjuk $\ker(A)$ valamelyik vektorát. Formalizálva ugyanez:

$$v_{i0} + \ker(A) = \{v_{i0} + v \mid v \in \ker(A)\} = V_i\,.$$

Azt mondhatjuk, hogy minden osztály a lineáris leképezés magterének *eltolásával* kapható.

2.3 Mátrix reprezentáció

Ebben a részben véges dimenziós vektorterek lineáris leképezéseihez mátrixokat rendelünk, és megvizsgáljuk, hogy a lineáris leképezésekkel végzett műveleteknek milyen mátrixműveletek felelnek meg. Meghatározzuk a lineáris transzformáció és inverze mátrixának kapcsolatát, levezetjük az általános bázistranszformáció egyenletét, végül pedig megvizsgáljuk, hogy egy lineáris transzformáció mátrixa, hogyan függ a vektortér bázisának megválasztásától.

Megjegyezzük, hogy mint minden vektortér esetében, itt is a lineáris leképezések terének elemeihez is rendelhetnénk skalár komponensekből felépített oszlopokat, de mint látni fogjuk, a mátrixok használatának olyan előnyei vannak, amelyről nem szeretnénk lemondani.

Mint már említettük, egy $A \in L(V,W)$ lineáris leképezést teljesen meghatároz az, hogy V valamely bázisának vektorait milyen vektorokra képezi. Valóban, ha $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ egy bázisa V-nek, akkor bármely $v \in V$ vektor egyértelműen kifejezhető az $\mathcal{X}$ vektorainak

$$v = \varepsilon_1 x_1 + \varepsilon_2 x_2 + \cdots + \varepsilon_n x_n$$

lineáris kombinációjaként, és akkor

$$A(v) = \varepsilon_1 A(x_1) + \varepsilon_2 A(x_2) + \cdots \varepsilon_n A(x_n).$$

Mivel a képvektorok a W vektortérben vannak, előállíthatók a W egy $\mathcal{Y} = \{y_1, y_2, \ldots, y_m\}$ bázisa vektorainak lineáris kombinációiként. Igy, ha

$$A(x_j) = \alpha_{1j} y_1 + \alpha_{2j} y_2 + \cdots + \alpha_{mj} y_m \quad j = 1, \ldots, n,$$

akkor

$$\begin{aligned}
A(v) &= \varepsilon_1(\alpha_{11} y_1 + \alpha_{21} y_2 + \cdots + \alpha_{m1} y_m) + \\
&+ \varepsilon_2(\alpha_{12} y_1 + \alpha_{22} y_2 + \cdots + \alpha_{m2} y_m) + \\
&\vdots \\
&+ \varepsilon_n(\alpha_{1n} y_1 + \alpha_{2n} y_2 + \cdots + \alpha_{mn} y_m) = \\
&= (\varepsilon_1 \alpha_{11} + \varepsilon_2 \alpha_{12} + \cdots + \varepsilon_n \alpha_{1n}) y_1 + \\
&+ (\varepsilon_1 \alpha_{21} + \varepsilon_2 \alpha_{22} + \cdots + \varepsilon_n \alpha_{2n}) y_2 + \\
&\vdots \\
&+ (\varepsilon_1 \alpha_{m1} + \varepsilon_2 \alpha_{m2} + \cdots + \varepsilon_n \alpha_{mn}) y_m.
\end{aligned}$$

Az α_{ij} skalárokat a lineáris leképezés $\mathcal{X}$, $\mathcal{Y}$ bázispárra vonatkozó koordinátáinak nevezzük és az

$$\mathbf{A}_{\mathcal{Y}}^{\mathcal{X}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{bmatrix}$$

mátrixot az $A \in L(V,W)$ lineáris leképezés $\mathcal{X}$–$\mathcal{Y}$ bázispárra vonatkozó mátrixának mondjuk. A leképezés mátrixát, csakúgy, mint a vektorok koordináta vektorait, vastagon szedett betűvel jelöljük, a felső index jelöli a V, az úgynevezett tárgyvektortér alapul vett bázisát, míg az alsó indexszel utalunk a W, az úgynevezett képvektortérben rögzített bázisra. Időnként, a jelölés egyszerűsítése érdekében az alsó és felső indexeket, elhagyjuk, ha a szövegkörnyezetből kiderül, hogy melyik bázispárra vonatkoznak a szóbanforgó leképezés koordinátái.

Felhívjuk az olvasó figyelmét arra, hogy a leképezés mátrixát alkotó oszlopok éppen az $\mathcal{X}$ bázis vektorai képének $\mathcal{Y}$ bázisra vonatkozó

$$\mathbf{A}(x_1)_{\mathcal{Y}} = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{bmatrix}, \ \mathbf{A}(x_2)_{\mathcal{Y}} = \begin{bmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{bmatrix}, \dots, \mathbf{A}(x_n)_{\mathcal{Y}} = \begin{bmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{bmatrix}$$

koordináta vektorai. Meg kell említenünk, hogy amennyiben egy $A \in L(V)$ lineáris transzformációt koordinatizálunk, akkor V-nek csak egy bázisát rögzítjük, mondjuk $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$-et, és A mátrixa

$$\mathbf{A}_{\mathcal{X}} = [\mathbf{A}(x_1)_{\mathcal{X}}, \mathbf{A}(x_2)_{\mathcal{X}}, \dots, \mathbf{A}(x_n)_{\mathcal{X}}]$$

felépítésű lesz, és tekintve, hogy ebben a mátrixban minden oszlop n-elmű, ez egy n-edrendű kvadratikus mátrix.

Egy $v \in V$ vektor $A(v)$ képének $\mathcal{Y}$ bázisra vonatkozó koordináta vektora

$$\mathbf{A}(v)_{\mathcal{Y}} = \begin{bmatrix} \varepsilon_1\alpha_{11} + \varepsilon_2\alpha_{12} + \cdots + \varepsilon_n\alpha_{1n} \\ \varepsilon_1\alpha_{21} + \varepsilon_2\alpha_{22} + \cdots + \varepsilon_n\alpha_{2n} \\ \vdots \\ \varepsilon_1\alpha_{m1} + \varepsilon_2\alpha_{m2} + \cdots + \varepsilon_n\alpha_{mn} \end{bmatrix}, \tag{2.3}$$

vagy ugyanez mésképpen:

$$\mathbf{A}(v)_{\mathcal{Y}} = \varepsilon_1\mathbf{A}(x_1)_{\mathcal{Y}} + \varepsilon_2\mathbf{A}(x_2)_{\mathcal{Y}} + \cdots + \varepsilon_n\mathbf{A}(x_n)_{\mathcal{Y}}\,, \tag{2.4}$$

ahol az

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

oszlop éppen $\mathbf{v}_{\mathcal{X}}$, a v vektor $\mathcal{X}$ bázisra vonatkozó koordináta vektora. Mivel a koordináta vektorokat is kezelhetjük ugyanolyan mátrixként, mint a lineáris leképezésekhez rendelt mátrixokat, azok szorzásának értelmezése szerint a (2.3) egyenlet röviden így írható

$$\mathbf{A}(v)_{\mathcal{Y}} = \mathbf{A}_{\mathcal{Y}}^{\mathcal{X}} \cdot \mathbf{v}_{\mathcal{X}}, \tag{2.5}$$

amit a *lineáris leképezés egyenletének* nevezünk.

Ha A egy V vektortér lineáris transzformációja, akkor a (2.5) egyenlet kicsit egyszerűsödik. Az $A \in L(V)$ lineáris transzformáció egyenlete az $\mathcal{X}$ bázisban

$$\mathbf{A}(v)_{\mathcal{X}} = \mathbf{A}_{\mathcal{X}} \cdot \mathbf{v}_{\mathcal{X}} \tag{2.6}$$

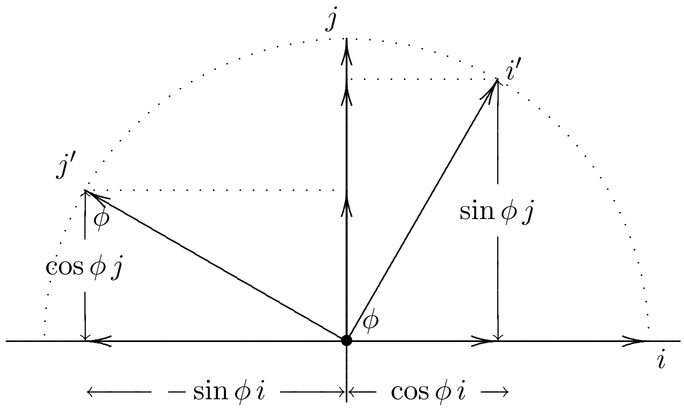
alakú lesz.

A (2.4) egyenletből következik, hogy az A lineáris leképezés $\operatorname{im}(A)$ képtere izomorf a mátrixát alkotó $\{\mathbf{A}(x_1)_{\mathcal{Y}}, \dots, \mathbf{A}(x_n)_{\mathcal{Y}}\}$ vektorrendszer lineáris burkával és ezért az A rangja egyenlő e vektorrendszer rangjával, amit az **A** *mátrix rangjának* is nevezünk és $\rho(\mathbf{A})$-val jelölünk.

1 Példa. *Legyen a 2-dimenziós tér egy bázisa az $\mathcal{I} = \{i, j\}$ egymásra merőleges, egységnyi hosszúságú helyvektorok rendszere. Határozzuk meg a tér azon F transzformációjának a mátrixát ebben a bázisban, amely minden vektort az origó körül ϕ szöggel elforgat az óramutató járásával ellentétes irányba.*

Az 2.1 ábráról leolvasható, hogy az i vektor i' képe, illetve a j vektor j' képe:

$$i' = \cos\phi\, i + \sin\phi\, j \quad \text{és} \quad j' = -\sin\phi\, i + \cos\phi\, j.$$



2.1. ábra: A sík forgatása

Így

$$\mathbf{F}(i)_{\mathcal{I}} = \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} \quad \text{és} \quad \mathbf{F}(j)_{\mathcal{I}} = \begin{bmatrix} -\sin\phi \\ \cos\phi \end{bmatrix}$$

tehát

$$\mathbf{F}_{\mathcal{I}} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

a transzformáció mátrixa az $\mathcal{I}$ bázisban. □

2 Példa. *Az $\mathbb{F}$ test feletti V vektortér egy $y^* : V \to \mathbb{F}$ lineáris függvényének, mint lineáris leképezésnek határozzuk meg a mátrixát a V vektortér $\mathcal{X} = \{x_1, \ldots, x_n\}$ és az $\mathbb{F}$, mint 1-dimenziós vektortér $\{1\}$ bázisára vonatkozóan, ahol 1 az $\mathbb{F}$ test egységelemét jelöli.*

Mivel $\mathbf{y}^{*\mathcal{X}}_{\{1\}} = [\mathbf{y}^*(\mathbf{x}_1)_{\{1\}}, \ldots, \mathbf{y}^*(\mathbf{x}_n)_{\{1\}}]$, így ha $y^*(x_i) = \xi_i$ minden $i(= 1, \ldots, n)$-re, akkor az y^* lineáris függvény mátrixa a választott bázispárra vonatkozóan az $1 \times n$ típusú

$$\mathbf{y}^{*\mathcal{X}}_{\{1\}} = [\xi_1, \ldots, \xi_n],$$

azaz egy sormátrix. Ez volt az oka annak, hogy egy V vektortér duálisát V^*-gal jelöltük. Mint láttuk, az a fordított állítás is igaz, hogy minden n komponensű $\mathbb{F}^{n*}$ sorvektor meghatározza V egy lineáris függvényét, mert a komponensekkel megadva a lineáris függvénynek az $\mathcal{X}$ bázis vektoraihoz rendelt skalárokat, a lineáris függvény egyértelműen meghatározott. $\square$

Legyen a V vektortér N lineáris transzformációja olyan, hogy egy $\mathcal{X} = \{x_1, \ldots, x_n\}$ bázisának vektorait az $N(x_1) = \lambda_1 x_1, \ldots, N(x_n) = \lambda_n x_n$ vektorokba viszi, ahol $\lambda_1, \ldots \lambda_n \in \mathbb{F}$ skalárok, ami valós vektortér esetében, geometriailag annyit jelent, hogy a transzformáció a koordinátarendszer tengelyeinek irányában nyújt, zsugorít és esetleg origóra tükröz. Akkor az N transzformáció mátrixa az $\mathcal{X}$ bázisban nagyon egyszerű

$$\mathbf{N} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$

diagonális mátrix.

2.3.1 A lineáris leképezésekkel és mátrixokkal végzett műveletek kapcsolata

Ebben a részben megmutatjuk, hogy milyen kapcsolat van a mátrixműveletek és a lineáris leképezések műveletei között. A mátrixműveletek értelmezésekor tulajdonképpen tekintettel voltunk arra, hogy a mátrixok lineáris leképezésekhez tartoznak, ez indokolja például a mátrixok szorzásának kicsit bonyolult értelmezését. Azt akartuk ugyanis, hogy a lineáris leképezések összegének mátrixa legyen egyenlő a tagok mátrixainak összegével, lineáris leképezés skalárszorosának mátrixa legyen a leképezés mátrixának skalárszorosa, és lineáris leképezések szorzatának mátrixa legyen a tényező leképezések mátrixainak szorzata.

Mátrixok és lineáris leképezések összeadásnak kapcsolata

Tekintsünk két $A, B \in L(V, W)$ lineáris leképezést, és tegyük fel, hogy V egy $\mathcal{X} = \{x_1, \ldots, x_n\}$ és W egy $\mathcal{Y} = \{y_1, \ldots, y_m\}$ bázisára vonatkozó mátrixaik az

$$\mathbf{A} = [\mathbf{A}(x_1)_\mathcal{Y}, \mathbf{A}(x_2)_\mathcal{Y}, \ldots, \mathbf{A}(x_n)_\mathcal{Y}] = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{bmatrix}$$

és

$$\mathbf{B} = [\mathbf{B}(x_1)_\mathcal{Y}, \mathbf{B}(x_2)_\mathcal{Y}, \ldots \mathbf{B}(x_n)_\mathcal{Y}] = \begin{bmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1} & \beta_{m2} & \ldots & \beta_{mn} \end{bmatrix}.$$

Az $A + B$ lineáris leképezéshez tartozó mátrix minden oszlopa ugyancsak megkapható, ha vesszük az $\mathcal{X}$ bázis vektorai $A + B$ leképezéssel kapott képeinek $\mathcal{Y}$ bázisra vonatkozó koordináta vektorait. Mivel a lineáris leképezések összegének definíciója szerint $(A + B)(v) = A(v) + B(v) \quad (v \in V)$ és mert az a hozzárendelés, amely egy vektortér vektoraihoz koordináta vektoraikat rendeli izomorfizmus, és így összegvektor koordináta vektora egyenlő a tagok koordináta vektorainak összegével, kapjuk, hogy

$$(\mathbf{A}+\mathbf{B})_{\mathcal{Y}}^{\mathcal{X}} = [\mathbf{A}(x_1)_{\mathcal{Y}}+\mathbf{B}(x_1)_{\mathcal{Y}}, \mathbf{A}(x_2)_{\mathcal{Y}}+\mathbf{B}(x_2)_{\mathcal{Y}}, \ldots, \mathbf{A}(x_n)_{\mathcal{Y}}+\mathbf{B}(x_n)_{\mathcal{Y}}] =$$

$$= \begin{bmatrix} \alpha_{11}+\beta_{11} & \alpha_{12}+\beta_{12} & \ldots & \alpha_{1n}+\beta_{1n} \\ \alpha_{21}+\beta_{21} & \alpha_{22}+\beta_{22} & \ldots & \alpha_{2n}+\beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1}+\beta_{m1} & \alpha_{m2}+\beta_{m2} & \ldots & \alpha_{mn}+\beta_{mn} \end{bmatrix} =$$

$$\mathbf{A}_{\mathcal{Y}}^{\mathcal{X}} + \mathbf{B}_{\mathcal{Y}}^{\mathcal{X}}\,.$$

A fentiek alapján azt mondhatjuk, hogy két lineáris leképezés összegének mátrixa megegyezik mátrixaik összegével.

Mátrixok és lineáris leképezések skalárral való szorzásának kapcsolata

A γA lineáris leképezés mátrixának oszlopait úgy kapjuk, hogy vennünk kell $\mathcal{X}$ bázis vektorainak γA képeit, és azoknak az $\mathcal{Y}$ bázisra vonatkozó koordináta vektorait. Mivel $(\gamma A)(v) = \gamma A(v)$ $(v \in V)$, és egy vektor skalárszorosának koordináta vektora egyenlő koordináta vektorának skalárszorosával, kapjuk, hogy

$$\gamma \mathbf{A}_{\mathcal{Y}}^{\mathcal{X}} = [\gamma\mathbf{A}(x_1)_{\mathcal{Y}}, \gamma\mathbf{A}(x_2)_{\mathcal{Y}}, \ldots, \gamma\mathbf{A}(x_n)_{\mathcal{Y}}] =$$

$$= \begin{bmatrix} \gamma\alpha_{11} & \gamma\alpha_{12} & \ldots & \gamma\alpha_{1n} \\ \gamma\alpha_{21} & \gamma\alpha_{22} & \ldots & \gamma\alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma\alpha_{m1} & \gamma\alpha_{m2} & \ldots & \gamma\alpha_{mn} \end{bmatrix} =$$

$$\gamma \cdot \mathbf{A}_{\mathcal{Y}}^{\mathcal{X}}\,.$$

Tehát egy lineáris leképezés skalárszorosának mátrixa egyenlő a lineáris leképezés mátrixának skalárszorosával.

Ha lerögzítjük mind V, mind W egy bázisát, akkor minden $A \in L(V, W)$ lineáris leképezéshez egyértelműen hozzárendelhető egy mátrix, amelynek elemei a vektorterek közös operátortartománya, az $\mathbb{F}$ testből való skalárok és típusa $\dim W \times \dim V$.

Fordítva, ha veszünk egy tetszőleges $\dim W \times \dim V$ típusú mátrixot, amelynek elemei $\mathbb{F}$-ből valók, egy $\mathcal{X}$ bázist V-ben, és egy $\mathcal{Y}$ bázist W-ben, akkor pontosan egy olyan lineáris leképezés van V-ből W-be, amelynek $\mathcal{X}$–$\mathcal{Y}$ bázispárra vonatkozó mátrixa éppen az adott mátrix. Ugyanis, az adott mátrix oszlopai, $\mathcal{X}$ vektorai képének koordináta vektorai az $\mathcal{Y}$ bázisban, tehát meghatározzák minden bázisvektor képét, és V bázisvektorainak képei egyértelműen meghatározzák V minden vektorának képét.

$\mathbb{F}_{m\times n}$-nel jelölve az összes $m\times n$ típusú $\mathbb{F}$ test feletti mátrixok vektorterét. Legyen $\mathbf{E}_{ij}$ az a mátrix, amelynek i-edik sorában a j-edik elem 1 (az $\mathbb{F}$ egységeleme), minden más elem pedig 0 (az $\mathbb{F}$ zéróeleme). Nem nehéz belátni, hogy az $\mathbb{F}_{m\times n}$ térben az

$$\mathcal{M} = \{\mathbf{E}_{ij} \;\; i = 1, \ldots, m; j = 1, \ldots, n\}$$

mátrixrendszer bázis. Ezért igaz az alábbi tétel.

2.3.1 Tétel. *Ha V n-dimenziós és W m-dimenziós $\mathbb{F}$ test feletti vektorterek, akkor az $L(V,W)$ lineáris leképezések tere izomorf az $\mathbb{F}$ fölötti mátrixok $\mathbb{F}_{m\times n}$ vektorterével, így $m \cdot n$-dimenziós.*

Mátrixok, illetve lineáris leképezések szorzásának kapcsolata

Legyenek $D \in L(V,W)$ és $C \in L(W,Z)$ lineáris leképezések és $\mathcal{X} = \{x_1, \dots, x_n\}$, $\mathcal{Y} = \{y_1, \dots, y_m\}$ és $\mathcal{Z} = \{z_1, \dots, z_p\}$ bázisok V-ben, W-ben, illetve Z-ben. A $CD \in L(V,Z)$ lineáris leképezés mátrixának oszlopai az $\mathcal{X}$ bázis vektorai CD leképezéssel kapott képeinek a $\mathcal{Z}$ bázisra vonatkozó koordináta vektorai. Jelölje a D leképezés mátrixát

$$\mathbf{D}_{\mathcal{Y}}^{\mathcal{X}} = \begin{bmatrix} \delta_{11} & \delta_{12} & \dots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \dots & \delta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{m1} & \delta_{m2} & \dots & \delta_{mn} \end{bmatrix},$$

ami azt jelenti, hogy

$$\begin{array}{rcccccccc} D(x_1) & = & \delta_{11}y_1 & + & \delta_{21}y_2 & + & \cdots & + & \delta_{m1}y_m \\ D(x_2) & = & \delta_{12}y_1 & + & \delta_{22}y_2 & + & \cdots & + & \delta_{m2}y_m \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots \\ D(x_n) & = & \delta_{1n}y_1 & + & \delta_{2n}y_2 & + & \cdots & + & \delta_{mn}y_m \end{array}$$

és a C leképezés mátrixát pedig

$$\mathbf{C}_{\mathcal{Z}}^{\mathcal{Y}} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{p1} & \gamma_{p2} & \cdots & \gamma_{pm} \end{bmatrix},$$

amiből pedig azt tudhatjuk, hogy

$$\begin{array}{rcccccccc} C(y_1) & = & \gamma_{11}z_1 & + & \gamma_{21}z_2 & + & \cdots & + & \gamma_{p1}z_p \\ C(y_2) & = & \gamma_{12}z_1 & + & \gamma_{22}z_2 & + & \cdots & + & \gamma_{p2}z_p \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots \\ C(y_m) & = & \gamma_{1m}z_1 & + & \gamma_{2m}z_2 & + & \cdots & + & \gamma_{pm}z_p \end{array}.$$

Akkor — felhasználva, hogy minden $v(\in V)$-re $(CD)(v) = C(D(v))$ — kapjuk, hogy

$$\begin{array}{rcccccccc} (CD)(x_1) & = & \delta_{11}C(y_1) & + & \delta_{21}C(y_2) & + & \cdots & + & \delta_{m1}C(y_m)\,, \\ (CD)(x_2) & = & \delta_{12}C(y_1) & + & \delta_{22}C(y_2) & + & \cdots & + & \delta_{m2}C(y_m)\,, \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots \\ (CD)(x_n) & = & \delta_{1n}C(y_1) & + & \delta_{2n}C(y_2) & + & \cdots & + & \delta_{mn}C(y_m)\,. \end{array}$$

A fenti egyenlőségekben minden $C(y_j)$-t a $\mathcal{Z}$ vektorainak fönt megadott lineáris kombinációival helyettesítjük:

$$\begin{array}{rclllllll}
(CD)(x_1) & = & \delta_{11}(\gamma_{11}z_1 & + & \gamma_{21}z_2 & + & \cdots & + & \gamma_{p1}z_p) \; + \\
 & + & \delta_{21}(\gamma_{12}z_1 & + & \gamma_{22}z_2 & + & \cdots & + & \gamma_{p2}z_p) \; + \\
 & & \vdots & & & & & & \\
 & + & \delta_{m1}(\gamma_{1m}z_1 & + & \gamma_{2m}z_2 & + & \cdots & + & \gamma_{pm}z_p) \; , \\
(CD)(x_2) & = & \delta_{12}(\gamma_{11}z_1 & + & \gamma_{21}z_2 & + & \cdots & + & \gamma_{p1}z_p) \; + \\
 & + & \delta_{22}(\gamma_{12}z_1 & + & \gamma_{22}z_2 & + & \cdots & + & \gamma_{p2}z_p) \; + \\
 & & \vdots & & & & & & \\
 & + & \delta_{m2}(\gamma_{1m}z_1 & + & \gamma_{2m}z_2 & + & \cdots & + & \gamma_{pm}z_p) \; , \\
\vdots & & \vdots & & \vdots & & \ddots & & \vdots \\
(CD)(x_n) & = & \delta_{1n}(\gamma_{11}z_1 & + & \gamma_{21}z_2 & + & \cdots & + & \gamma_{p1}z_p) \; + \\
 & + & \delta_{2n}(\gamma_{12}z_1 & + & \gamma_{22}z_2 & + & \cdots & + & \gamma_{p2}z_p) \; + \\
 & & \vdots & & & & & & \\
 & + & \delta_{mn}(\gamma_{1m}z_1 & + & \gamma_{2m}z_2 & + & \cdots & + & \gamma_{pm}z_p) \; .
\end{array}$$

Ebből már kiolvasható a CD lineáris leképezés mátrixa, csupán az egyes $\mathcal{Z}$-beli bázisvektorokra vonatkozó koordinátákat kell összegyűjteni és oszlopokba rendezni:

$$(\mathbf{CD})_{\mathcal{Z}}^{\mathcal{X}} = \begin{bmatrix} \sum_{k=1}^{m}\gamma_{1k}\delta_{k1}, & \sum_{k=1}^{m}\gamma_{1k}\delta_{k2}, & \ldots & , \sum_{k=1}^{m}\gamma_{1k}\delta_{kn} \\ \sum_{k=1}^{m}\gamma_{2k}\delta_{k1}, & \sum_{k=1}^{m}\gamma_{2k}\delta_{k2}, & \ldots & , \sum_{k=1}^{m}\gamma_{2k}\delta_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{m}\gamma_{pk}\delta_{k1}, & \sum_{k=1}^{m}\gamma_{pk}\delta_{k2}, & \ldots & , \sum_{k=1}^{m}\gamma_{pk}\delta_{kn} \end{bmatrix}.$$

Ezek szerint a lineáris leképezések szorzatának mátrixa

$$(\mathbf{CD})_{\mathcal{Z}}^{\mathcal{X}} = \mathbf{C}_{\mathcal{Z}}^{\mathcal{Y}} \cdot \mathbf{D}_{\mathcal{Y}}^{\mathcal{X}}$$

megegyezik a tényezők mátrixainak szorzatával.

Lineáris leképezések és transzformációk szorzásának tulajdonságai

A lineáris leképezések, illetve lineáris transzformációk szorzásának és mátrixaik szorzásának kapcsolata alapján a mátrixok szorzása tulajdonságainak ismeretében a következőt állíthatjuk.

2.3.2 Állítás. *(a) A lineáris leképezések és transzformációk szorzása nem kommutatív.*
(b) A lineáris leképezések és transzformációk szorzása asszociatív.
(c) A lineáris leképezések, illetve transzformációk szorzása azok összeadására vonatkozóan disztributív.

A mátrixok szorzására mutatott példáknál láttuk, hogy ha egy $\mathbf{A} \in \mathbb{F}_{m\times n}$ mátrixot szorzunk egy $\mathbf{x}(\in \mathbb{F}_{n\times 1})$ oszlopvektorral, akkor a szorzatmátrix típusa $m \times 1$, azaz egy oszlopvektor $\mathbb{F}_{m\times 1}$-ben. Ez az oszlopvektor az $\mathbf{A}$ mátrix oszlopainak lineáris kombinációja, amelynél a skalár együtthatók éppen az $\mathbf{x}$ oszlopvektor komponensei.

Formailag nincs különbség az $\mathbb{F}_{n\times 1}$-beli mátrixok és az $\mathbb{F}^n$ koordináta tér elemei között, ezért azonosíthatjuk azokat. Így értelmezhetjük $m \times n$ tipusú mátrixoknak $\mathbb{F}^n$-beli vektorokkal való szorzását. Akkor a mátrixok szorzásának tulajdonságai alapján állíthatjuk, hogy az

$$x \longrightarrow \mathbf{A}x \quad (x \in \mathbb{F}^n, \quad (\mathbf{A} \in \mathbb{F}_{m\times n})$$

hozzárendelés az $\mathbb{F}^n$ koordinátatérnek az $\mathbb{F}^m$ koordinátatérbe való A' lineáris leképezése. Ennek birtokában a lineáris leképezések reprezentációját kicsit más szemszögből is tekinthetjük:

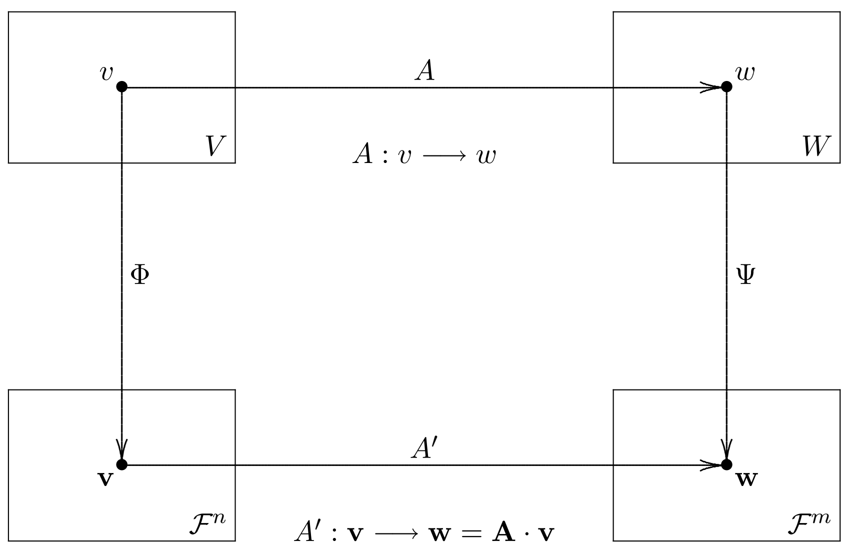
Legyen az $\mathbb{F}$ feletti n-dimenziós V vektortérnek az m-dimenziós W vektortérbe való $A \in L(V,W)$ lineáris leképezése olyan, amelynek mátrixa a V-beli $\mathcal{X} = \{x_1,\dots,x_n\}$ és W-beli $\mathcal{Y} = \{y_1,\dots,y_m\}$ bázispárra vonatkozóan $\mathbf{A}$. A

$$V \xrightarrow{A} W$$

lineáris leképezéssel "párhuzamosan", azt mintegy szimulálva, lejátszódik egy

$$\mathbb{F}^n \xrightarrow{A'} \mathbb{F}^m$$

lineáris leképezés a megfelelő koordinátaterek között, és az A' lineáris leképezés képvektorai, a tárgyvektorokból egyszerű mátrixszorzással kaphatók. Megmutatjuk, hogy a koordinátaterek közötti A' lineáris leképezés hű modellje az $A \in L(V,W)$ lineáris leképezésnek.



2.2. ábra: Lineáris leképezés "szimulációja" a koordináta terekben

Az $\mathcal{X}$ bázis meghatároz egy $\Phi : V \rightleftharpoons \mathbb{F}^n$ és az $\mathcal{Y}$ bázis pedig egy $\Psi : W \rightleftharpoons \mathbb{F}^m$ izomorf leképezést, amelyek V és W vektoraihoz koordináta vektoraikat rendeli. Mivel tetszőleges $v(\in V)$ vektorra

$$A'(\Phi(v)) = \Psi(A(v))$$

teljesül, az A' leképezés valóban az A leképezés hű mása. A 2.2 ábra jól szemlélteti az elmondottakat.

Ez teszi lehetővé, hogy bármely, lineáris leképezésekkel, vagy transzformációkkal kapcsolatos probléma numerikus megoldásakor a tárgy– illetve a képvektortér vektorai helyett, azok koordináta vektoraival számolhatunk, és a lineáris leképezések hatását, azok mátrixaival való szorzással számszerűsíthetjük.

2.3.2 Lineáris transzformációk inverzének mátrixa

Egy véges dimenziós V vektortér bármely $A \in L(V)$ lineáris transzformációjához úgy rendeltünk mátrixot, V egy $\mathcal{X} = \{x_1, \ldots, x_n\}$ bázisában, hogy meghatároztuk a bázisvektorok $\{A(x_1), \ldots, A(x_n)\}$ képeinek $\{\mathbf{A}(x_1)_\mathcal{X}, \ldots, \mathbf{A}(x_n)_\mathcal{X}\}$ koordináta vektorait és ezeket az oszlopokat gyűjtöttük össze az

$$\mathbf{A}_\mathcal{X} = [\mathbf{A}(x_1)_\mathcal{X}, \ldots, \mathbf{A}(x_n)_\mathcal{X}]$$

$n \times n$ tipusú mátrixban. Az $I \in L(V)$ identikus transzformáció mátrixa bármely bázisban ugyanolyan alakú, hiszen

$$\mathbf{I}(x_1)_\mathcal{X} = \mathbf{x}_{1\mathcal{X}} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{I}(x_n)_\mathcal{X} = \mathbf{x}_{n\mathcal{X}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

tehát

$$\mathbf{I}_\mathcal{X} = [\delta_{ij}], \text{ ahol } \delta_{ij} = \begin{cases} 1 & \text{ha} \quad i = j, \\ 0 & \text{ha} \quad i \neq j, \end{cases} \quad (i, j = 1, \ldots, n)$$

a már jól ismert Kronecker-szimbólum. Ezt a mátrixot $\mathbf{E}$-vel jelöltük és n-edrendű egységmátrixnak neveztük. Ha $A \in R(V)$ reguláris lineáris transzformáció, akkor $A^{-1} \in R(V)$, inverze kielégíti az

$$A \cdot A^{-1} = I \qquad \text{és} \qquad A^{-1} \cdot A = I$$

egyenleteket. Ebből, tekintve, hogy lineáris transzformációk szorzatának mátrixa egyenlő az egyes tényezők mátrixainak szorzatával, az következik, hogy az inverz transzformáció mátrixa, $\mathbf{A}_\mathcal{X}^{-1}$ kielégíti az

$$\mathbf{A}_\mathcal{X} \cdot \mathbf{A}_\mathcal{X}^{-1} = \mathbf{E} \qquad \text{és} \qquad \mathbf{A}_\mathcal{X}^{-1} \cdot \mathbf{A}_\mathcal{X} = \mathbf{E}$$

mátrixegyenleteket. A (2.2.3) következmény alapján azt is állíthatjuk, hogy ezen egyenletek bármelyike egyértelműen meghatározza $A \in R(V)$ inverzének mátrixát az $\mathcal{X}$ bázisban. Tehát reguláris lineáris transzformáció mátrixa invertálható a mátrixok szorzására vonatkozóan.

Ennek megfelelően most már egy $n \times n$ tipusú $\mathbf{A} \in \mathbb{F}_{n\times n}$ mátrixot is *regulárisnak* vagy más elnevezéssel *nemszingulárisnak* nevezünk, ha létezik megoldása az

$$\mathbf{AX} = \mathbf{E} \qquad \text{és} \qquad \mathbf{YA} = \mathbf{E}$$

mátrixegyenleteknek. Ha a mátrix nem invertálható, akkor a mátrixot is, csakúgy, mint a neminvertálható lineáris transzformációt *szingulárisnak* mondjuk. Sokszor egy mátrixról csupán azt kell eldönteni, hogy reguláris vagy szinguláris anélkül, hogy magára az inverzre (még ha létezik is) szükségünk lenne. Ilyenkor kihasználhatjuk, hogy a mátrixot alkotó oszlopok lineáris burka a mátrixhoz tartozó lineáris transzformáció képterével izomorf altere az $\mathbb{F}^n$ koordinátatérnek, ami persze akkor és csak akkor izomorf az egész vektortérrel, ha az oszlopok a koordinátatér bázisát alkotják, következésképpen lineárisan független vektorrendszert alkotnak. Meg kell jegyeznünk, hogy a (2.2.3) következmény szerint az $\mathbf{AX}=\mathbf{E}$ mátrixegyenlet akkor és csak akkor oldható meg, ha az $\mathbf{YA}=\mathbf{E}$ mátrixegyenletnek van megoldása, ezért bármelyik egyenlet használható lesz a reguláris mátrixok inverzének numerikus meghatározására.

2.4 Általános bázistranszformáció

Már láttuk, hogy egy vektortér elemeinek koordináta vektorai a tér bázisának megválasztásától függ, sőt az elemi bázistranszformációról szóló részben azt is megvizsgáltuk, hogy a tér vektorainak koordináta vektora hogyan változik meg, ha az új bázist az eredetiből egyetlen bázisvektornak egy új vektorral való kicserélésével kapjuk. Itt most két, tetszőlegesen választott bázisra vonatkozó koordináta vektorok kapcsolatát fogjuk jellemezni.

Legyen az $\mathbb{F}$ test feletti n-dimenziós V vektortérnek $\mathcal{X} = \{x_1, \ldots, x_n\}$ és $\mathcal{Y} = \{y_1, \ldots, y_n\}$ két bázisa és legyen egy tetszőleges $v \in V$ vektor koordináta vektora

$$\mathbf{v}_{\mathcal{X}} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \text{ az } \mathcal{X} \text{ bázisban és } \mathbf{v}_{\mathcal{Y}} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

az $\mathcal{Y}$ bázisban. Ez utóbbi pontosan azt jelenti, hogy

$$v = \epsilon_1 y_1 + \cdots + \epsilon_n y_n. \tag{2.7}$$

Mint tudjuk, a $\mathbf{v}_{\mathcal{X}}$ a v vektor izomorf képe azon $\Phi_{\mathcal{X}} : V \rightleftharpoons \mathbb{F}^n$ izomorf leképezés mellett, amely az $x_i \in \mathcal{X}$ bázisvektort az $e_i = [0, \ldots 0, 1, 0 \ldots, 0]$, úgynevezett i-edik egységvektorba viszi minden $i(= 1, \ldots, n)$-re, ezért a (2.7) egyenletből az is következik, hogy

$$\mathbf{v}_{\mathcal{X}} = \epsilon_1 \mathbf{y}_{1_{\mathcal{X}}} + \cdots + \epsilon_n \mathbf{y}_{n_{\mathcal{X}}}. \tag{2.8}$$

Gyűjtsük össze az $\mathbf{y}_{1_{\mathcal{X}}}, \ldots, \mathbf{y}_{n_{\mathcal{X}}}$ oszlopvektorokat a $\mathbf{B}$ mátrixba, amelyet a bázistranszformáció *átmenet mátrixának* nevezünk, azaz legyen az átmenet mátrix

$$\mathbf{B} = [\mathbf{y}_{1_{\mathcal{X}}}, \ldots, \mathbf{y}_{n_{\mathcal{X}}}].$$

A $\mathbf{B}$ átmenet mátrix azon $B \in L(V)$ lineáris transzformációnak a mátrixa az $\mathcal{X}$ bázisban, amely az $\mathcal{X}$ bázis elemeinek az $\mathcal{Y}$ bázis elemeit felelteti meg, egész pontosan amelyre $B(x_1) = y_1, \ldots, B(x_n) = y_n$. Ennek megfelelően a B transzformációt *bázistranszformációnak* nevezzük. Valóban a B transzformációnak a mátrixa az $\mathcal{X}$ bázisban

$$\mathbf{B}_{\mathcal{X}} = [\mathbf{B}(x_1)_{\mathcal{X}}, \ldots, \mathbf{B}(x_n)_{\mathcal{X}}] = [\mathbf{y}_{1_{\mathcal{X}}}, \ldots, \mathbf{y}_{n_{\mathcal{X}}}]$$

éppen az átmenet mátrix. Ezt felhasználva, a (2.8) egyenlet az egyszerűbb

$$\mathbf{v}_{\mathcal{X}} = \mathbf{B}_{\mathcal{X}} \cdot \mathbf{v}_{\mathcal{Y}} \tag{2.9}$$

alakot ölti. Mivel a B nyilvánvalóan reguláris lineáris transzformáció, mátrixának van inverze, így a (2.9) egyenlet ekvivalens a

$$\mathbf{B}_{\mathcal{X}}^{-1} \cdot \mathbf{v}_{\mathcal{X}} = \mathbf{v}_{\mathcal{Y}} \tag{2.10}$$

egyenlettel, amit a *bázistranszformáció egyenletének* nevezünk. A bázistranszformáció egyenletét úgy interpretálhatjuk, hogy amennyiben ismerjük a V vektortér egy v vektorának a koordináta vektorát az $\mathcal{X}$ bázisban és ismerjük egy új $\mathcal{Y}$ bázis vektorainak is a koordináta vektorait az $\mathcal{X}$ bázisban, (ezek a bázistranszformáció mátrixának, az átmenet mátrixnak az oszlopai), akkor a v vektor $\mathcal{Y}$ bázisra vonatkozó koordináta vektora megkapható, ha az átmenet mátrix inverzével szorozzuk a v $\mathcal{X}$ bázisra vonatkozó koordináta vektorát.

Persze a bázistranszformáció egyenletének inkább elméleti, mint gyakorlati jelentősége van, konkrét numerikus feladatok megoldásánál elemi bázistranszformációk sorozatán keresztül határozzuk meg egy vektor új bázisra vonatkozó koordináta vektorát.

2.4.1 Lineáris transzformáció mátrixa új bázisban

A bázistranszformáció egyenletének ismeretében nagyon könnyű kideríteni, hogy milyen kapcsolat van egy lineáris transzformáció két különböző bázisra vonatkozó mátrixai között.

Ezt a kapcsolatot leírandó, legyen A az n-dimenziós $\mathbb{F}$ test feletti V vektortér egy lineáris transzformációja, és jelölje $\mathbf{A}_{\mathcal{X}}$, illetve $\mathbf{A}_{\mathcal{Y}}$ az A mátrixát az $\mathcal{X} = \{x_1, \ldots, x_n\}$, illetve az $\mathcal{Y} = \{y_1, \ldots, y_n\}$ bázisban. Mint azt jól tudjuk,

$$\mathbf{A}_{\mathcal{X}} = [\mathbf{A}(x_1)_{\mathcal{X}}, \ldots, \mathbf{A}(x_n)_{\mathcal{X}}]$$

és

$$\mathbf{A}_{\mathcal{Y}} = [\mathbf{A}(y_1)_{\mathcal{Y}}, \ldots, \mathbf{A}(y_n)_{\mathcal{Y}}].$$

Jelölje megint B a V vektortér azon lineáris transzformációját, amelyre $B(x_1) = y_1, \ldots, B(x_n) = y_n$. Akkor $A(y_1) = A(B(x_1)), \ldots, A(y_n) = A(B(x_n))$, következésképpen a lineáris transzformációk (2.6) egyenlete alapján

$$\mathbf{A}(y_1)_{\mathcal{X}} = \mathbf{A}_{\mathcal{X}} \cdot \mathbf{B}(x_1)_{\mathcal{X}}, \ldots, \mathbf{A}(y_n)_{\mathcal{X}} = \mathbf{A}_{\mathcal{X}} \cdot \mathbf{B}(x_n)_{\mathcal{X}}.$$

Ebből az általános bázistranszformáció (2.10) egyenletét felhasználva kapjuk, hogy

$$\begin{aligned}
\mathbf{A}(y_1)_{\mathcal{Y}} &= \mathbf{B}_{\mathcal{X}}^{-1}\mathbf{A}(y_1)_{\mathcal{X}} = \mathbf{B}_{\mathcal{X}}^{-1}\mathbf{A}_{\mathcal{X}}\mathbf{B}(x_1)_{\mathcal{X}}\,, \\
&\;\vdots \\
\mathbf{A}(y_n)_{\mathcal{Y}} &= \mathbf{B}_{\mathcal{X}}^{-1}\mathbf{A}(y_n)_{\mathcal{X}} = \mathbf{B}_{\mathcal{X}}^{-1}\mathbf{A}_{\mathcal{X}}\mathbf{B}(x_n)_{\mathcal{X}}\,.
\end{aligned}$$

Vegyük észre, hogy $\mathbf{B}(x_1)_{\mathcal{X}}, \ldots, \mathbf{B}(x_n)_{\mathcal{X}}$ éppen a B lineáris transzformáció $\mathcal{X}$ bázisra vonatkozó $\mathbf{B}_{\mathcal{X}}$ mátrixának oszlopai, így eredményünk a mátrixaritmetikai ismereteinket is felhasználva a tömörebb

$$\mathbf{A}_{\mathcal{Y}} = \mathbf{B}_{\mathcal{X}}^{-1} \cdot \mathbf{A}_{\mathcal{X}} \cdot \mathbf{B}_{\mathcal{X}} \tag{2.11}$$

formába írható.

2.4.1 Definíció. *Azt mondjuk, hogy az* $\mathbf{A}_{\mathcal{Y}} = \mathbf{B}_{\mathcal{X}}^{-1} \cdot \mathbf{A}_{\mathcal{X}} \cdot \mathbf{B}_{\mathcal{X}}$ *mátrix az* $\mathbf{A}_{\mathcal{X}}$ *mátrixnak* konjugáltja*. Két mátrixot* hasonlónak *mondunk, ha az egyik a másiknak konjugáltja.*

Tehát egy lineáris transzformáció különböző bázisokra vonatkozó mátrixai hasonlók, amit $\mathbf{A}_{\mathcal{Y}} \sim \mathbf{A}_{\mathcal{X}}$-vel jelölünk.

Nem nehéz megmutatni, hogy a $\sim$ hasonlósági reláció, ekvivalencia reláció az azonos tipusú négyzetes mátrixok halmazán. Valóban,

(a) egy mátrixnak az identikus transzformációval való konjugáltja önmaga, tehát a $\sim$ reláció reflexív,

(b) ha

$$\mathbf{A} = \mathbf{B}^{-1}\mathbf{A}'\mathbf{B} \Longrightarrow \mathbf{A}' = \mathbf{B}\mathbf{A}\mathbf{B}^{-1},$$

mutatva, hogy $\sim$ szimmetrikus, és

(c) ha

$$\mathbf{A} = \mathbf{B}^{-1}\mathbf{A}'\mathbf{B} \quad \text{és} \quad \mathbf{A}' = \mathbf{C}^{-1}\mathbf{A}''\mathbf{C} \Longrightarrow$$

$$\Longrightarrow \mathbf{A} = \mathbf{B}^{-1}\mathbf{C}^{-1}\mathbf{A}''\mathbf{C}\mathbf{B} = (\mathbf{C}\mathbf{B})^{-1}\mathbf{A}''(\mathbf{C}\mathbf{B}),$$

igazolva, hogy $\sim$ reláció tranzitív is.

Reguláris lineáris transzformációk inverzének mátrixát könnyen meg tudjuk határozni a következő eredmény ismeretében:

2.4.2 Tétel. *Legyen a V vektortérnek $\mathcal{X} = \{x_1, \ldots, x_n\}$ és $\mathcal{Y} = \{y_1, \ldots, y_n\}$ két tetszőleges bázisa, és $B \in R(V)$ az a lineáris transzformációja, amelyre $B(x_i) = y_i$ minden $i(= 1, \ldots, n)$-re. Akkor teljesülnek az alábbi egyenlőségek:*

(i) $\mathbf{B}_{\mathcal{Y}} = \mathbf{B}_{\mathcal{X}}$,

(ii) $\mathbf{B}_{\mathcal{Y}}^{-1} = \mathbf{B}_{\mathcal{X}}^{-1}$,

Bizonyítás. A (2.11) egyenlet felhasználásával kapjuk, hogy

$$\mathbf{B}_{\mathcal{Y}} = \mathbf{B}_{\mathcal{X}}^{-1}\mathbf{B}_{\mathcal{X}}\mathbf{B}_{\mathcal{X}} = \mathbf{E}\mathbf{B}_{\mathcal{X}} = \mathbf{B}_{\mathcal{X}},$$

és teljesen hasonlóan

$$\mathbf{B}_{\mathcal{Y}}^{-1} = \mathbf{B}_{\mathcal{X}}^{-1}\mathbf{B}_{\mathcal{X}}^{-1}\mathbf{B}_{\mathcal{X}} = \mathbf{B}_{\mathcal{X}}^{-1}\mathbf{E} = \mathbf{B}_{\mathcal{X}}^{-1}.$$

□

3 Példa. *Legyen a V valós vektortér egy bázisa az $\mathcal{X} = \{x_1, x_2, x_3\}$ vektorrendszer és az A lineáris transzformáció mátrixa ebben a bázisban legyen*

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

Határozzuk meg az A mátrixát az $\mathcal{Y} = \{y_1, y_2, y_3\}$ bázisban, ahol

$$\begin{array}{rcrcrcr} y_1 &=& x_1 &+& 2x_2 &-& x_3, \\ y_2 &=& & & x_2 &+& x_3, \\ y_3 &=& x_1 &+& x_2. & & \end{array}$$

A bázistranszformáció

$$\mathbf{B} = [\mathbf{y}_{1\mathcal{X}}, \mathbf{y}_{2\mathcal{X}}, \mathbf{y}_{3\mathcal{X}}] = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

átmenet mátrixa adott $\mathcal{Y}$ vektorainak megadásával, így használhatnánk az $\mathbf{A}_{\mathcal{Y}} = \mathbf{B}^{-1}\mathbf{A}\mathbf{B}$ képletet a transzformáció mátrixának meghatározására az új $\mathcal{Y}$ bázisban, ehhez azonban invertálnunk kellene az átmenet mátrixot és elvégezni két mátrixszorzást. Egy kis munkát megspórolhatunk, ugyanis az $A(y_i)$ $(i = 1, 2, 3)$ vektorok $\mathcal{X}$ bázisra vonatkozó $\mathbf{A}(y_i)_{\mathcal{X}} = \mathbf{A} \cdot \mathbf{y}_{i\mathcal{X}}$ $(i = 1, 2, 3)$ koordináta vektorai könnyen meghatározhatók és akkor elemi bázistranszformációk sorozatával kiszámíthatók az $\mathcal{Y}$ bázisra vonatkozó koordináta vektoraik is, amelyek éppen a keresett $\mathbf{A}_{\mathcal{Y}}$ mátrix oszlopai. Ezeket a számításokat követheti nyomon az olvasó az alábbi táblázatok tanulmányozásával: (a táblázatok fejlécében a bázisra utaló indexeket elhagytuk, a táblázatok baloldali oszlopában ugyis fel vannak tüntetve a bázisvektorok)

	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$	$\mathbf{A}(\mathbf{y}_1)$	$\mathbf{A}(\mathbf{y}_2)$	$\mathbf{A}(\mathbf{y}_3)$
x_1	$\boxed{1}$	0	1	0	1	1
x_2	2	1	1	3	0	1
x_3	−1	1	0	−4	0	−2

↓

	$\mathbf{y}_2$	$\mathbf{y}_3$	$\mathbf{A}(\mathbf{y}_1)$	$\mathbf{A}(\mathbf{y}_2)$	$\mathbf{A}(\mathbf{y}_3)$
y_1	0	1	0	1	1
x_2	$\boxed{1}$	−1	3	−2	−1
x_3	1	1	−4	1	−1

↓

	$\mathbf{y}_3$	$\mathbf{A}(\mathbf{y}_1)$	$\mathbf{A}(\mathbf{y}_2)$	$\mathbf{A}(\mathbf{y}_3)$
y_1	1	0	1	1
y_2	−1	3	−2	−1
x_3	$\boxed{2}$	−7	3	0

↓

	$\mathbf{A}(\mathbf{y}_1)$	$\mathbf{A}(\mathbf{y}_2)$	$\mathbf{A}(\mathbf{y}_3)$
y_1	7/2	−1/2	1
y_2	−1/2	−1/2	−1
y_3	−7/2	3/2	0

Az utolsó táblázatból már kiolvasható:

$$\mathbf{A}_{\mathcal{Y}} = \begin{bmatrix} 7/2 & -1/2 & 1 \\ -1/2 & -1/2 & -1 \\ -7/2 & 3/2 & 0 \end{bmatrix}$$

Első ránézésre nem volt nyilvánvaló, hogy a megadott $\mathcal{Y}$ vektorrendszer valóban bázis, azonban az eljárás során ez is kiderült, hiszen $\mathcal{X}$ vektorait lépésről lépésre ki lehetett cserélni $\mathcal{Y}$ vektoraival. □

4 Példa. *Legyen a V valós vektortér egy bázisa az $\mathcal{X} = \{x_1, x_2, x_3\}$ vektorrendszer, és az A lineáris transzformáció mátrixa ebben a bázisban legyen*

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

Állapítsuk meg, hogy A reguláris–e, és ha igen, akkor határozzuk meg az A^{-1} inverz transzformáció mátrixát!

Az A mátrixából kiolvashatjuk, hogy

$$\begin{aligned} A(x_1) &= y_1 = x_1 - x_3, \\ A(x_2) &= y_2 = x_2 - x_3, \\ A(x_3) &= y_3 = x_1 - x_2 + x_3. \end{aligned}$$

Az inverz transzformációra persze $A^{-1}(y_i) = x_i$ $(i = 1,2,3)$ teljesül, amiből kiolvashatjuk az $\mathbf{A}^{-1}(y_i)_{\mathcal{X}}$ $(i = 1,2,3)$ koordináta vektorokat is. Ezekből elemi bázistranszformációk sorozatával meghatározhatók az $\mathbf{A}^{-1}(y_i)_{\mathcal{Y}}$ $(i = 1,2,3)$ koordináta vektorok, és nekünk éppen ezekre van szükségünk. A számításokat az alábbi táblázatokban végeztük:

	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$	$\mathbf{A}(\mathbf{y}_1)$	$\mathbf{A}(\mathbf{y}_2)$	$\mathbf{A}(\mathbf{y}_3)$
x_1	$\boxed{1}$	0	1	1	0	0
x_2	0	1	−1	0	1	0
x_3	−1	−1	1	0	0	1

↓

	$\mathbf{y}_2$	$\mathbf{y}_3$	$\mathbf{A}(\mathbf{y}_1)$	$\mathbf{A}(\mathbf{y}_2)$	$\mathbf{A}(\mathbf{y}_3)$
y_1	0	1	1	0	0
x_2	$\boxed{1}$	−1	0	1	0
x_3	−1	2	1	0	1

↓

	$\mathbf{y}_3$	$\mathbf{A}(\mathbf{y}_1)$	$\mathbf{A}(\mathbf{y}_2)$	$\mathbf{A}(\mathbf{y}_3)$
y_1	1	1	0	0
y_2	−1	0	1	0
x_3	$\boxed{1}$	1	1	1

$$\begin{array}{c|rrr|}
 & \mathbf{A}(\mathbf{y}_1) & \mathbf{A}(\mathbf{y}_2) & \mathbf{A}(\mathbf{y}_3) \\
 & \downarrow & & \\
\hline
y_1 & 0 & -1 & -1 \\
y_2 & 1 & 2 & 1 \\
y_3 & 1 & 1 & 1 \\
\hline
\end{array}$$

Az utolsó táblázatból már kiolvasható:

$$\mathbf{A}_{\mathcal{Y}}^{-1} = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

ami a (2.4.2) tétel szerint ugyanaz, mint $\mathbf{A}_{\mathcal{X}}^{-1}$. Persze, amint az előző példában is, csak a számítások elvégzése közben derült ki, hogy A reguláris mátrix, hogy oszlopai valóban egy bázis különböző vektorainak koordináta vektorai. $\square$

2. Gyakorlatok, feladatok

1. Legyen A a 3-dimenziós valós tér azon lineáris transzformációja, amely minden helyvektorhoz az origóra, mint középpontra vonatkozó tükörképét rendeli. Határozza meg az A mátrixát egy tetszőlegesen választott bázisban!

2. Legyen T a 3-dimenziós valós tér azon lineáris transzformációja, amely minden helyvektorhoz egy nemzéró v vektor irányú tengelyre vonatkozó tükörképét rendeli.

   (a) Határozza meg a T mátrixát egy, a v vektort tartalmazó, de azontúl tetszőleges bázisban!

   (b) Határozza meg a T^{-1} transzformációt és annak mátrixát az (a) feladat megoldásakor választott bázisban!

   (Megjegyzés: Legyenek x és y a bázis további vektorai és $x' = T(x)$, $y' = T(y)$. Nyilván x, x' és v valamint y, y' és v egy síkban vannak, továbbá a tükrözés miatt v az $x + x'$ vektornak is és az $y + y'$ vektornak is skalárszorosa.)

3. Legyen $\{u, v, w\}$ bázisa a 3-dimenziós valós térnek, és rendelje a P lineáris transzformáció minden $x = \alpha u + \beta v + \gamma w$ helyvektorhoz az $x' = \alpha u + \beta v$ vektort, (amit az x vektor u és v síkjába eső *vetületének* nevezünk).

   (a) Határozza meg a P transzformáció mátrixát előbb az $\{u, v, w\}$, majd az $\{u + w, v + w, u + v\}$ bázisokban!

   (b) Számitsa ki a P^n transzformáció mátrixát az előbbi bázisokban ($n = 2, \ldots, 100$)-ra!

   (c) Döntse el, hogy P reguláris, vagy szinguláris transzformáció-e!

4. Legyen $\{u, v, w\}$ a 3-dimenziós valós vektortér egy bázisa, és rendelje az A leképezés minden $x = \alpha u + \beta v + \gamma w$ helyvektorhoz az $x' = (\alpha + \gamma)u + (\beta + \gamma)v + \gamma w$ vektort.

(a) Mutassa meg, hogy A lineáris transzformáció!

(b) Határozza meg a transzformáció mátrixát előbb az $\{u, v, w\}$, majd az $\{u + v, v + w, w\}$ bázisokban!

(c) Állapítsa meg, hogy létezik–e A-nak inverze!

2.5 Mátrixok bázisfaktorizációja

A numerikus alkalmazások, így például a lineáris egyenletrendszerek megoldásakor segítségünkre lesz a mátrixok bázisfaktorizációja.

Legyen $\mathbf{A} \in \mathbb{F}_{m\times n}$, azaz egy olyan $m \times n$ tipusú mátrix, amelynek elemei az $\mathbb{F}$ testből valók, és tételezzük fel, hogy $\mathbf{A}$ oszlopvektorainak rendszere r rangú, azaz oszlopvektorrendszere az $\mathbb{F}^m$ koordinátatérnek r-dimenziós alterét generálja. Ezt az alteret az $\mathbf{A}$ mátrix oszlopvektorterének nevezzük. Emlékezzünk arra, hogy annak a lineáris leképezésnek, amely minden $x \in \mathbb{F}^n$ vektorhoz az $\mathbf{A}x \in \mathbb{F}^m$ vektort rendeli, a képtere éppen az $\mathbf{A}$ oszlopvektortere, és annak dimenziója az $\mathbf{A}$ mátrix $\rho(\mathbf{A})$ rangja. Ha $\mathbf{A}$ oszlopvektorai közül kiválasztunk egy r elemű lineárisan független vektorrendszert, akkor az, $\mathbf{A}$ oszlopvektorterének egy bázisa. Ezért az oszlopvektortér minden vektora, így az $\mathbf{A}$ mátrix oszlopai is előállíthatók ezeknek a lineáris kombinációjaként. Jelölje az $\mathbf{A}$ oszlopait $a_1, \dots, a_n$ és tegyük fel hogy ennek a vektorrendszernek a $\mathcal{B} = \{a_{i1}, \dots, a_{ir}\}$ egy lineárisan független részrendszere. Akkor ez bázis $\mathbf{A}$ oszlopvektorterében, mert feltevésünk szerint $\rho(\mathbf{A}) = r$. Legyenek az $a_j \quad (j = 1, \dots, n)$ oszlopvektoroknak az $\mathcal{B}$ bázisra vonatkozó koordináta vektorai rendre $\mathbf{a}_{j\mathcal{B}} \quad (j = 1, \dots, n)$. Gyűjtsük össze a $\mathcal{B} = \{a_{i1}, \dots, a_{ir}\}$ bázis vektorait a $\mathbf{B}$ mátrixba, azaz legyen

$$\mathbf{B} = [a_{i1}, \dots, a_{ir}]$$

és az $\mathbf{a}_{j\mathcal{B}} \quad (j = 1, \dots, n)$ koordináta vektorokat pedig a $\mathbf{C}$ mátrixba, tehát legyen

$$\mathbf{C} = [\mathbf{a}_{1\mathcal{B}}, \dots, \mathbf{a}_{n\mathcal{B}}] \, .$$

A $\mathbf{B} \cdot \mathbf{C}$ mátrix oszlopai

$$\mathbf{B}\mathbf{a}_{1\mathcal{B}} = a_1, \dots, \mathbf{B}\mathbf{a}_{n\mathcal{B}} = a_n \, ,$$

hiszen, mint azt a mátrixok szorzására adott példáknál láttuk, egy mátrixnak egy oszlopmátrixszal való szorzata, a mátrix oszlopainak olyan lineáris kombinációja, amelyben a skalár együtthatók, éppen az oszlopmátrix komponensei és az $\mathbf{a}_{j\mathcal{B}}$ oszlopvektor éppen az a_j vektornak a $\mathbf{B}$ oszlopaira vonatkozó koordinátáit tartalmazta minden $j(= 1, \dots, n)$-re. A mátrixok szorzását illusztráló példák között azt is láttuk, hogy két mátrix szorzatának i-edik oszlopa, a baloldali mátrixtényezőnek a jobboldali mátrixszorzó i-edik oszlopával való szorzata. Mindezek figyelembevételével kapjuk, hogy az $\mathbf{A}$ mátrix a $\mathbf{B}$ és $\mathbf{C}$ mátrixok szorzata, tehát

$$\mathbf{A} = \mathbf{B} \cdot \mathbf{C}. \tag{2.12}$$

Az $\mathbf{A}$ mátrix (2.12) egyenlet szerinti szorzatként való előállítását *bázisfaktorizációnak* hívjuk.

A 2.12 egyenlőségből az is következik, hogy az $\mathbf{A}$ mátrix minden sora a $\mathbf{C}$ mátrix sorainak lineáris kombinációja. Mivel a $\mathbf{B}$ tipusa $m \times r$ és a $\mathbf{C}$ tipusa pedig $r \times n$, azonnal kapjuk, hogy $\mathbf{A}$ sorvektorai rendszerének legfeljebb r lehet a rangja.

A fenti szorzatra bontást elvégezve az $\mathbf{A}$ mátrix transzponáltjára $\mathbf{A}^*$-ra, arra a következtetésre juthatunk, hogy az $\mathbf{A}^*$ sorvektorrendszerének rangja nem lehet nagyobb oszlopvektorrendszerének rangjánál. Nyilvánvaló, hogy az $\mathbf{A}$ mátrix sorvektorai által generált vektortér — $\mathbf{A}$ *sorvektorterének* nevezzük — és az $\mathbf{A}^*$ mátrix oszlopvektortere izomorfak. Ezeket az észrevételeket összegezve állíthatjuk, hogy

2.5.1 Állítás. *[Mátrixok rangszám tétele] Az* $\mathbf{A}$ *mátrix sorvektorrendszerének rangja egyenlő oszlopvektorrendszerének rangjával.*

Bemutatunk egy numerikus példát annak illusztrálására, hogy elemi bázistranszformáció alkalmazásával, hogyan lehet egy adott mátrixot bázisfaktorizálni.

5 Példa. *Bázisfaktorizáljuk az*

$$
\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 & 3 & 1 \\ 2 & -1 & 1 & 2 & 3 \\ 1 & 7 & -1 & 7 & 0 \\ -1 & -7 & 1 & -7 & 0 \end{bmatrix}
$$

mátrixot.

Meg kell határoznunk az $\mathbf{A}$ mátrix oszlopvektorrendszerének rangját. Ezt az elemi bázistranszformáció módszerével végezzük. Az induló táblában az $\mathbf{A}$ oszlopait $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$, és $\mathbf{a}_5$ jelölik, ezek megegyeznek az $\mathbb{F}^m$ tér $\mathcal{E} = \{e_1, e_2, e_3, e_4\}$ egységvektorai alkotta bázisra vonatkozó koordináta vektoraikkal. Ez a magyarázata annak, hogy az alábbi táblázatok fejlécében vastagon szedett betűkkel jelöltük azokat.

$$
\begin{array}{c|rrrrr}
 & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 \\ \hline
e_1 & \boxed{1} & 2 & 0 & 3 & 1 \\
e_2 & 2 & -1 & 1 & 2 & 3 \\
e_3 & 1 & 7 & -1 & 7 & 0 \\
e_4 & -1 & -7 & 1 & -7 & 0
\end{array}
\rightarrow
\begin{array}{c|rrrr}
 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 \\ \hline
a_1 & 2 & 0 & 3 & 1 \\
e_2 & -5 & \boxed{1} & -4 & 1 \\
e_3 & 5 & -1 & 4 & -1 \\
e_4 & -5 & 1 & -4 & 1
\end{array}
\rightarrow
$$

$$
\rightarrow
\begin{array}{c|rrr}
 & \mathbf{a}_2 & \mathbf{a}_4 & \mathbf{a}_5 \\ \hline
a_1 & 2 & 3 & 1 \\
a_3 & -5 & -4 & 1 \\
e_3 & 0 & 0 & 0 \\
e_4 & 0 & 0 & 0
\end{array}
$$

Amint az utolsó táblázatból látható, $\rho(\mathbf{A}) = 2$, és az $\mathbf{A}$ mátrix első és harmadik oszlopa bázis $\mathbf{A}$ oszlopvektorterében. Az is kiolvasható a táblázatból, hogy melyek a mátrix oszlopainak erre a bázisra vonatkozó koordináta vektorai. Nyilván $a_1 = 1 \cdot a_1 + 0 \cdot a_3$ és $a_3 = 0 \cdot a_1 + 1 \cdot a_3$,
míg a többi oszlop koordinátáit a táblázatból olvasva

$$
a_2 = 2 \cdot a_1 - 5 \cdot a_3, a_4 = 3 \cdot a_1 - 4 \cdot a_3 \quad \text{és} \quad a_5 = 1 \cdot a_1 + 1 \cdot a_3 \,.
$$

Ezek szerint az **A** bázisfaktorizációjának tényezői

$$\mathbf{B} = [a_1, a_3] = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}$$

és

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & 0 & 3 & 1 \\ 0 & -5 & 1 & -4 & 1 \end{bmatrix}.$$

Annak ellenőrzését, hogy valóban $\mathbf{A} = \mathbf{B} \cdot \mathbf{C}$ bárki gyorsan elvégezheti. □

3. Gyakorlatok, feladatok

1. Igazoljuk, hogy ha $\mathbf{A} = \mathbf{B} \cdot \mathbf{C}$, akkor

$$\rho(\mathbf{A}) \leq \min(\rho(\mathbf{B}), \rho(\mathbf{C}))$$

2. Mutassuk meg, hogy ha $\mathbf{A} = \mathbf{B} \cdot \mathbf{C}$ és **B** oszlopai lineárisan független rendszert alkotnak, akkor $\rho(\mathbf{A}) = \rho(\mathbf{C})$ és ha **C** sorai függetlenek, akkor pedig $\rho(\mathbf{A}) = \rho(\mathbf{B})$.

3. Bizonyítsuk be, hogy ha $\mathbf{A} = \mathbf{B} \cdot \mathbf{C}$, akkor

$$\mathbf{A}^* = \mathbf{C}^* \cdot \mathbf{B}^* .$$

3. Fejezet

Alkalmazások

Ebben a részben a vektorterek elméletének néhány nevezetes alkalmazásával ismerkedünk meg. Tárgyaljuk a lineáris egyenletrendszerek megoldhatóságának kérdéseit és az elemi bázistranszformációra támaszkodva megoldási algoritmust is adunk. Vizsgáljuk bizonyos tipusú mátrixegyenletek megoldhatóságát is, és belátjuk, hogy ezek megoldása visszavezethető lineáris egyenletrendszerek megoldására. Ezt követően ismertetünk egy módszert mátrixok inverzének viszonylag gyors meghatározására.

3.1 Lineáris egyenletrendszerek

Egy m egyenletből álló n ismeretlenes lineáris egyenletrendszer általános alakja:

$$\begin{array}{ccccccccc} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & + & \cdots & + & \alpha_{1n}\xi_n & = & \beta_1 \\ \alpha_{21}\xi_1 & + & \alpha_{22}\xi_2 & + & \cdots & + & \alpha_{2n}\xi_n & = & \beta_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ \alpha_{m1}\xi_1 & + & \alpha_{m2}\xi_2 & + & \cdots & + & \alpha_{mn}\xi_n & = & \beta_m \end{array}$$

ahol az ξ_i-k az ismeretlenek, míg az α_{ij} együtthatók és az egyenletek jobboladalán álló β_i konstansok adott, valamely számtestből való skalárok. A lineáris egyenletrendszert *inhomogén*nek nevezzük, ha a jobboldalon szereplő skalárok között van nemzéró, míg ellenkező esetben, ha az egyenletek jobboldalán lévő skalárok mindegyike nulla, akkor az egyenletrendszert *homogén*nek mondjuk. Természetesen a legtöbb esetben gyakorlati problémák modellezésekor fellépő lineáris egyenletrendszerek esetében ezek a skalárok valós számok, de a megoldás módszerét illetően ez nem jelent különösebb könnyebbséget, hacsak azt nem vesszük figyelembe, hogy a valós számokkal való számolásban nagyobb gyakorlatunk van. Mindenesetre mi ezen skalárok testét $\mathbb{F}$-fel fogjuk jelölni, hangsúlyozandó, hogy az bármely test lehet.

Az egyes ismeretlenek együtthatóit és a jobboldali konstansokat oszlopokba rendezve a

$$\xi_1 \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{bmatrix} + \xi_2 \begin{bmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{bmatrix} + \cdots \xi_n \begin{bmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}$$

vektoregyenletet kapjuk, ami az

$$a_1 = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{bmatrix}, \quad a_2 = \begin{bmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{bmatrix}, \dots, a_n = \begin{bmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{bmatrix} \quad \text{és} \quad b = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

jelölések bevezetése után, egyszerűen

$$\xi_1 a_1 + \xi_2 a_2 + \cdots + \xi_n a_n = b$$

formát ölt. Ha még az $a_1, a_2, \dots, a_n$ oszlopokat az $\mathbf{A}$ mátrixba gyűjtjük össze, és az ismeretleneket pedig az x oszlopba, tehát az

$$\mathbf{A} = [a_1, a_2, \dots, a_n] \qquad \text{és} \qquad x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$$

jelölések bevezetése után, az egyenletrendszer röviden

$$\mathbf{A}x = b \tag{1}$$

alakban írható fel.

Az egyenletrendszerek megoldhatóságának kritériuma többféleképpen is megfogalmazható:

(a) Mivel az $\mathbf{A} \cdot x$ szorzat egy olyan oszlopvektor, ami az $\mathbf{A}$ mátrix oszlopainak az x vektor komponenseivel képzett lineáris kombinációja, az (1) egyenletrendszer pontosan akkor oldható meg, ha a b vektor az $\mathbf{A}$ oszlopainak lineáris kombinációja. Tekintve, hogy az $\mathbf{A}$ mátrix $a_1, \dots, a_n$ oszlopai és a b vektor is az $\mathbb{F}^m$ koordinátatérnek a vektorai, azt is mondhatjuk, hogy az (1) egyenletrendszer akkor és csak akkor oldható meg, ha $b \in \operatorname{lin}(a_1, \dots, a_n)$, amit úgy mondtunk, hogy b kompatibilis az $\{a_1, \dots, a_n\}$ vektorrendszerre nézve.

(b) Az egyenletrendszerek megoldhatóságának már klasszikusnak mondható *Kronecker–Capelli-féle kritériuma* így szól:
Az (1) egyenletrendszer megoldhatóságának szükséges és elegendő feltétele, hogy $\rho(\mathbf{A}) = \rho([\mathbf{A}, b])$ teljesüljön, ahol az $[\mathbf{A}, b]$ bővített mátrixot úgy kapjuk, hogy az $\mathbf{A}$ oszlopai mellé, még a b oszlopot is hozzávesszük a mátrixhoz.

Annak az oka, hogy a lineáris egyenletrendszerek tárgyalását nem részleteztük a kompatibilitási vizsgálatokhoz kapcsoltan, az, hogy szeretnénk kihasználni a megoldási algoritmus konstruálásakor azt a tényt, hogy az $x \to \mathbf{A} \cdot x$ hozzárendelés az $\mathbb{F}^n$ vektortérnek az $\mathbb{F}^m$ vektortérbe való A lineáris leképezése. Az $\mathbb{F}^n$ térben véve az $\mathcal{E} = \{e_1, \dots, e_n\}$ bázist, míg az $\mathbb{F}^m$-ben az $\mathcal{E}' = \{e'_1, \dots, e'_m\}$ bázist, (ahol e_i, (e'_j) az az n komponensű (m komponensű) egységvektor, amelynek i-edik (j-edik) komponense 1) az A lineáris leképezés mátrixa e bázispárra vonatkozóan éppen $\mathbf{A}$, ami azt jelenti, hogy az $e_i \in \mathcal{E} \quad (i = 1, \dots, n)$ vektor $A(e_i)$ képe éppen a_i. Az (1) egyenletrendszer megoldhatóságának szükséges és elegendő feltétele tehát így is megfogalmazható:

(c) Az (1) egyenletrendszer pontosan akkor oldható meg, ha az $x \to \mathbf{A} \cdot x$ hozzárendelési szabály által adott A lineáris leképezés képtere tartalmazza a $b \in \mathbb{F}^m$ vektort.

Az $\mathcal{E}$ és $\mathcal{E}'$ bázisok abból a szempontból nagyon jónak mondhatók, hogy minden $\mathbb{F}^n$-beli, illetve $\mathbb{F}^m$-beli vektor komponensei és e bázisokra vonatkozó koordinátái egyenlők, de ahhoz, hogy könnyen eldönthető legyen az (1) egyenletrendszer megoldhatóságának kérdése, szerencsésebb olyan bázispárt választani, amelyre vonatkozó koordináta vektorok ismeretében azonnal megsmondható, hogy a b vektor benne van-e a leképezés képterében, és ha igen, melyek azok az $\mathbb{F}^n$-beli vektorok, amelyeknek a képe b. Ha a b vektornak meghatározzuk a koordináta vektorát az $\mathbb{F}^m$ tér egy olyan bázisában, amely tartalmazza a leképezés képterének egy bázisát, akkor abból azonnal kiolvasható, hogy b eleme-e a képtérnek, nevezetesen pontosan akkor eleme, ha csak a képtér bázisát alkotó vektorokra vonatkozó koordinátái különböznek zérótól, de minden más koordinátája nulla. Ha a leképezés képterének bázisvektorait az $\mathbf{A}$ mátrix oszlopai közül választjuk, — ez mindig lehetséges, mert $\mathbf{A}$ oszlopai generálják a képteret — véve, mondjuk az $a_{i1}, \ldots, a_{ir}$ vektorokat, és ezt egészítjük ki az $\mathbb{F}^m$ tér bázisává, akkor ahhoz, hogy b benne lehessen a képtérben, e bázisra vonatkozó koordináta vektorának csak az $a_{i1}, \ldots, a_{ir}$ vektorokra vonatkozó elemei lehetnek nullától különbözőek, minden más koordináta 0 kell legyen. Ha történetesen

$$b = \delta_{i1} a_{i1} + \cdots + \delta_{ir} a_{ir} \,,$$

akkor azt is azonnal kapjuk, hogy az

$$x_0 = \delta_{i1} e_{i1} + \cdots + \delta_{ir} e_{ir} \in \mathbb{F}^n$$

vektor

$$A(x_0) = \delta_{i1} A(e_{i1}) + \cdots + \delta_{ir} A(e_{ir})$$

képe b, tehát az x_0 vektor egy megoldása (1)–nek. Persze ha x egy tetszőleges eleme az A leképezés $\ker(A)$ magterének, akkor az

$$A(x + x_0) = A(x) + A(x_0) = 0 + b = b \,,$$

mutatva, hogy az $x + x_0$ vektor is megoldása az (1) egyenletrendszernek, sőt minden megoldás megkapható ilyen összeg alakban. Ezt az alábbi tételben be is bizonyítjuk.

3.1.1 Tétel. *Legyen $\mathbf{A} \in \mathbb{F}_{m \times n}$ tetszőleges mátrix és $b \in \mathbb{F}^m$ tetszőleges vektor. Az $\mathbf{A} \cdot x = b$ inhomogén lineáris egyenletrendszer minden megoldása megkapható egy x_0 megoldásának és az $\mathbf{A} \cdot x = 0$ homogén lineáris egyenletrendszer valamely megoldásának összegeként.*

Bizonyítás. Ha x' tetszőleges megoldása az inhomogén egyenletrendszernek, akkor $\mathbf{A} \cdot x' = \mathbf{A} \cdot x_0 = b$, amiből

$$\mathbf{A} \cdot (x' - x_0) = \mathbf{A} \cdot x' - \mathbf{A} \cdot x_0 = b - b = 0 \,,$$

tehát $x' - x_0$ a homogén egyenletrendszer megoldása, és nyilván $x' = x_0 + (x' - x_0)$, amint állítottuk. □

A fenti tétel mutatja, hogy a homogén lineáris egyenletrendszerek összes megoldásának meghatározása kulcsfontosságú a lineáris egyenletrendszerek megoldásainak keresésekor.

3.1.1 Homogén lineáris egyenletrendszerek megoldása

Az $\mathbf{A} \cdot x = 0$ homogén lineáris egyenletrendszert megoldani annyit jelent, mint meghatározni azon $A \in L(\mathbb{F}^n, \mathbb{F}^m)$ lineáris leképezés magterét, amelyet az $x \to \mathbf{A} \cdot x$ hozzárendelési szabály definiál. Mivel egy lineáris leképezés magtere altér, tehát legalább a nullvektort tartalmazza, a homogén lineáris egyenletrendszereknek mindig van megoldása, és *megoldáshalmaza* altere $\mathbb{F}^n$-nek. Ennek az altérnek a meghatározásához felhasználjuk az $\mathbf{A}$ mátrix bázisfaktorizációját. Felbontva az $\mathbf{A}$ mátrixot a bázisfaktorizációnak megfelelően $\mathbf{B} \cdot \mathbf{C}$ szorzatra, az

$$\mathbf{A} \cdot x = \mathbf{B} \cdot \mathbf{C} \cdot x = 0$$

alakhoz jutunk, ahol a $\mathbf{B}$ mátrix oszlopai az A leképezés képterének bázisa. Ennélfogva a $\mathbf{B} \cdot \mathbf{C} \cdot x = 0$ akkor és csak akkor teljesül, ha $\mathbf{C} \cdot x = 0$. Ha $\rho(\mathbf{A}) = r$, és történetesen az $\mathbf{A}$ mátrix $a_1, \ldots, a_r$ oszlopai független rendszert alkotnak, akkor ezek a vektorok alkothatják $\operatorname{im}(A)$ bázisát. Így $\mathbf{B} = [a_1, \ldots, a_r]$ $m \times r$ tipusú mátrix és $\mathbf{C}$ pedig, tekintettel arra, hogy oszlopai, az $\mathbf{A}$ mátrix oszlopainak a koordináta vektorai az $\{a_1, \ldots, a_r\}$ bázisban, $r \times n$ tipusú és első, második, $\ldots, r$-edik oszlopa rendre az $e_1, \ldots, e_r$ r komponensű egységvektor, azaz $\mathbf{C} = [\mathbf{E}, \mathbf{D}]$ alakú. Persze azt kérdezheti az olvasó, hogy mi biztosítja, hogy az $\mathbf{A}$ mátrixnak éppen az első r oszlopa lineárisan független. A válasz az, hogy semmi, de az egyenletrendszer ismeretlenjeinek átindexelésével, ami az $\mathbf{A}$ oszlopainak cseréjével jár együtt, ez mindig elérhető, így feltevésünk valójában nem csorbítja az általánosságot, ugyanakkor nagy segítségünkre van a formalizálásban. Ha a $\mathbf{C}$ mátrix particionálásának megfelelően az ismeretleneket tartalmazó x vektort is az

$$x_1 = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_r \end{bmatrix} \qquad \text{és} \qquad x_2 = \begin{bmatrix} \xi_{r+1} \\ \vdots \\ \xi_n \end{bmatrix}$$

vektorokra bontjuk, akkor a

$$\mathbf{C} \cdot x = [\mathbf{E}, \mathbf{D}] \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 + \mathbf{D} \cdot x_2 = 0$$

alakhoz jutunk, amiből kiolvashatjuk, hogy egy x vektor akkor és csak akkor lesz megoldása az $\mathbf{A} \cdot x = 0$ homogén lineáris egyenletrendszernek, ha komponensei között fennáll az

$$(*) \qquad x_1 = -\mathbf{D} x_2$$

kapcsolat, ahol az x_1 vektorba gyűjtöttük össze az $\mathbf{A}$ mátrix oszlopvektorterének bázisát alkotó oszlopaihoz tartozó ismeretleneket, amelyeket *kötött ismeretleneknek* nevezünk, míg a többi ismeretlen alkotja az x_2 vektor komponenseit. A $\mathbf{D}$ mátrix oszlopai az $\mathbf{A}$ mátrix azon oszlopainak koordináta vektorai, amelyek nincsenek az oszlopvektortér választott bázisában. Ha az x_2 komponenseinek tetszőlegesen adunk értékeket, és az x_1 vektort a (*) feltételt kielégítendő $-\mathbf{D} \cdot x_2$-vel tesszük egyenlővé, akkor az így kapott

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\mathbf{D} \cdot x_2 \\ x_2 \end{bmatrix}$$

vektor az $\mathbf{A}\cdot x = 0$ egyenletrendszer megoldása lesz. Ezért az x_2 komponenseit alkotó ismeretleneket *szabad ismeretleneknek* hívjuk. A szabad ismeretlenek s számát az egyenletrendszer *szabadság fokának* nevezzük. Sokszor célszerű a megoldáshalmazt

$$\begin{bmatrix} -\mathbf{D} \\ \mathbf{E} \end{bmatrix} \cdot t \qquad (t \in \mathbb{F}^s)$$

alakban megadni, ahol $\mathbf{E}$ $s \times s$ tipusú egységmátrix. Ezt az alakot a homogén lineáris egyenletrendszer *általános megoldásának* is mondjuk, míg a megoldáshalmaz egyes vektorait *partikuláris megoldásoknak* nevezzük. Az általános megoldásként említett alakból kiderül, hogy a

$$\begin{bmatrix} -\mathbf{D} \\ \mathbf{E} \end{bmatrix}$$

mátrix oszlopai az egyenletrendszerhez tartozó A lineáris leképezés $\ker(A)$ magterének bázisvektorai. A (2.1.6) tétel alapján nyilvánvaló, hogy az egyenletrendszer szabadságfoka: $s = n - \rho(\mathbf{A})$. Ebből azonnal adódik:

3.1.2 Következmény. *Az $\mathbf{A} \cdot x = 0$ $(\mathbf{A} \in \mathbb{F}_{m\times n})$ homogén lineáris egyenletrendszernek pontosan akkor van nemzéró megoldása, ha $\mathbf{A}$ rangja kisebb, mint n, azaz, ha az együtthatók mátrixának oszlopai lineárisan összefüggő rendszert alkotnak. Ha $\rho(\mathbf{A}) = n$, tehát ha $\mathbf{A}$ oszlopai lineárisan független rendszert alkotnak, akkor pedig a nullvektor az egyetlen megoldás.*

1 Példa. *Határozzuk meg a*

$$\begin{array}{rcrcrcrcl}
\xi_1 & & & + & \xi_3 & + & 2\xi_4 & = & 0 \\
2\xi_1 & + & \xi_2 & + & 4\xi_3 & + & 6\xi_4 & = & 0 \\
-\xi_1 & + & 2\xi_2 & + & 3\xi_3 & + & 2\xi_4 & = & 0 \\
\xi_1 & - & \xi_2 & - & \xi_3 & & & = & 0
\end{array}$$

homogén lineáris egyenletrendszer megoldáshalmazát és az(oka)t a partikuláris megoldás(oka)t, ahol a $\xi_3 = 2$.

Az egyenletrendszer együttható mátrixa

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 4 & 6 \\ -1 & 2 & 3 & 2 \\ 1 & -1 & -1 & 0 \end{bmatrix}$$

Először meg kell határoznunk az $\mathbf{A}$ oszlopvektorterének egy bázisát és az $\mathbf{A}$ oszlopainak e bázisra vonatkozó koordináta vektorait. Ezt elemi bázistranszformációval végezhetjük. Az alábbi táblázatban az $\mathbf{A}$ mátrix oszlopait a megfelelő ismeretlenekkel fejléceztük, ami lineáris egyenletrendszerek megoldásakor azért előnyös, mert a táblázatból kiolvasható lesz, hogy melyek a kötött ismeretlenek, míg az eredeti bázis vektorai az $\mathbb{R}^4$ valós koordinátatér $e_1, \ldots, e_4$ egységvektorai.

$$\begin{array}{c|cccc|}
 & \xi_1 & \xi_2 & \xi_3 & \xi_4 \\ \hline
e_1 & \boxed{1} & 0 & 1 & 2 \\
e_2 & 2 & 1 & 4 & 6 \\
e_3 & -1 & 2 & 3 & 2 \\
e_4 & 1 & -1 & -1 & 0 \\ \hline
\end{array}
\;\rightarrow\;
\begin{array}{c|ccc|}
 & \xi_2 & \xi_3 & \xi_4 \\ \hline
\xi_1 & 0 & 1 & 2 \\
e_2 & \boxed{1} & 2 & 2 \\
e_3 & 2 & 4 & 4 \\
e_4 & -1 & -2 & -2 \\ \hline
\end{array}
\;\rightarrow\;
\begin{array}{c|cc|}
 & \xi_3 & \xi_4 \\ \hline
\xi_1 & 1 & 2 \\
\xi_2 & 2 & 2 \\
e_3 & 0 & 0 \\
e_4 & 0 & 0 \\ \hline
\end{array}$$

Az utolsó táblázatból kiolvasható, hogy $\rho(\mathbf{A}) = 2$, így a szabadságfok is $2(= 4-2)$, továbbá, hogy az $\mathbf{A}$ mátrix

$$a_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 1 \end{bmatrix} \qquad \text{és} \qquad a_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ -1 \end{bmatrix}$$

oszlopvektorai bázist alkotnak az $\mathbf{A}$ oszlopvektorterében, ezért a nekik megfelelő ξ_1 és ξ_2 ismeretlenek lesznek a kötött ismeretlenek és a $\xi_3\,, \xi_4$ ismeretlenek pedig a szabad ismeretlenek. Az is látható, hogy az $\mathbf{A}$ a_3 és a_4 oszlopainak a választott bázisra vonatkozó koordináta vektorai az

$$\mathbf{a}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{és} \qquad \mathbf{a}_4 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

és így az $\mathbf{A}$ bázisfaktorizációja a választott bázis mellett

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 2 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 2 \end{bmatrix}$$

Ennek megfelelően a kötött és szabad ismeretlenek között fenn kell álljon a

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = -\begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} \xi_3 \\ \xi_4 \end{bmatrix}$$

kapcsolat. Ennek alapján most már az általános megoldás:

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ -2 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot t \quad \text{ahol} \quad t \in \mathbb{R}^2$$

Azokat a partikuláris megoldásokat kell még megadnunk, ahol a $\xi_3 = 2$. Ebből a szempontból, mint látható, körültekintően választottuk $\mathbf{A}$ oszlopvektorterének bázisát, a ξ_3 ismeretlent szabad ismeretlennek hagytuk (a neki megfelelő a_3 vektort nem választottuk bázisvektornak), így most az általános megoldás felhasználásával ezeket a partikuláris megoldásokat gyorsan megkaphatjuk, ha a t vektor első koordinátájaként 2-t választunk, míg a második koordináta továbbra is tetszőleges valós szám. Tehát a keresett partikuláris megoldások:

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} = \begin{bmatrix} -2-2\alpha \\ -4-2\alpha \\ 2 \\ \alpha \end{bmatrix} \qquad (\alpha \in \mathbb{R})\,.$$

□

3.1.2 Inhomogén lineáris egyenletrendszerek megoldása

Az $\mathbf{A} \cdot x = b$ inhomogén lineáris egyenletrendszer megoldásait a (3.1.1) tétel értelmében megkaphatjuk, ha meghatározzuk egy megoldását és ahhoz rendre hozzáadjuk az $\mathbf{A} \cdot x = 0$ homogén lineáris egyenletrendszer megoldásait. Minthogy a homogén egyenletrendszer megoldási algoritmusa már adott, csupán azt kell még tisztáznunk, hogy hogyan lehetne az inhomogén egyenletrendszer egyetlen megoldását megkonstruálni.

Mint azt már tudjuk ilyen megoldás pontosan akkor létezik, ha b benne van az $\mathbf{A}$ mátrix oszlopvektorterében. Ha $\mathbf{A}$ bázisfaktorizációja $\mathbf{A} = \mathbf{B} \cdot \mathbf{C}$, éppen a $\mathbf{B}$ oszlopai alkotják az $\mathbf{A}$ oszlopvektorterének egy bázisát, következésképpen, ha van megoldása az inhomogén egyenletrendszernek, akkor b a $\mathbf{B}$ mátrix oszlopainak lineáris kombinációja. Jelölje $\mathbf{b}$ a b vektor az oszlopvektortér e bázisára vonatkozó koordináta vektorát, akkor $\mathbf{Bb} = b$ és az egyenletrendszer

$$(**) \qquad \mathbf{A} \cdot x = \mathbf{B} \cdot \mathbf{C} \cdot x = b = \mathbf{Bb}$$

alakot ölt. Tekintve, hogy $\mathbf{A}$ oszlopvektorterének minden vektora egyértelműen állítható elő bázisvektorainak, esetünkben a $\mathbf{B}$ oszlopainak, lineáris kombinációjaként a (**) akkor és csak akkor teljesülhet, ha

$$\mathbf{C} \cdot x = \mathbf{b}\,.$$

A könnyebb formalizálhatóság kedvéért megint feltéve, hogy az $\mathbf{A}$ mátrix első $r(= \rho(\mathbf{A}))$ oszlopa alkotja az $\mathbf{A}$ oszlopvektorterének a bázisát (ezek az oszlopok alkotják a $\mathbf{B}$ mátrixot), kapjuk, hogy $\mathbf{C} = [\mathbf{E}, \mathbf{D}]$ és ennek megfelelően felbontva az x vektort az x_1 kötött ismeretleneket tartalmazó és az x_2 szabad ismeretleneket tartalmazó vektorokra az

$$[\mathbf{E}, \mathbf{D}] \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 + \mathbf{D} \cdot x_2 = \mathbf{b}$$

egyenletet kapjuk. Ebben az egyenletben x_1 és $\mathbf{b}$ r komponensű, míg x_2 $s(= n - r)$ komponensű vektor, és az egyenletet nyilvánvalóan kielégíti az $x_1 = \mathbf{b}$ és $x_2 = \mathbf{0}$ részekből összerakott

$$x = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}$$

vektor. Ez tehát egy megoldása az inhomogén egyenletrendszernek, amit *bázismegoldásnak* nevezünk. Ezekután az inhomogén egyenletrendszer általános megoldása már könnyen kapható, a bázismegoldás és a homogén egyenletrendszer általános megoldásának összegeként:

$$x = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{D} \\ \mathbf{E} \end{bmatrix} \cdot t \quad \text{ahol} \quad t \in \mathbb{F}^s$$

A különböző partikuláris megoldások persze most is úgy kaphatók, hogy a t paraméter vektornak konkrét értéket adunk. Speciálisan $t = \mathbf{0}$ választással adódik a már említett bázismegoldás. Természetesen az inhomogén lineáris egyenletrendszerek megoldásakor a bázismegoldás és a homogén egyenletrendszer általános megoldásának előállítása egyidejűleg történik, az $\mathbf{A}$ oszlopvektortere bázisának meghatározásakor a b vektorról azt is eldönthetjük, hogy előállítható–e a bázisvektorok lineáris kombinációjaként (van–e megoldás), és ha igen, akkor megkaphatjuk e bázisra vonatkozó **b** koordináta vektorát is.

A Következő példa bemutatja a fentiekben leírt megoldási algoritmus működését a gyakorlatban.

2 Példa. *Határozzuk meg az alábbi inhomogén lineáris egyenletrendszer megoldáshalmazát:*

$$\begin{array}{rcrcrcrcr} \xi_1 & & & + & 2\xi_3 & - & \xi_4 & = & 2 \\ 2\xi_1 & + & \xi_2 & + & 5\xi_3 & & & = & 6 \\ 2\xi_1 & + & 2\xi_2 & + & 6\xi_3 & + & 2\xi_4 & = & 8 \\ & - & \xi_2 & - & \xi_3 & - & 2\xi_4 & = & -2 \end{array}$$

A megoldás csak annyiban tér el a megfelelő homogén lineáris egyenletrendszer megoldásától, hogy az elemi bázistranszformációs táblázatban követnünk kell az egyenletrendszer jobboldalán szereplő konstansokból felépített b vektor koordináta vektorának változását is a bázis megváltoztatásakor, ugyanis szükségünk van az együtthatómátrix oszlopvektortere bázisára vonatkozó koordináta vektorára, hogy az inhomogén egyenletrendszer bázismegoldását meghatározzuk. Az induló táblában az eredeti bázis az $\mathbb{R}^4$ valós kordinátatér egységvektorai alkotta bázis.

$$\begin{array}{c|cccc|c} & \xi_1 & \xi_2 & \xi_3 & \xi_4 & b \\ \hline e_1 & \boxed{1} & 0 & 2 & -1 & 2 \\ e_2 & 2 & 1 & 5 & 0 & 6 \\ e_3 & 2 & 2 & 6 & 2 & 8 \\ e_4 & 0 & -1 & -1 & -2 & -2 \\ \hline \end{array} \rightarrow \begin{array}{c|ccc|c} & \xi_2 & \xi_3 & \xi_4 & b \\ \hline \xi_1 & 0 & 2 & -1 & 2 \\ e_2 & 1 & \boxed{1} & 2 & 2 \\ e_3 & 2 & 2 & 4 & 4 \\ e_4 & -1 & -1 & -2 & -2 \\ \hline \end{array} \rightarrow$$

$$\begin{array}{c|cc|c} & \xi_2 & \xi_4 & b \\ \hline \xi_1 & -2 & -5 & -2 \\ \xi_3 & 1 & 2 & 2 \\ e_3 & 0 & 0 & 0 \\ e_4 & 0 & 0 & 0 \\ \hline \end{array}$$

Az utolsó táblázat mutatja, hogy **A** oszlopvektorterének egy bázisát alkotják az **A** a_1 és a_3 oszlopai, és mivel a b vektor kifejezhető ezek lineáris kombinációjaként, az egyenletrendszernek van megoldása. Az $\{a_1, a_3\}$ bázisban a b vektor koordináta vektora: $\mathbf{b} = [-2, 2]^*$. Tekintettel arra, hogy most az **A** mátrixnak nem az első két oszlopát választottuk az oszlopvektortér bázisának (tehettük volna, csak szeretnénk bemutatni, hogy ilyen esetben hogyan írhatók fel az egyenletrendszer megoldásai), a **b** vektor és a kiegészítő nullvektor komponensei permutálódnak. Egyenletrendszerünk szabadságfoka 2, a ξ_2 és ξ_4 ismeretlenek szabad ismeretlenek. A bázismegoldás és a megfelelő homogén lineáris egyenletrendszer általános megoldása:

$$x_0 = \begin{bmatrix} -2 \\ 0 \\ 2 \\ 0 \end{bmatrix} \qquad \text{és} \qquad x_h = \begin{bmatrix} 2 & 5 \\ 1 & 0 \\ -1 & -2 \\ 0 & 1 \end{bmatrix} \cdot t \quad \text{ahol} \quad t \in \mathbb{R}^2$$

Így most már az inhomogén lineáris egyenletrendszer általános megodása:

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 & 5 \\ 1 & 0 \\ -1 & -2 \\ 0 & 1 \end{bmatrix} \cdot t \quad \text{ahol} \quad t \in \mathbb{R}^2$$

Fel kell hívjuk az olvasó figyelmét arra is, hogy a megfelelő homogén lineáris egyenletrendszer megoldása is tükrözi azt a tényt, hogy az első és harmadik ismeretlen a kötött– míg a második és negyedik a szabad ismeretlen. □

Az alábbi példában egy olyan inhomogén lineáris egyenletrendszert oldunk meg, amelyben az együtthatók mátrixa paramétert tartalmaz. Megvizsgáljuk, hogy a paraméter milyen értékei mellett van megoldás, illetve, hogy a paraméter értékének változtatásával hogyan változik az egyenletrendszer szabadságfoka.

3 Példa. *Vizsgáljuk meg, hogy az α paraméter milyen értékei mellett van megoldása az alábbi lineáris egyenletrendszernek, és ha van megoldás, mennyi a szabadságfok.*

$$\begin{array}{rcrcrcl} \xi_1 & + & \alpha\xi_2 & + & \xi_3 & = & 1 \\ \xi_1 & + & \xi_2 & + & \alpha\xi_3 & = & 1 \\ \alpha\xi_1 & + & \xi_2 & + & \xi_3 & = & 1 \end{array}$$

Az induló táblázat:

$$\begin{array}{c|ccc|c} & \xi_1 & \xi_2 & \xi_3 & b \\ \hline e_1 & \boxed{1} & \alpha & 1 & 1 \\ e_2 & 1 & 1 & \alpha & 1 \\ e_3 & \alpha & 1 & 1 & 1 \\ \hline \end{array} \rightarrow \begin{array}{c|cc|c} & \xi_2 & \xi_3 & b \\ \hline \xi_1 & \alpha & 1 & 1 \\ e_2 & \boxed{1-\alpha} & \alpha - 1 & 0 \\ e_3 & 1-\alpha^2 & 1-\alpha & 1-\alpha \\ \hline \end{array}$$

Mint az látható, nem választható olyan generáló elem, amely ne lenne α polinomja. Mivel $\alpha = 1$ esetén táblázatunk:

$$\begin{array}{c|cc|c} & \xi_2 & \xi_3 & b \\ \hline \xi_1 & 1 & 1 & 1 \\ e_2 & 0 & 0 & 0 \\ e_3 & 0 & 0 & 0 \\ \hline \end{array}$$

abból kiolvashatjuk, hogy $\alpha = 1$ estén van megoldás, a szabadságfok 2 és az általános megoldás:

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot t \quad \text{ahol} \quad t \in \mathbb{R}^2$$

Ha $\alpha \neq 1$ akkor az együtthatómátrix második oszlopa független az elsőtől, bevonható az oszlopvektortér bázisába. Az új bázisra vonatkozó koordináta vektorok pedig az alábbi táblázatban láthatók:

$$\begin{array}{c|c|c} & \xi_3 & b \\ \hline \xi_1 & 1+\alpha & 1 \\ \xi_2 & -1 & 0 \\ e_3 & \boxed{2-\alpha-\alpha^2} & 1-\alpha \\ \hline \end{array}$$

Az együtthatómátrix harmadik oszlopa csak akkor független az első kettőtől, ha $2-\alpha-\alpha^2=(1-\alpha)(2+\alpha)\neq 0$, amiből látható hogy meg kell vizsgálnunk az $\alpha=-2$ esetet. Ebben az esetben a táblázat:

	ξ_3	b
ξ_1	-1	1
ξ_2	-1	0
e_3	0	3

mutatja, hogy a b vektor nincs benne az együtthatók mátrixának oszlopvektorterében, tehát nincs megoldása az egyenletrendszernek. Ha α mind 1-től, mind -2-től különböző, akkor az együtthatók mátrixának oszlopvektortere megegyezik az $\mathbb{R}^3$ koordinátatérrel és így a szabadságfok 0, tehát egyetlen megoldás létezik, amelyet a harmadik oszlop bázisba vonása után ki is olvashatunk a

	b
ξ_1	$\frac{1}{2+\alpha}$
ξ_2	$\frac{1}{2+\alpha}$
ξ_3	$\frac{1}{2+\alpha}$

táblázatból: $x=\frac{1}{2+\alpha}[1,1,1]^*$. □

3.2 Mátrixegyenletek

Ha V , W és Z az $\mathbb{F}$ test feletti m , n és p-dimenziós vektorterek, továbbá $B\in L(V,Z)$ és $A\in L(W,Z)$ lineáris leképezések, felvethető a kérdés, hogy létezik–e olyan $X\in L(V,W)$ lineáris leképezés, hogy $A\cdot X=B$ teljesül. A válasz könnyen megsejthető, nevezetesen, hogy amennyiben a B leképezés $\operatorname{im}(B)$ képtere altere az A leképezés $\operatorname{im}(A)$ képterének, akkor létezik ilyen X leképezés. Ha ismert az A leképezés $\mathbf{A}\in\mathbb{F}_{p\times n}$ és a B leképezés $\mathbf{B}\in\mathbb{F}_{p\times m}$ mátrixa, akkor előző kérdésünk a mátrixaritmetika nyelvére lefordítva így hangzik:
Milyen feltételek mellett oldható meg az

$$\mathbf{A}\cdot\mathbf{X}=\mathbf{B} \tag{3.1}$$

mátrixegyenlet? A mátrixok szorzási szabályát ismerve, világos, hogy csak $n\times m$ tipusú $\mathbf{X}$ mátrix elégítheti ki a (3.1) egyenletet.
Jelöljük az ismeretlen $\mathbf{X}$ mátrix oszlopait $x_1,x_2,\ldots,x_m$-mel, és a $\mathbf{B}$ oszlopait pedig $b_1,b_2,\ldots,b_m$-mel. Akkor kihasználva, hogy az $\mathbf{A}\cdot\mathbf{X}$ mátrix oszlopai rendre $\mathbf{A}\cdot x_1,\mathbf{A}\cdot x_2,\ldots,\mathbf{A}\cdot x_m$, kapjuk, hogy a (3.1) egyenlet ekvivalens az

$$\mathbf{A}\cdot x_1=b_1,\mathbf{A}\cdot x_2=b_2,\ldots,\mathbf{A}\cdot x_m=b_m$$

lineáris egyenletrendszerek rendszerével. Ezeknek az egyenletrendszereknek mindegyike pontosan akkor oldható meg, ha a $\mathbf{B}$ minden oszlopa benne van az $\mathbf{A}$ mátrix oszlopvektorterében, azaz, ha a $\mathbf{B}$ mátrix oszlopvektortere altere az $\mathbf{A}$ oszlopvektorterének. Ez éppen fenti sejtésünk mátrixaritmetikai megfelelője.

A fenti egyenletrendszerek együtthatómátrixa közös. Ez lehetővé teszi, hogy azokat egyszerre oldjuk meg. Ezt kívánja illusztrálni az alábbi példa:

4 Példa. *Oldjuk meg az*

$$\begin{bmatrix} 1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 2 \\ 1 & 2 & 3 & 1 \end{bmatrix} \cdot \mathbf{X} = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 7 & 6 \end{bmatrix}$$

mátrixegyenletet!

Két inhomogén lineáris egyenletrendszer megoldásai adják az **X** mátrix első és második oszlopát. Megoldásukat elemi bázistranszformációs módszerrel végeztük:

	a_1	a_2	a_3	a_4	b_1	b_2
e_1	$\boxed{1}$	0	1	-1	1	2
e_2	-1	1	0	2	2	0
e_3	1	2	3	1	7	6

$\rightarrow$

	a_2	a_3	a_4	b_1	b_2
a_1	0	1	-1	1	2
e_2	$\boxed{1}$	1	1	3	2
e_3	2	2	2	6	4

$\rightarrow$

	a_3	a_4	b_1	b_2
a_1	1	-1	1	2
a_2	1	1	3	2
e_3	0	0	0	0

Az utolsó táblázatból kiolvashatjuk a $\mathbf{A} \cdot x_1 = b_1$ és $\mathbf{A} \cdot x_2 = b_2$ egyenletrendszerek általános megoldásait:

$$x_1 = \begin{bmatrix} \xi_{11} \\ \xi_{21} \\ \xi_{31} \\ \xi_{41} \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot t_1$$

és

$$x_2 = \begin{bmatrix} \xi_{12} \\ \xi_{22} \\ \xi_{32} \\ \xi_{42} \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot t_2\,,$$

ahol t_1, t_2 tetszőlegesen választható kétkomponensű vektorok. Ezek alapján a mátrixegyenlet általános megoldása:

$$\mathbf{X} = \begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \\ \xi_{31} & \xi_{32} \\ \xi_{41} & \xi_{42} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \mathbf{T} \quad \text{ahol} \quad \mathbf{T} = [t_1, t_2] \in \mathbb{R}_{2\times 2}$$

Végtelen sok megoldás létezik, mert a megfelelő egyenletrendszerek szabadságfoka 2. A 2×2-es **T** mátrix elemei tetszőlegesen választhatók. □

3.3 Mátrix inverzének numerikus meghatározása

A transzformációk inverzéről szóló alpontban láttuk, hogy ha a V véges dimenziós vektortér egy A reguláris transzformációjának mátrixa valamely bázisban $\mathbf{A}$, akkor A^{-1} inverzének $\mathbf{A}^{-1}$ mátrixa kielégíti az

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{E}$$

mátrixegyenletet, sőt egy példán be is mutattuk, hogy az inverz transzformáció mátrixát hogyan lehet numerikusan meghatározni. Ezért ebben a részben csak arra kívánunk rámutatni, hogy a numerikus számítások lerövidíthetők. Amint láttuk, ha A a V tér egy $\mathcal{X} = \{x_1, \dots, x_n\}$ bázisát egy másik $\mathcal{Y} = \{y_1, \dots, y_n\}$ bázisba transzformálja, akkor A és A^{-1} mátrixa az $\mathcal{X}$ bázisban ugyanaz, mint az $\mathcal{Y}$ bázisban. Másrészt, $\mathbf{A}_{\mathcal{X}}$ oszlopai éppen az $y_1, \dots, y_n$ vektorok koordináta vektorai az $\mathcal{X}$ bázisban, míg $\mathbf{A}_{\mathcal{Y}}^{-1}$ oszlopai az $x_1, \dots, x_n$ vektorok $\mathcal{Y}$ bázisra vonatkozó koordináta vektorai. Ez gyakorlatilag azt jelenti, hogy a $\mathbf{A}_{\mathcal{X}}$ mátrix inverzének meghatározása nem más, mint az $\mathcal{X}$ bázis kicserélése az $\mathcal{Y}$ bázisra és az eredeti bázisvektorok új bázisra vonatkozó koordináta vektorainak kiszámítása. Amint azt az elemi bázistranszformáció tárgyalásakor láttuk, ha az

$$y = \psi_1 x_1 + \cdots + \psi_{i-1} x_{i-1} + \psi_i x_i + \psi_{i+1} x_{i+1} + \cdots + \psi_n x_n$$

vektort az x_i vektor helyére a bázisba vonjuk, (ezt megtehetjük, ha a ψ_i generáló elem nem nulla) akkor

$$x_i = -\frac{\psi_1}{\psi_i} x_1 - \cdots - -\frac{\psi_{i-1}}{\psi_i} x_{i-1} + \frac{1}{\psi_i} y - \frac{\psi_{i+1}}{\psi_i} x_{i+1} - \cdots - \frac{\psi_n}{\psi_i} x_n$$

lesz, azaz a bázisból elhagyott x_i vektor koordinátái a bázisba bevont y vektor eredeti koordinátáiból könnyen kiszámíthatók; az új, y bázisvektorra vonatkozó koordináta a generáló elem reciproka, míg minden más $x_j \ \ (j \neq i)$ bázisvektorra vonatkozó koordináta az y megfelelő koordinátájának ellentettje és a generáló elem reciprokának a szorzata. Az alábbi táblákon jól nyomon követhető a szavakkal nehezen megfogalmazható változás:

$$\begin{array}{r|c}
 & \mathbf{y} \\ \hline
x_1 & \psi_1 \\
\vdots & \vdots \\
x_{i-1} & \psi_{i-1} \\
x_i & \boxed{\psi_i} \\
x_{i+1} & \psi_{i+1} \\
\vdots & \vdots \\
x_n & \psi_n \\ \hline
\end{array}
\quad \longrightarrow \quad
\begin{array}{r|c}
 & \mathbf{x}_i \\ \hline
x_1 & -\frac{\psi_1}{\psi_i} \\
\vdots & \vdots \\
x_{i-1} & -\frac{\psi_{i-1}}{\psi_i} \\
y & \frac{1}{\psi_i} \\
x_{i+1} & -\frac{\psi_{i+1}}{\psi_i} \\
\vdots & \vdots \\
x_n & -\frac{\psi_n}{\psi_i} \\ \hline
\end{array}$$

A következő példában bemutatjuk egy reguláris mátrix inverzének meghatározását az elmondottak felhasználásával.

5 Példa. *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 2 & 4 \end{bmatrix}$$

mátrix inverzét!

Jelöljék az x_1, x_2, x_3 vektorok az eredeti bázisvektorokat és legyenek

$$\begin{aligned} y_1 &= x_1 - x_2 + x_3\,, \\ y_2 &= x_2 + 2x_3\,, \\ y_1 &= x_1 + 4x_3\,. \end{aligned}$$

Akkor $\mathbf{A}$ annak az A lineáris transzformációnak a mátrixa, amelyre $A(x_i) = y_i$ $(i = 1, 2, 3)$. Elemi bázistranszformációk sorozatával áttérünk az $\{x_1, x_2, x_3\}$ bázisról az $\{y_1, y_2, y_3\}$ bázisra és kiszámítjuk a bázisból kihagyott vektorok új bázisra vonatkozó koordináta vektorait:

$$\begin{array}{c|rrr} & \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 \\ \hline x_1 & \boxed{1} & 0 & 1 \\ x_2 & -1 & 1 & 0 \\ x_3 & 1 & 2 & 4 \end{array} \rightarrow \begin{array}{c|rrr} & \mathbf{x}_1 & \mathbf{y}_2 & \mathbf{y}_3 \\ \hline y_1 & 1 & 0 & 1 \\ x_2 & 1 & \boxed{1} & 1 \\ x_3 & -1 & 2 & 3 \end{array} \rightarrow \begin{array}{c|rrr} & \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{y}_3 \\ \hline y_1 & 1 & 0 & 1 \\ y_2 & 1 & 1 & 1 \\ x_3 & -3 & -2 & \boxed{1} \end{array} \rightarrow$$

$$\begin{array}{c|rrr} & \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \\ \hline y_1 & 4 & 2 & -1 \\ y_2 & 4 & 3 & -1 \\ y_3 & -3 & -2 & 1 \end{array}$$

Az utolsó táblázat már az $\mathbf{A}^{-1}$ mátrixot tartalmazza. □

Meg kell jegyeznünk, hogy a bázisvektorok sorrendje lényeges, ezért ha az elemi transzformációk során valamelyik x_i vektort nem az $A(x_i) = y_i$ vektorral cseréljük ki, akkor a teljes báziscsere után, sor– és oszlopcserékkel helyre kell állítanunk a bázisvektorok eredeti sorrendjét. Erre is mutatunk egy példát.

6 Példa. *Invertáljuk az*

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

mátrixot.

$$\begin{array}{c|rrr} & \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 \\ \hline x_1 & 0 & -1 & 1 \\ x_2 & \boxed{1} & 0 & 1 \\ x_3 & 1 & 1 & 1 \end{array} \rightarrow \begin{array}{c|rrr} & \mathbf{x}_2 & \mathbf{y}_2 & \mathbf{y}_3 \\ \hline x_1 & 0 & \boxed{-1} & 1 \\ y_1 & 1 & 0 & 1 \\ x_3 & -1 & 1 & 0 \end{array} \rightarrow \begin{array}{c|rrr} & \mathbf{x}_2 & \mathbf{x}_1 & \mathbf{y}_3 \\ \hline y_2 & 0 & -1 & -1 \\ y_1 & 1 & 0 & 1 \\ x_3 & -1 & 1 & \boxed{1} \end{array} \rightarrow$$

	$\mathbf{x}_2$	$\mathbf{x}_1$	$\mathbf{x}_3$
y_2	-1	0	1
y_1	2	-1	-1
y_3	-1	1	1

Az első két sor és első két oszlop cseréje után kapjuk az $\mathbf{A}$ mátrix inverzét:

$$\mathbf{A}^{-1} = \begin{bmatrix} -1 & 2 & -1 \\ 0 & -1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

$\square$

4. Fejezet

Euklideszi terek

A jegyzet első fejezetét azzal kezdtük, hogy a vektorterek elmélete a középiskolai koordináta geometria, illetve analitikus geometria általánosításának tekinthető. Mindeddig azonban nem foglalkoztunk olyan kérdésekkel, amelyek a tér elemeinek távolságával, vagy egyenesek által bezárt szögekkel kapcsolatosak, sőt azt sem tudhatjuk az eddigiekből, hogy melyek a köznapi értelemben vett 3-dimenziós tér olyan objektumainak, mint a pont, egyenes, sík stb., megfelelői az absztrakt többdimenziós vektorterekben. A valós számsíkon elhelyezkedő pontok távolságát a pontok koordinátáiból származtattuk, nevezetesen vettük a megfelelő koordináták különbségének négyzetösszegéből vont négyzetgyököt. A távolság nemnegatív valós, azon belül is esetleg irracionális számnak adódott, általában még akkor is, ha a pontok koordinátái egész, vagy racionális számok voltak, például az A(1,0) és B(0,1) pontok távolsága $\sqrt{2}$ irracionális szám. Annak érdekében, hogy megőrizhessük azt a tulajdonságot, hogy a tér elemeinek távolsága nemnegatív valós számmal kifejezhető és a tér elemeinek koordinátáiból származtatható, vizsgálatainkat leszűkítjük véges dimenziós valós vektorterekre, de ahol olyan eredményeket kapunk, amelyek érvényesek végtelen dimenziós vektorterekre is, akkor a tér véges dimenziós voltát nem hangsúlyozzuk.

4.1 Skaláris szorzatos terek

A valós számtest, mint 1-dimenziós valós vektortér, és egy egyenes pontjai között úgy létesítettünk egy-egyértelmű megfeleltetést, hogy kijelöltünk az egyenesen két pontot, az egyikhez a 0, a másikhoz az 1 valós számot rendeltük, majd ezután a λ számhoz azt a pontot rendeltük, amely az origótól, azaz a 0-nak megfelelő ponttól $|\lambda|$ távolságra van, a 0 és az 1-gyel jelölt pontok távolságát véve egységnek, az egyenesnek az origótól az 1-et tartalmazó félegyenesen, ha $\lambda > 0$, illetve az ellentétes irányú félegyenesen, ha $\lambda < 0$. Eljárhattunk volna a következőképpen is: felveszünk egy helyvektort, azaz egy irányított szakaszt, melynek a hosszát egységnyinek tekintjük, majd annak λ valós számsorosait úgy értelmezzük, hogy vesszük az ugyanabból a kezdőpontból induló, $|\lambda|$ hosszúságú, és az alapul vett helyvektor irányával megegyező, vagy azzal ellentétes iányú helyvektorokat, aszerint, hogy $\lambda > 0$, vagy $\lambda < 0$. A két eljárás ekvivalens abban az értelemben, hogy a helyvektorok végpontjai azonosak az első eljárás szerinti egyenes pontjaival. Így amikor a számegyenes pontjainak koordinátáiról beszélünk, tulajdonképpen annak a vektornak a koordinátáiról van szó, amely az origóból a szóbanforgó pontba mutat, éspedig az alapul vett helyvektorra, mint bázisra vonatkozóan. A valós számsík pontjainak koordinátái is helyvektorok koordinátái, ahol a bázist, általában egymásra merőleges, két vektor alkotja.

Ennek megfelelően a továbbiakban vizsgált vektorterek vektorait a tér pontjainak fogjuk nevezni. Értelmezni fogjuk a pontok távolságát, a pontokból képezünk félegyeneseket, definiáljuk azok hajlásszögét és megadjuk néhány más geometriai objektum megfelelőit tetszőleges n-dimenziós valós vektorterekben. A 2-dimenziós valós koordinátatérben, a valós számsíkon az

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \quad \text{és} \quad y = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$

pontok távolsága $d = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2}$. A középiskolai tanulmányaink során értelmezett skaláris szorzat segítségével ez a távolság, a két vektor különbségének önmagával alkotott skaláris szorzatából vont pozitív négyzetgyöknek adódott, tehát a skaláris szorzat lehet az az eszköz, amelyre támaszkodhatunk, a távolság értelmezésekor. Mivel nem kívánjuk vizsgálatainkat a koordinátaterekre korlátozni, a vektorok skaláris szorzatának fogalmát úgy kell definiálni, hogy egyrészt ne csak koordinátaterekben lehessen a vektorok skaláris szorzatát értelmezni, másrészt a középiskolában megismert skaláris szorzat a mi értelmezésünk szerint is skaláris szorzat maradjon. A következő definícióban a skaláris szorzat általános fogalmát adjuk meg.

4.1.1 Definíció. *Legyen V valós vektortér és*

$$\langle , \rangle : V \times V \to \mathbb{R}$$

olyan függvény, amely minden $v, w \in V$ vektorpárhoz egy $\langle v, w\rangle$-vel jelölt $\mathbb{R}$-beli skalárt rendel. A $\langle , \rangle$ függvényt skaláris szorzatnak nevezzük, ha eleget tesz az alábbi feltételeknek: bármely $v, w, z \in V$ vektorokra és tetszőleges $\lambda \in \mathbb{R}$ skalárra

(1) $\langle v, w\rangle = \langle w, v\rangle$,

(2) $\langle \lambda v, w\rangle = \lambda \langle v, w\rangle$,

(3) $\langle v + w, z\rangle = \langle v, z\rangle + \langle w, z\rangle$,

(4) $\langle v, v\rangle \geq 0$ *és* $\langle v, v\rangle = 0 \iff v = 0$.

Érdemes felfigyelni arra, hogy a skaláris szorzat értelmezésénél nincs szükség arra, hogy a V vektortér véges dimenziós legyen, például a valós együtthatós polinomok $\mathbb{R}[t]$ végtelen dimenziós terében két, $p(t), q(t)$ polinom skaláris szorzata értelmezhető a

$$\langle p(t), q(t)\rangle \overset{\text{def}}{=} \int_0^1 p(t)q(t)\, dt$$

definiáló egyenlőséggel. (Az (1)–(4) feltételek teljesülésének ellenőrzését az olvasóra bizzuk!) Vannak azonban olyan végtelen dimenziós valós vektorterek is, amelyekben nem definiálható skaláris szorzat, ilyen például az összes valós számsorozatok vektortere. Ennek az állításnak az igazolása nem könnyű és kivül esik vizsgálataink körén.

A skaláris szorzat (1)-es és (2)-es axiómáját kihasználva, könnyen kapjuk, hogy

$$\langle v, \lambda w\rangle = \langle \lambda w, v\rangle = \lambda \langle w, v\rangle = \lambda \langle v, w\rangle$$

is teljesül, azaz a második változó skalárszorzója is kiemelhető a vektorok skaláris szorzatából.

Vegyük észre azt is, hogy tetszőleges v, w és z vektorokra a

$$\langle v, w + z\rangle = \langle w + z, v\rangle = \langle w, v\rangle + \langle z, v\rangle = \langle v, w\rangle + \langle v, z\rangle$$

azonosság is igaz.

Fel kell hívjuk az olvasó figyelmét arra, hogy amennyiben a skaláris szorzatban rögzítjük az első változót, akkor az a második változójának lineáris függvénye, azaz bármely rögzített $v \in V$ vektor esetén a

$$\langle v, \cdot\rangle : V \to \mathbb{R}$$

leképezés lineáris függvénye V-nek, tehát létezik $f^* \in V^*$, hogy minden $w \in V$-re

$$\langle v, w\rangle = f^*(w)\,.$$

Teljesen hasonlóan, amennyiben a második változót rögzítjük, akkor a

$$\langle \cdot, w\rangle : V \to \mathbb{R}$$

leképezés is lineáris függvénye V-nek.

4.1.2 Definíció. *Az olyan véges dimenziós valós vektortereket, amelyekben skaláris szorzat van értelmezve* euklideszi tereknek *nevezzük.*

Bár már a mátrixaritmetikai részben értelmeztük $\mathbb{F}^n$-beli vektorok belső szorzatát, most mégis megismételjük annak definícióját, de mostmár kiterjesztve az értelmezést, tetszőleges véges dimenziós $\mathbb{F}$ test feletti vektortér vektoraira.

4.1.3 Definíció. *Legyen V n-dimenziós $\mathbb{F}$ test feletti vektortér, és annak $\mathcal{X} = \{x_1, \ldots, x_n\}$ egy rögzített bázisa. Tetszőleges*

$$v = \sum_{i=1}^{n} \varepsilon_i x_i \quad \text{és} \quad w = \sum_{i=1}^{n} \omega_i x_i$$

vektorok $\mathcal{X}$ bázis szerinti belső szorzatán a

$$[v, w]_{\mathcal{X}} \stackrel{\text{def}}{=} \sum_{i=1}^{n} \varepsilon_i \omega_i$$

összeget értjük.

Vegyük észre, hogy két vektor valamely $\mathcal{X}$ bázisra vonatkozó belső szorzata úgy kapható, hogy vesszük az első vektor koordináta vektorának, mint egyetlen oszlopból álló mátrixnak a transzponáltját és azt szorozzuk a mátrixszorzás szabályai szerint a második vektor koordináta vektorával. Tehát a v és $w \in V$ vektorok $\mathcal{X}$ bázisra vonatkozó belső szorzata $\mathbf{v}_{\mathcal{X}}^* \cdot \mathbf{w}_{\mathcal{X}}$.

Véges dimenziós valós vektorterekben belső szorzat mindig értelmezhető és az egyszersmind skaláris szorzat. Következésképpen, minden véges dimenziós valós vektortér euklideszivé tehető. Ezt bizonyítjuk a következő tételben.

4.1.4 Tétel. *Ha V véges dimenziós valós vektortér, akkor értelmezhető skaláris szorzat V-ben.*

Bizonyítás. Legyen $\mathcal{X} = \{x_1, \dots, x_n\}$ V-nek tetszőleges bázisa és bármely $v, w \in V$-re legyen

$$\langle v, w\rangle \stackrel{\text{def}}{=} [v, w]_{\mathcal{X}} = \sum_{i=1}^{n} \varepsilon_i \omega_i \,,$$

ahol ε_i, illetve ω_i $(i = 1, \dots, n)$ a v, illetve w vektorok $\mathcal{X}$ bázisra vonatkozó koordinátái. Be kell látnunk, hogy a belső szorzat kielégíti a skaláris szorzattól megkövetelt tulajdonságokat. Ezeket egszerű számolássl igazolhatjuk.

(1) $[v, w]_{\mathcal{X}} = \sum_{i=1}^{n} \varepsilon_i \omega_i = \sum_{i=1}^{n} \omega_i \varepsilon_i = [w, v]_{\mathcal{X}}$.

(2) Kihasználva, hogy

$$\lambda v = \sum_{i=1}^{n} \lambda \varepsilon_i x_i \,,$$

kapjuk, hogy

$$[\lambda v, w]_{\mathcal{X}} = \sum_{i=1}^{n} \lambda \varepsilon_i \omega_i = \lambda \sum_{i=1}^{n} \varepsilon_i \omega_i = \lambda [v, w]_{\mathcal{X}} \,.$$

(3) Ha $z = \sum_{i=1}^{n} \zeta_i$ egy tetszőleges harmadik vektor V-ben, akkor, kihasználva, hogy

$$v + w = \sum_{i=1}^{n} (\varepsilon_i + \omega_i) x_i \,,$$

és azt, hogy bármely testben a szorzás az összeadásra nézve disztributív, azt kapjuk, hogy

$$[v + w, z]_{\mathcal{X}} = \sum_{i=1}^{n} (\varepsilon_i + \omega_i) \zeta_i = \sum_{i=1}^{n} \varepsilon_i \zeta_i +$$

$$+ \sum_{i=1}^{n} \omega_i \zeta_i = [v, z]_{\mathcal{X}} + [w, z]_{\mathcal{X}}$$

(4) Mivel minden $\alpha \in \mathbb{R}$-re $\alpha^2 \geq 0$,

$$[v, v]_{\mathcal{X}} = \sum_{i=1}^{n} \varepsilon_i \varepsilon_i = \sum_{i=1}^{n} \varepsilon_i^2 \geq 0$$

és valós számok négyzeteinek összege akkor és csakis akkor nulla, ha mindegyik nulla, ezért $[v, v]_{\mathcal{X}} = 0 \Longleftrightarrow v = 0$. □

4.1.5 Definíció. *Egy V valós vektorteret normált térnek hívunk, ha létezik olyan $\|\cdot\| : V \longrightarrow \mathbb{R}$ norma-függvény, amely bármely $v \in V$-re és $\varepsilon \in \mathbb{R}$-re teljesíti a*

(1) $$\|v\| \geq 0 \quad \textit{és} \quad \|v\| = 0 \Longleftrightarrow v = 0 \,,$$

$$\|\varepsilon v\| = |\varepsilon| \|v\| \,, \tag{2}$$

$$\|v + w\| \leq \|v\| + \|w\| \tag{3}$$

feltételeket. A $\|v\|$ nemnegatív számot a v vektor normájának nevezzük.

4.1.6 Tétel. *Ha egy V valós vektortéren skaláris szorzat van értelmezve, akkor a $\|v\| \overset{\mathrm{def}}{=} \sqrt{\langle v, v \rangle}$ definiáló egyenlőséggel megadott függvény norma.*

Bizonyítás. Az, hogy $v(\neq 0) \in V$-re $\|v\| = \sqrt{\langle v, v \rangle} > 0$ közvetlen következménye a skaláris szorzat (4)-es tulajdonságának.
Ha ε tetszőleges eleme $\mathbb{R}$-nek, akkor

$$\|\varepsilon v\| = \sqrt{\langle \varepsilon v, \varepsilon v \rangle} = \sqrt{\varepsilon^2 \langle v, v \rangle} =$$

$$|\varepsilon| \sqrt{\langle v, v \rangle} = |\varepsilon| \|v\| \,.$$

A normától megkövetelt 3-dik tulajdonság teljesülését bizonyítandó, először vegyük észre, hogy

$$\|v + w\| \leq \|v\| + \|w\| \Longleftrightarrow \|v + w\|^2 \leq (\|v\| + \|w\|)^2 \,,$$

mert az egyenlőtlenség mindkét oldalán nem-negatív valós számok állnak. Kiszámítva az egyenlőtlenség mindkét oldalát:

$$\|v+w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + 2 \langle v, w \rangle + \langle w, w \rangle = \|v\|^2 + 2 \langle v, w \rangle + \|w\|^2 \,, \tag{4.1}$$

illetve

$$(\|v\| + \|w\|)^2 = \|v\|^2 + 2 \cdot \|v\| \cdot \|w\| + \|w\|^2 \,. \tag{4.2}$$

Összehasonlítva (4.1)-t és (4.2)-t kapjuk, hogy

$$\|v + w\|^2 \leq (\|v\| + \|w\|)^2 \Longleftrightarrow \langle v, w \rangle \leq \cdot \|v\| \cdot \|w\| \,.$$

Ennél egy kicsivel több is igaz, nevezetesen

$$| \langle v, w \rangle | \leq \|v\| \cdot \|w\| \,. \tag{4.3}$$

Ez utóbbi, a *Cauchy-Schwarz-egyenlőtlenség* néven ismert. Igazolása során, amelyet fontosságára való tekintettel kiemelünk, a skaláris szorzat tulajdonságait használjuk ki.

Bizonyítás. [*Cauchy-Schwarz-egyenlőtlenség*] Bármely λ valós számra a skaláris szorzat (4)–es tulajdonsága miatt, $\langle v + \lambda w, v + \lambda w \rangle \geq 0$, ami részletesebben

$$\langle v, v \rangle + 2\lambda \langle v, w \rangle + \lambda^2 \langle w, w \rangle = \|v\|^2 + 2\lambda \langle v, w \rangle + \lambda^2 \|w\|^2 \geq 0$$

alakban írható. Mivel ez λ-nak másodfokú polinomja pozitív főegyütthatóval, csak akkor lehet nemnegatív, ha ezen polinom diszkriminánsa nempozitív. (Pozitív diszkrimináns esetén két különböző, λ_1, λ_2 értéknél a polinom zérussá, míg a (λ_1, λ_2) intervallumban negatívvá válna!) Így tehát

$$4(\langle v, w\rangle)^2 - 4\|v\|^2 \cdot \|w\|^2 \leq 0\,,$$

amiből átrendezéssel, 4-gyel való egyszerűsítéssel, majd mindkét oldalból való négyzetgyök vonással kapjuk a kívánt $|\langle v, w\rangle| \leq \|v\| \cdot \|w\|$ Cauchy-Schwarz-egyenlőtlenséget.

Tekintve, hogy a normától elvárt 3-dik tulajdonság, amint azt már megmutattuk adódik a Cauchy–Schwarz–egyenlőtlenségből, ezzel a tétel bizonyítása is teljes. □

A skaláris szorzathoz rendelt norma esetén minden $v \in V$ vektorra $\langle v, v\rangle = \|v\|^2$, megjegyzendő azonban, hogy az általában nem igaz, hogy minden norma származtatható skaláris szorzatból. Ezért megkülönböztetésül a skaláris szorzatból származtatott normát *euklideszi-normának* fogjuk nevezni. Példa nem euklideszi-normára, az $\mathbb{R}^n$ valós koordinátatérben az

$$\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \longrightarrow \max_{1\leq i\leq n} |\xi_i|$$

függvény, amelyet nevezzünk *maximum-normának.* Annak igazolását, hogy ez valóban norma, az olvasóra bizzuk. Azt, hogy az így definiált norma nem euklideszi-norma $n > 1$ esetén, a következőképpen láthatjuk be: Vegyük észre, hogy euklideszi-norma esetén tetszőleges két x, y valós vektorra

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2\,,$$

tehát skaláris szorzatuk

$$\langle x, y\rangle = \frac{\|x + y\|^2 - \|x\|^2 - \|y\|^2}{2}$$

kifejezhető a norma függvényeként. Tekintsük az $x = [-1, -2, 0, \ldots, 0]$ és $y = [2, 1, 0, \ldots, 0]$ $\mathbb{R}^n$-beli vektorokat, ahol a 0 komponensek száma $n - 2$. Ha a maximum-norma skaláris szorzatból származna, akkor ezekre a vektorokra

$$\langle x, y\rangle = \frac{1 - 4 - 4}{2} = -\frac{7}{2} \quad \text{és} \quad \langle -x, y\rangle = \frac{9 - 4 - 4}{2} = \frac{1}{2}$$

teljesülne, ellentétben a skaláris szorzat (2)-es tulajdonságával. Ez a kis ellenpélda mutatja, hogy $n \geq 2$-re a maximum-norma nem euklideszi-norma.

Analízis tanulmányainkból tudjuk, hogy normált terekben definiálható *távolság-függvény*, vagy más elnevezéssel *metrika.* A metrika teszi lehetővé azoknak a topológiai alapfogalmaknak az értelmezését, ami a többváltozós analízis lineáris algebrai alapokra helyezését lehetővé teszi. A metrikától az alábbi elvárásaink vannak.

4.1.7 Definíció. *A V vektortéren értelmezet $d : V \oplus V \longrightarrow \mathbb{R}$ függvényt metrikának nevezzük, ha minden $v, w, z \in V$-re eleget tesz az alábbi feltételeknek:*

(i) $d(v, w) \geq 0$ *és* $d(v, w) = 0 \Longleftrightarrow v = w,$

(ii) $d(v, w) = d(w, v)$,

(iii) $d(v, z) \leq d(v, w) + d(w, z)$

Ezek a feltételek elég természetesek: (i) azt kívánja, hogy két pont távolsága nemnegatív legyen és csak akkor lehessen nulla, ha a két pont egybeesik. (ii) azt fejezi ki, hogy két pont távolsága független kell legyen attól, hogy azt melyik pontból kiindulva mérjük. Ez a követelmény is természetes. Végül (iii) a háromszög egyenlőtlenség azt mondja, hogy két pont között a legrövidebb út a pontokat összekötő egyenes mentén van.

Bár analízis tanulmányaink során igazoltuk, a teljesség kedvéért itt is bizonyítjuk, hogy ha egy valós vektortér normált tér, akkor abban metrika is definiálható.

4.1.8 Tétel. *Ha a valós V vektortér normált tér, akkor az a $d : V \oplus V \longrightarrow \mathbb{R}$ függvény, amely tetszőleges $v, w \in V$ pontokhoz a $d(v, w) \stackrel{\text{def}}{=} \|v - w\|$ valós számot rendeli metrika.*

Bizonyítás. (a) Mivel minden $v(\neq 0) \in V$ vektor normája pozitív, másrészt $\|0\| = \|0 \cdot v\| = |0| \cdot \|v\| = 0$, kapjuk, hogy

$$d(v, w) = \|v - w\| \geq 0 \quad \text{és}$$

$$d(v, w) = \|v - w\| = 0 \Longleftrightarrow v - w = 0 \Longleftrightarrow v = w,$$

igazolva ezzel, hogy a metrika (i) axiómája teljesül.
(b) A norma második tulajdonsága alapján

$$d(v, w) = \|v - w\| = \| - 1(w - v))\| = | - 1|\|w - v\| = \|w - v\| = d(w, v).$$

Ez mutatja, hogy a (ii) feltételnek is eleget tesz a definiált függvény.
(c) A háromszög egyenlőtlenség igazolása a norma harmadik tulajdonságából következik. Bevezetve a $v - w = x$ és $w - z = y$ jelöléseket, tekintve, hogy $v - z = x + y$, kapjuk, hogy

$$d(v, z) = \|v - z\| = \|x + y\| \leq \|x\| + \|y\| =$$

$$\|v - w\| + \|w - z\| = d(v, w) + d(w, z)\,.$$

$\square$

Mint ismeretes, egy metrikával ellátott halmazt *metrikus térnek* nevezünk, tehát az előző tételt úgy is megfogalmazhattuk volna, hogy egy normált valós vektortér metrikus tér. Az euklideszi-norma által meghatározott metrikát *euklideszi metrikának* fogjuk nevezni.
Az euklideszi terek pontjainak távolságára felső becslést biztosít a háromszög egyenlőtlenség, alsó becslést pedig, hogy a távolság nemnegatív. Időnként azonban szükség van finomabb alsó becslésre. Ezt szolgálja a következő tetszőleges metrikus terekben is igaz állítás:

4.1.9 Állítás. *Ha v, w és z az E euklideszi tér pontjai, akkor*

$$|d(v, w) - d(w, z)| \leq d(v, z)\,.$$

Bizonyítás. A háromszög egyenlőtlenség miatt, felhasználva a távolságfüggvény változóinak felcserélhetőségét is, kapjuk, hogy

$$(a)\,. \qquad d(v,w) \le d(v,z) + d(w,z) \qquad \text{és} \qquad d(w,z) \le d(v,w) + d(v,z)$$

Az (a) egyenlőtlenségek átrendezésével adódnak a

$$(b) \qquad d(v,w) - d(w,z) \le d(v,z) \qquad \text{és} \qquad d(w,z) - d(v,w) \le d(v,z)$$

egyenlőtlenségek és a (b) ekvivalens a bizonyítandó

$$|d(v,w) - d(w,z)| \le d(v,z)$$

egyenlőtlenséggel. □

Vezessük be az E euklideszi tér egy $v \neq 0$ pontján átmenő, origó kezdőpontú félegyenes fogalmát. Ezen egy $\mathbb{R}_+v$-vel jelölt ponthalmazt fogunk érteni, amelynek elemei a v pont nemnegatív valós skalárszorosai. Formalizálva:

$$\mathbb{R}_+v = \{\lambda v \ \mid \ \lambda \in \mathbb{R}, \lambda \ge 0\}\,.$$

Definiáljuk két tetszőleges $\mathbb{R}_+v$ és $\mathbb{R}_+w$ $(v \neq 0 \neq w)$ félegyenes hajlásszögét is, értve ezen azt a ϕ $(0 \le \phi \le \pi)$ szöget, amelyre

$$\cos\phi \stackrel{\text{def}}{=} \frac{\langle v,w\rangle}{\|v\| \cdot \|w\|}$$

áll fenn. A Cauchy-Schwarz-egyenlőtlenségből következik, hogy

$$-1 \le \frac{\langle v,w\rangle}{\|v\| \cdot \|w\|} \le 1\,,$$

másrészt, ha α, β pozitív valós számok, akkor

$$\frac{\langle \alpha v, \beta w\rangle}{\|\alpha v\| \cdot \|\beta w\|} = \frac{\alpha\beta\,\langle v,w\rangle}{\alpha\|v\| \cdot \beta\|w\|} = \frac{\langle v,w\rangle}{\|v\| \cdot \|w\|}\,,$$

és így, a hajlásszög fogalma jól-definiált. Hangsúlyozni kívánjuk, hogy távolságot, amint láttuk, normált terekben is mérhetünk, amihez nem feltétlenül kell legyen skaláris szorzat, de a szög bevezetéséhez használtuk a skaláris szorzatot is, és a skaláris szorzatból származtatott normát is. A továbbiakban mindig feltesszük, hogy az euklideszi térben a norma a skaláris szorzattal van értelmezve. A definícióból azonnal adódik, hogy két $\mathbb{R}_+v$ és $\mathbb{R}_+w$ félegyenes merőleges egymásra, ha a v és w pontok skaláris szorzata nulla. A derékszöget bezáró félegyeneseket meghatározó vektorokat *ortogonálisaknak* nevezzük. Itt érdemes megjegyeznünk, hogy a zéróvektornak bármely vektorral való skaláris szorzata nulla. Ez azonnal adódik a skaláris szorzat (2)-es axiómájából. Ugyanakkor a zéróvektor nem határoz meg félegyenest, így célszerű a vektorok ortogonalitását újradefiniálni, hogy a zéró vektort se kelljen kizárni a számba vehető vektorok köréből. Ezért két vektort akkor fogunk ortogonálisnak mondani, ha skaláris szorzatuk nulla. Az olyan vektorrendszereket, amelyek a zéró vektort nem tartalmazzák és vektorai páronként ortogonálisak *ortogonális rendszernek* hívjuk. A két, illetve 3-dimenziós valós terekben az ortogonális rendszerek lineárisan függetlenek, így nem meglepő a következő állítás.

4.1.10 Tétel. *Egy E euklideszi tér bármely ortogonális rendszere lineárisan független.*

Bizonyítás. Legyen $\mathcal{X} = \{x_1, \ldots, x_i, \ldots, x_n\}$ ortogonális rendszere E-nek, és tételezzük fel, hogy

$$\xi_1 x_1 + \cdots + \xi_{i-1} x_{i-1} + \xi_i x_i + \xi_{i+1} x_{i+1} + \cdots + \xi_n x_n = 0\,. \tag{4.4}$$

Képezve ezen vektoregyenlet mind baloldalán, mind jobboldalán lévő vektorának az $x_i \in \mathcal{X}$ $(1 \leq i \leq n)$ vektorral való skaláris szorzatát, kihasználva az $\mathcal{X}$ ortogonális rendszer voltát, kapjuk, hogy

$$\sum_{j=1}^{n} \langle x_i, \xi_j x_j \rangle = \xi_i \langle x_i, x_i \rangle = \xi_i \cdot \|x_i\|^2 = \langle 0, x_i \rangle = 0\,.$$

Mivel feltevésünk szerint $x_i \neq 0$, ez csak úgy lehetséges, ha $\xi_i = 0$. Tekintve, hogy x_i tetszőleges vektora volt $\mathcal{X}$-nek, ez azt jelenti, hogy az (4.4) lineáris kombináció minden skalár együtthatója nulla kell, hogy legyen, azaz az $\mathcal{X}$ ortogonális rendszer lineárisan független. □

A 2- és 3-dimenziós valós euklideszi terek Descartes-féle koordináta rendszereire gondolva, felmerül a kérdés, hogy vajon tetszőleges n-dimenziós euklideszi terekben lehet-e olyan bázist választani, amelynek vektorai páronként ortogonálisak és egységnyi hosszúságúak. Be fogjuk látni, hogy ez minden euklideszi térben lehetséges. Első lépésben ismertetjük, a *Gram-Schmidt-féle ortogonalizációs eljárást.*

4.1.11 Tétel. *Ha $\mathcal{X} = \{x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n\}$ az E euklideszi tér egy tetszőleges lineárisan független vektorrendszere, akkor van E-nek olyan $\mathcal{Y} = \{y_1, \ldots, y_{i-1}, y_i, y_{i+1}, \ldots, y_n\}$ ortogonális rendszere, amelynek minden y_i vektora az $x_1, \ldots, x_i$ vektoroknak lineáris kombinációja.*

Bizonyítás. A bizonyítás teljes indukcióval történik, ami egyúttal azt is megmutatja, hogy $\mathcal{Y}$ vektorait milyen módon konstruálhatjuk meg $\mathcal{X}$ vektoraiból. Legyen

$$y_1 = x_1\,,$$

majd képezzük az $y_2 = x_2 + \lambda y_1$ lineáris kombinációt, amelyben a λ együtthatót úgy választjuk meg, hogy az y_2 és y_1 ortogonálisak legyenek, azaz

$$\langle y_1, y_2 \rangle = \langle y_1, x_2 \rangle + \lambda \langle y_1, y_1 \rangle = \langle y_1, x_2 \rangle + \lambda \|y_1\|^2 = 0$$

teljesüljön. Ez a

$$\lambda = -\frac{\langle y_1, x_2 \rangle}{\|y_1\|^2}$$

skalár együttható választással megvalósul, tehát az

$$y_2 = x_2 - \frac{\langle y_1, x_2 \rangle}{\|y_1\|^2} y_1$$

vektor ortogonális y_1-re, és mert y_1 egyenlő volt x_1-gyel, az is igaz, hogy y_2 az x_1 és x_2 vektorok lineáris kombinációja. Legyen az az indukciós feltevésünk, hogy már sikerült úgy meghatároznunk az $y_1, \ldots, y_i$ vektorokat, hogy azok páronként ortogonálisak és mindegyik y_j $(1 \leq j \leq i)$ vektor az $x_1, \ldots, x_j$ vektorok lineáris kombinációja. Akkor a következőt

$$y_{i+1} = x_{i+1} + \sum_{j=1}^{i} \lambda_j y_j$$

alakban állítjuk elő, ahol a skalár együtthatókat annak a követelménynek megfelelően választjuk, hogy y_{i+1} ortogonális legyen az $y_1, \ldots, y_i$ vektorok mindegyikére. Az

$$\langle y_k, y_{i+1} \rangle = \langle y_k, x_{i+1} \rangle + \sum_{j=1}^{i} \lambda_j \langle y_k, y_j \rangle = \langle y_k, x_{i+1} \rangle + \lambda_k \langle y_k, y_k \rangle =$$

$$= \langle y_k, x_{i+1} \rangle + \lambda_k \|y_k\|^2 = 0 \quad (k = 1, \ldots, i)$$

feltételeket kihasználva, az adódik, hogy a

$$\lambda_k = -\frac{\langle y_k, x_{i+1} \rangle}{\|y_k\|^2} \quad (k = 1, \ldots, i)$$

választás mellett

$$y_{i+1} = x_{i+1} - \sum_{j=1}^{i} \frac{\langle y_j, x_{i+1} \rangle}{\|y_j\|^2} y_j$$

vektor ortogonális az $y_1, \ldots, y_i$ vektorok mindegyikére és az $x_1, \ldots, x_i, x_{i+1}$ vektorok lineáris kombinációja. Az eljárás n lépésben befejeződik, az n elemű lineárisan független $\mathcal{X}$ vektorrendszer elemeinek lineáris kombinációiként egy n elemű $\mathcal{Y}$ ortogonális rendszert konstruáltunk. $\square$

A fentiekben bemutatott eljárással kapott ortogonális rendszer y_1 vektora megegyezett az eredeti lineárisan független $\mathcal{X}$ vektorrendszer x_1 vektorával. Nyilvánvaló, hogy ha $\mathcal{X}$ az euklideszi tér egy bázisa, akkor az eljárással kapott $\mathcal{Y}$ ortogonális rendszer is bázis, amelyet *ortogonális bázisnak* nevezünk. Ezért kijelenthetjük:

4.1.12 Következmény. *Egy n-dimenziós E euklideszi tér bármely nemzéró vektora benne van a tér valamely ortogonális bázisában.*

A két, illetve 3-dimenziós valós tér helyvektorainak hossza ugyanaz, mint a skaláris szorzattal definiált norma, így egységnyi hosszúságú vektorok helyett normált terekben egységnyi normájú vektorokról beszélhetünk. Bármely nemzéró vektor valós skalárszorosaként kaphatunk egységnyi normájú vektort. Valóban ha $v(\neq 0)$ a normált V vektortér tetszőleges vektora, akkor

$$\| \frac{1}{\|v\|} v \| = | \frac{1}{\|v\|} | \cdot \|v\| = \frac{\|v\|}{\|v\|} = 1\,,$$

tehát a $v_0 = \frac{1}{\|v\|} \cdot v$ vektor, amelyet a v normáltjának mondunk, egységnyi normájú.

Egy n-dimenziós euklideszi tér bármely bázisából a Gram-Schmidt-féle eljárással készíthetünk ortogonális bázist, majd annak vektorait normálhatjuk. Az így kapott vektorrendszert *ortonormált bázisnak* mondjuk. Kimondhatjuk tehát az alábbi tételt:

4.1.13 Tétel. *Bármely n-dimenziós euklideszi térnek van ortonormált bázisa.*

Ortonormált bázis meghatározására bemutatunk két példát:

1 Példa. *A 3-dimenziós valós V vektortérben legyen $\mathcal{V} = \{v_1, v_2, v_3\}$ egy tetszőleges bázis és értelmezzünk skaláris szorzatot a vektorok $\mathcal{V}$ bázisra vonatkozó koordináta vektorainak belső szorzataként. Határozzunk meg olyan ortonormált bázist, amelynek vektorai előállíthatók az ugyancsak bázist alkotó $w_1 = v_1 - v_2, w_2 = v_2 - v_3, w_3 = v_3$ vektorok lineáris kombinációiként.*

Mindenekelőtt megjegyezzük, hogy az $\mathcal{V}$ bázis is ortonormált a definiált skaláris szorzat mellett, mert nyilván

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \text{és} \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

a v_1, v_2 és v_3 vektoroknak a $\mathcal{V}$ bázisra vonatkozó koordináta vektorai és ezért $[v_i, v_j]_{\mathcal{V}} = \delta_{ij}$. A feladat viszont a

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \qquad \text{és} \quad \mathbf{w}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

koordináta vektorú w_1, w_2, w_3 vektorok lineáris kombinációiból felépíteni ortonormált bázist. A Gram-Schmidt-módszerrel előbb ortogonális rendszert keresünk, majd annak vektorait normáljuk. A meghatározandó ortogonális rendszer vektorait jelölje x_1, x_2, x_3 . Akkor $x_1 = w_1$, és így $\mathbf{x}_1 = [1, -1, 0]$.

$$x_2 = w_2 - \frac{[x_1, w_2]}{\|x_1\|^2} x_1 = w_2 - \frac{-1}{2} x_1 = w_2 + \frac{1}{2} w_1 ,$$

ennélfogva $\mathbf{x}_2 = [\frac{1}{2}, \frac{1}{2}, -1]$. Végül

$$x_3 = w_3 - \frac{[x_1, w_3]}{\|x_1\|^2} x_1 - \frac{[x_2, w_3}{\|x_2\|^2} x_2 =$$

$$= w_3 - \frac{-1}{\frac{3}{2}} x_2 = w_3 + \frac{2}{3} w_2 + \frac{1}{3} w_1 ,$$

ezért $\mathbf{x}_3 = \frac{1}{3}[1, 1, 1]$. Az $\|x_1\| = \sqrt{2}$, $\|x_2\| = \frac{\sqrt{3}}{\sqrt{2}}$ és az $\|x_3\| = \frac{1}{\sqrt{3}}$. Akkor a keresett ortonormált bázis vektorai $\frac{\sqrt{2}}{2} w_1$, $\frac{\sqrt{2}}{\sqrt{3}} w_2 + \frac{1}{\sqrt{6}} w_1$, $\sqrt{3} w_3 + \frac{2}{\sqrt{3}} w_2 + \frac{1}{\sqrt{3}} w_1$, és koordináta vektoraik az eredeti $\mathcal{V}$ bázisban:

$$\mathbf{x}_1 = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \mathbf{x}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \qquad \text{és} \quad \mathbf{x}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Ezekből a koordináta vektorokból az is kiolvasható, hogy az új ortonormált bázis vektorai az eredeti $\mathcal{V}$ bázis vektorainak $x_1 = \frac{\sqrt{2}}{2}(v_1 - v_2)$, $x_2 = \frac{1}{\sqrt{6}}(v_1 + v_2 - 2v_3)$, $x_3 = \frac{1}{\sqrt{3}}(v_1 + v_2 + v_3)$ lineáris kombinációja. □

A következő példa a numerikus számításokat illetően talán nehezebb, de a megoldás elvét tekintve teljesen ugyanaz, mint az előző.

2 Példa. *A legfeljebb másodfokú valós együtthatós polinomok $\mathbb{R}_2[t]$ terét euklideszivé teszi a*

$$\langle p(t), q(t)\rangle = \int_0^1 p(t)q(t)\,dt$$

skaláris szorzat. Az $\{1, t, t^2\}$ polinomok a tér egy bázisát alkotják. E bázis vektorainak lineáris kombinációiként állítsunk elő ortonormált bázist!

A Gram-Schmidt-módszerrel előbb ortogonális bázist állítunk elő, majd annak vektorait normáljuk. $p_1(t) = 1$.

$$p_2(t) = t - \frac{\int_0^1 1\cdot t\,dt}{\int_0^1 1^2\,dt}\cdot 1 = t - \frac{1}{2}\cdot 1 = t - \frac{1}{2}$$

$$p_3(t) = t^2 - \frac{\int_0^1 1\cdot t^2\,dt}{\int_0^1 1^2\,dt}\cdot 1 - \frac{\int_0^1 (t-\frac{1}{2})t^2\,dt}{\int_0^1 (t-\frac{1}{2})^2\,dt}\cdot (t - \frac{1}{2}) =$$

$$= t^2 - \frac{1}{3}\cdot 1 - 1\cdot (t - \frac{1}{2}) = t^2 - t + \frac{1}{6}$$

Az $\{1,\ t - \frac{1}{2},\ t^2 - t + \frac{1}{6}\}$ vektorrendszer ortogonális, és az 1 már normált is. Mivel

$$\|t - \frac{1}{2}\| = \left(\int_0^1 (t - \frac{1}{2})^2\,dt\right)^{\frac{1}{2}} = \frac{\sqrt{3}}{6}$$

és

$$\|t^2 - t + \frac{1}{6}\| = \left(\int_0^1 (t^2 - t + \frac{1}{6})^2\right)^{\frac{1}{2}} = \frac{\sqrt{105}}{30}\,,$$

a keresett ortonormált bázis az

$$1,\ \sqrt{3}(2t-1),\ \frac{2\sqrt{105}}{7}(t^2 - t + \frac{1}{6})$$

vektorrendszer. □

Egy vektortér valamely altere is vektortér, és ha az eredeti tér euklideszi tér, akkor altere is euklideszi tér, amelyben a skaláris szorzat egyszerűen az eredeti térben értelmezett skaláris szorzatnak az altérre való leszűkítése. A (4.1.13) tételnél ezért kicsit többet is állíthatunk:

4.1.14 Tétel. *Egy n-dimenziós euklideszi tér minden nemzéró alterének van ortonormált bázisa.*

4.1.15 Tétel. *Ha M részhalmaza az E euklideszi térnek, akkor jelölje $M^\perp$ azoknak az E-beli vektoroknak a halmazát, amelyek ortogonálisak M minden elemére. Akkor $M^\perp$ altere E-nek.*

Bizonyítás. Az állítás igazolásakor két esetet kell vizsgálnunk:
(a) ha M az üres halmaz, akkor $M^\perp = E$, így ebben az esetben igaz az állítás,
(b) ha M nemüres, akkor legyen v tetszőleges eleme M-nek és x, y pedig tetszőleges elemei $M^\perp$-nak. Bármely α, β skalárokkal képzett $\alpha x + \beta y$ lineáris kombinációra

$$\langle \alpha x + \beta y, v \rangle = \alpha \langle x, v \rangle + \beta \langle y, v \rangle = \alpha 0 + \beta 0 = 0\,,$$

így $\alpha x + \beta y \in M^\perp$ is teljesül, igazolva, hogy $M^\perp$ altér. □

Ha egy x vektor ortogonális mind a v, mind a w vektorra, akkor x ortogonális a v és w vektorok bármely lineáris kombinációjára is, mert

$$\langle \varepsilon v + \omega w, x \rangle = \varepsilon \langle v, x \rangle + \omega \langle w, x \rangle = \varepsilon 0 + \omega 0 = 0\,.$$

Ezért egy tetszőleges M részhalmazra $M^\perp = (\operatorname{lin}(M))^\perp$. Az M részhalmaz, sőt az általa generált altér, részhalmaza az $M^{\perp\perp} = (M^\perp)^\perp$ altérnek. Amennyiben M maga is altér, akkor $M^\perp$-t az M *ortogonális komplementerének* nevezzük. Az elnevezést alátámasztja a következő, úgynevezett projekciós tétel.

4.1.16 Tétel. **[Projekciós tétel]** *Ha M altere a véges dimenziós E euklideszi térnek, akkor E az M és $M^\perp$ altereinek direktösszege.*

Bizonyítás. Legyen $\mathcal{X} = \{x_1, \ldots, x_r\}$ az M-nek, és $\mathcal{Y} = \{y_1, \ldots, y_s\}$ az $M^\perp$-nak ortonormált bázisa. Megmutatjuk, hogy $\mathcal{X} \cup \mathcal{Y}$ ortonormált bázisa E-nek. Mivel $\mathcal{X} \cup \mathcal{Y}$ ortogonális rendszer, így lineárisan független az (4.1.10) tétel értelmében. Így már csak azt kell igazolnunk, hogy tetszőleges $v \in E$ vektor az $\mathcal{X} \cup \mathcal{Y}$ elemeinek lineáris kombinációja. Legyen

$$x = \sum_{i=1}^{r} \alpha_i x_i$$

ahol $\alpha_i = \langle x_i, v \rangle$. Akkor, kihasználva, hogy $\mathcal{X}$ vektorai páronként ortogonálisak és normáltak, kapjuk minden $(k = 1, \ldots r)$-ra, hogy

$$\langle x_k, v - x \rangle = \langle x_k, v \rangle - \langle x_k, v \rangle \|x_k\|^2 = \langle x_k, v \rangle - \langle x_k, v \rangle = 0\,,$$

de akkor $v - x$ az M minden elemére ortogonális, vagyis az $y = v - x$ vektor $M^\perp$-ben van, ezért $\mathcal{Y}$ elemeinek

$$y = \sum_{j=1}^{s} \beta_j y_j$$

lineáris kombinációja. Ebből következik, hogy

$$v = x + y = \sum_{i=1}^{r} \alpha_i x_i + \sum_{j=1}^{s} \beta_j y_j\,,$$

tehát valóban az $\mathcal{X} \cup \mathcal{Y}$ vektorrendszer generátorrendszere is E-nek. □

A tétel igazolásához elegendő lett volna azt megmutatni, hogy $M \cap M^\perp$ a zérus altér és E minden vektora egy M-beli és egy $M^\perp$-beli vektor összege. Bizonyításunk verifikálja a következő tételt is.

4.1.17 Tétel. *Egy véges dimenziós euklideszi tér bármely M alterének ortonormált bázisa kiegészíthető az egész tér ortonormált bázisává.*

Azon túl, hogy a továbbiakban szükségünk lesz rá, még esztétikailag is szép, hogy igaz a Pythagorasz-tétel tetszőleges véges dimenziós euklideszi terekre való általánosítása:

4.1.18 Tétel. [Pythagorasz-tétele] *Ha v és w a véges dimenziós E euklideszi tér egymásra ortogonális vektorai, akkor*

$$\|v+w\|^2 = \|v\|^2 + \|w\|^2 .$$

Bizonyítás. Mivel v és w ortogonálisak, $\langle v, w\rangle = 0$, ezért

$$\|v+w\|^2 = \langle v+w, v+w\rangle = \langle v, v\rangle + 2\mathrm{Re}(\langle v, w\rangle) + \langle w, w\rangle = \|v\|^2 + \|w\|^2 .$$

□

4.1.19 Definíció. *Ha v az E euklideszi tér tetszőleges pontja, és $\mathcal{X} = \{x_1, \ldots, x_r\}$ az E tér egy M alterének ortonormált bázisa, akkor az*

$$x = \sum_{i=1}^{r} \langle x_i, v\rangle\, x_i$$

pontot a v M-re eső ortogonális vetületének nevezzük.

Amint a (4.1.16) tétel bizonyításában láttuk, $v-x$ az M altér minden vektorára ortogonális. Megmutatjuk azt is, hogy az x pont M-nek v-hez legközelebb eső pontja. Legyen $x' = \sum_{i=1}^{r} \xi_i x_i$ egy tetszőleges M-beli pont, akkor $x - x' \in M$, ezért ortogonális $v-x$-re. Felhasználva Pythagorasz tételét, kapjuk, hogy

$$d^2(v, x') = \|v-x'\|^2 = \|(v-x) + (x-x')\|^2 =$$

$$= \|v-x\|^2 + \|x-x'\|^2 = d^2(v, x) + \|x-x'\|^2 ,$$

amiből már $d(v, x) \le d(v, x')$ nyilvánvalóan következik.

Analízis tanulmányaink során egy (X, d) metrikus tér valamely p pontjának egy H részhalmazától való távolságát a

$$d(p, H) = \inf_{h \in H} d(p, h)$$

összefüggéssel értelmeztük. Előző számításaink szerint, euklideszi tér valamely v pontjának egy M alterétől való távolságát a

$$d(v, M) = \|v-x\|$$

formula adja, ahol az x pont a v-nek M-re eső ortogonális vetülete.

3 Példa. *Legyen az E euklideszi tér egy ortonormált bázisa az $\mathcal{X} = \{x_1, x_2, x_3\}$ vektorrendszer. Határozzuk meg a $v = x_1 + 2x_2 - x_3$ vektornak az $a = x_1 + x_2 + x_3$ és $b = x_2 + 2x_3$ vektorok által generált M altértől való távolságát!*

Első lépésben meg kell keresnünk az altér egy ortonormált bázisát. Legyen $y_1 = a$ és

$$y_2 = b - \frac{\langle y_1, b\rangle}{\|y_1\|^2} y_1 = x_3 .$$

Akkor y_1 és y_2 az M altér ortogonális, míg az $\frac{1}{\sqrt{3}}y_1$ és y_2 pedig ortonormált bázisa. Akkor a v vektor M-re eső ortogonális vetülete:

$$z = \left\langle \frac{1}{\sqrt{3}}y_1, v \right\rangle \frac{1}{\sqrt{3}}y_1 + \langle y_2, v\rangle\, y_2 \,.$$

Kifejezve a z vektort az $\mathcal{X}$ bázis vektoraival, kapjuk, hogy

$$z = \frac{2}{3}(x_1 + x_2 + x_3) - x_3 = \frac{2}{3}x_1 + \frac{2}{3}x_2 - \frac{1}{3}x_3 \,,$$

és így

$$v - z = \frac{1}{3}x_1 + \frac{4}{3}x_2 - \frac{4}{3}x_3 \,,$$

amelynek

$$\|v - z\| = \frac{\sqrt{33}}{3}$$

normája a keresett távolság. □

Azt már előbb láttuk, hogy a valós véges dimenziós vektorterekben a tér valamely bázisára vonatkozó belső szorzat mindig skaláris szorzatot definiál. Most megmutatjuk, hogy a fordított állítás is igaz, nevezetesen, hogy az euklideszi terekben a skaláris szorzat mindig valamely ortonormált bázisra vonatkozó belső szorzat.

4.1.20 Tétel. *Ha az n-dimenziós E euklideszi térnek $\mathcal{X} = \{x_1, \ldots, x_n\}$ egy tetszőleges ortonormált bázisa, akkor*
(1) *bármely $v \in E$ vektor a bázisvektorok*

$$v = \langle x_1, v\rangle\, x_1 + \cdots + \langle x_n, v\rangle\, x_n$$

alakú lineáris kombinációja, és
(2) *tetszőleges $v, w \in E$ vektorokra*

$$\langle v, w\rangle = [v, w]_{\mathcal{X}} \,.$$

Más szavakkal megfogalmazva az állítást azt mondhatjuk, hogy egy euklideszi tér ortonormált koordináta rendszerében a tér vektorai felbomlanak a tengelyekre eső ortogonális vetületeik összegére, ugyanúgy, ahogy a két– illetve 3-dimenziós Descartes-féle koordináta rendszerben láttuk.

Bizonyítás. Azt tudjuk, hogy bármely $v \in E$ vektor kifejezhető $\mathcal{X}$ vektorainak

$$v = \sum_{j=1}^{n} \varepsilon_j x_j$$

lineáris kombinációjaként, tehát csak azt kell megmutatnunk, hogy minden $i(= 1, \ldots, n)$-re $\varepsilon_i = \langle x_i, v\rangle$ teljesül. Ezt kapjuk, ha összehasonlítjuk az alábbi

$$\langle x_i, v\rangle = \left\langle x_i, \sum_{j=1}^{n} \varepsilon_j x_j \right\rangle = \sum_{j=1}^{n} \varepsilon_j \langle x_i, x_j\rangle = \varepsilon_i$$

egyenlőségsorozat bal– és jobboldalát.

Legyen v és w két tetszőleges vektora E-nek. Akkor az előzőek szerint

$$v = \sum_{i=1}^{n} \langle x_i, v\rangle\, x_i \quad \text{és} \quad w = \sum_{j=1}^{n} \langle x_j, w\rangle\, x_j\,,$$

azaz koordináta vektoraik az $\mathcal{X}$ bázisban:

$$\mathbf{v}_{\mathcal{X}} = \begin{bmatrix} \langle x_1, v\rangle \\ \vdots \\ \langle x_n, v\rangle \end{bmatrix} \qquad \text{és} \qquad \mathbf{w}_{\mathcal{X}} = \begin{bmatrix} \langle x_1, w\rangle \\ \vdots \\ \langle x_n, w\rangle \end{bmatrix}.$$

Akkor

$$\langle v, w\rangle = \left\langle \sum_{i=1}^{n} \langle x_i, v\rangle\, x_i, \sum_{j=1}^{n} \langle x_j, w\rangle\, x_j \right\rangle =$$

$$= \left(\sum_{i=1}^{n} \langle x_i, v\rangle \sum_{j=1}^{n} \langle x_j, w\rangle \right) \langle x_i, x_j\rangle = \sum_{i=1}^{n} \langle x_i, v\rangle \langle x_i, w\rangle = [v, w]_{\mathcal{X}}\,.$$

Az egyenlőségsorozat bal– és jobboldalát összehasonlítva azt kapjuk, hogy $\langle v, w\rangle = [v, w]_{\mathcal{X}}$, amint állítottuk. □

Az (4.1.20) tétel kényelmes eszközt biztosít a tér vektorai skaláris szorzatának numerikus meghatározására. Ezt kihasználjuk a következő feladat megoldásakor, amikor egy vektor koordináta vektorát kell meghatároznunk egy ortonormált bázisban.

4 Példa. *Az (1) példában az $\mathbb{R}^3$ koordinátateret euklideszivé tettük. A skaláris szorzatot a vektorok $\mathcal{V} = \{v_1, v_2, v_3\}$ bázisra vonatkozó koordináta vektorainak belső szorzatával definiáltuk. Ezt követően áttértünk egy másik $\mathcal{X} = \{x_1 = \frac{\sqrt{2}}{2}(v_1 - v_2),\ x_2 = \frac{1}{\sqrt{6}}(v_1 + v_2 - 2v_3),\ x_3 = \frac{1}{\sqrt{3}}(v_1 + v_2 + v_3)\}$ ortonormált bázisra. Határozzuk most meg a $z = 2v_1 - v_2 + 3v_3$ vektor koordináta vektorát az $\mathcal{X}$ bázisban.*

Az (4.1.20) tétel szerint

$$\mathbf{z}_{\mathcal{X}} = \begin{bmatrix} \langle z, x_1\rangle \\ \langle z, x_2\rangle \\ \langle z, x_3\rangle \end{bmatrix} = \begin{bmatrix} \left\langle 2v_1 - v_2 + 3v_3, \frac{\sqrt{2}}{2}(v_1 - v_2)\right\rangle \\ \left\langle 2v_1 - v_2 + 3v_3, \frac{1}{\sqrt{6}}(v_1 + v_2 - 2v_3)\right\rangle \\ \left\langle 2v_1 - v_2 + 3v_3, \frac{1}{\sqrt{3}}(v_1 + v_2 + v_3)\right\rangle \end{bmatrix}$$

Esetünkben ezek a skaláris szorzatok:

$$\left\langle 2v_1 - v_2 + 3v_3, \frac{\sqrt{2}}{2}(v_1 - v_2)\right\rangle = [2, -1, 3] \cdot \frac{\sqrt{2}}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \frac{3\sqrt{2}}{2}\,,$$

$$\left\langle 2v_1 - v_2 + 3v_3, \frac{1}{\sqrt{6}}(v_1 + v_2 - 2v_3)\right\rangle = [2, -1, 3] \cdot \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = -\frac{5}{\sqrt{6}}\,,$$

és

$$\left\langle 2v_1 - v_2 + 3v_3, \frac{1}{\sqrt{3}}(v_1 + v_2 + v_3) \right\rangle = [2, -1, 3] \cdot \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{4}{\sqrt{3}}.$$

Tehát a keresett koordináta vektor:

$$\mathbf{z}_{\mathcal{X}} = \begin{bmatrix} \frac{3\sqrt{2}}{2} \\ -\frac{5}{\sqrt{6}} \\ \frac{4}{\sqrt{3}} \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 9\sqrt{2} \\ -5\sqrt{6} \\ 8\sqrt{3} \end{bmatrix}.$$

□

Az azonos dimenziójú valós vektorterek mind izomorfak, tehát algebrai szempontból csak vektoraik jelölésében térnek el egymástól. Ugyanakkor, az (4.1.4) tétel szerint ugyanabban a V véges dimenziós valós vektortérben több különböző skaláris szorzat is értelmezhető. Jelenti vajon ez azt, hogy több különböző n-dimenziós euklideszi tér van? Célszerű két euklideszi teret lényegében azonosnak tekintenünk, ha mint vektorterek izomorfak és az egymásnak megfeleltetett pontok távolsága azonos. Mivel az euklideszi tér pontjainak távolsága a térben értelmezett skaláris szorzattól függ, ezért az euklideszi terek izomorfiáját a következőképpen értelmezzük:

4.1.21 Definíció. *Az E és E' euklideszi tereket izomorfaknak mondjuk, ha*
(1) *mint vektorterek izomorfak, és*
(2) *bármely $v, w \in E$-re a $\langle v, w\rangle$ skaláris szorzat megegyezik a $v', w' \in E'$ izomorf képek (E'-beli) $\langle v', w'\rangle$ skaláris szorzatával.*

Az euklideszi terek izomorfiája erősebb megszorítás, mint a vektorterek izomorfiája, hiszen teljesülni kell annak a feltételnek is, hogy a tárgyvektorok skaláris szorzata meg kell egyezzen képeik skaláris szorzatával. Ez a feltétel azt jelenti, hogy az euklideszi terek izomorf leképezései távolságtartóak, ezért szokás *izometriának* is nevezni. (Az más kérdés, hogy ez nem jelenti a két térben a távolságegység egyenlő voltát. Tehát lehet, hogy az egyik térben a távolságegység mondjuk a méter, míg a másikban mondjuk a yard. Akkor az 1 méterre lévő pontok izomorf képei 1 yard távolságra vannak. Hasonló értelemben értendő, hogy az euklideszi terek izomorf leképezése megtartja a félegyenesek hajlásszögét is.) Ennek ellenére kicsit meglepő, hogy

4.1.22 Állítás. *Az azonos dimenziójú euklideszi terek izomorfak.*

Bizonyítás. Legyenek E és E' n-dimenziós euklideszi terek és $\mathcal{X} = \{v_1, \ldots, v_n\}$ illetve $\mathcal{X}' = \{v_1', \ldots, v_n'\}$ ortonormált bázisok E-ben, illetve E'-ben. Értelmezük a

$$\Phi : E \rightleftharpoons E'$$

leképezést úgy, hogy legyen $\Phi(v_i) = v_i'$ minden $i(= 1, \ldots, n)$-re, és ha $v \in E$ az $\mathcal{X}$ bázis vektorainak

$$v = \sum_{i=1}^{n} \varepsilon_i v_i$$

lineáris kombinációja, akkor legyen

$$\Phi(v) = \sum_{i=1}^{n} \varepsilon_i v_i'.$$

Akkor nyilván Φ az E vektortérnek E'-re való izomorf leképezése és mivel az E-beli, illetve az E'-beli skaláris szorzat az $\mathcal{X}$ bázisra vonatkozó, illetve az $\mathcal{X}'$ bázisra vonatkozó belső szorzatok, tetszőleges v, $w \in E$ vektorok skaláris szorzata megegyezik a $\Phi(v)$, $\Phi(w) \in E'$ vektorok skaláris szorzatával. $\square$

A fenti állításnak megfelelően az algebrai szempontból egyetlen n-dimenziós euklideszi teret a következőkben E_n-nel fogjuk jelölni és feltételezzük, hogy az E_n-ben rögzített egy $\mathcal{X}$ ortonormált bázis. Ez lehetővé teszi, hogy az E_n-beli pontokat koordináta vektoraikkal adjuk meg, és pontjaik skaláris szorzatát egyszerűen az $\mathcal{X}$ bázis által meghatározott belső szorzattal számíthassuk ki.

Az (4.1.20) tételnek és az (4.1.22) állításnak egy nagyon fontos következménye, hogy az euklideszi terek olyan bázistranszformációi, amelyek ortonormált bázist ortonormált bázisba visznek, megtartják a tér pontjainak távolságát és a tér félegyeneseinek hajlásszögét. Ezeknek a transzformációknak a vizsgálatára a későbbiekben még visszatérünk.

1. Gyakorlatok, feladatok

1. Állapítsa meg, hogy mi annak a szükséges és elegendő feltétele, hogy az E_n euklideszi tér v, w pontjaira teljesüljön az
$$\langle v, w\rangle = \|v\| \cdot \|w\|$$
egyenlőség.

2. Mutassa meg, hogy egy E_n euklideszi tér bármely v vektorának euklideszi-normája nem kisebb mint valamely ortonormált bázisa alapján kiszámított maximum-normája.

3. Mutassa meg, hogy az $m \times n$ tipusú valós mátrixok terében skaláris szorzatot kapunk, ha tetszőleges $\mathbf{A} = [\alpha_{ij}]$ és $\mathbf{B} = [\beta_{ij}]$ mátrixpárhoz az
$$\langle \mathbf{A}, \mathbf{B}\rangle \stackrel{\text{def}}{=} \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_{ij}\beta_{ij}$$
valós számot rendeljük.

4. Igazolja, hogy egy E_n euklideszi tér lineáris transzformációinak $L(E_n)$ tere is euklideszivé tehető. (Tanács: Lássa be, hogy, ha $\mathcal{X} = \{x_1, \ldots, x_n\}$ az E_n egy ortonormált bázisa, akkor az
$$\langle A, B\rangle \stackrel{\text{def}}{=} \sum_{i=1}^{n} \langle A(x_i), B(x_i)\rangle$$
egyenlőséggel definiált függvény skaláris szorzat az $L(E_n)$ téren.)

5. Legyen az E_n euklideszi térnek M egy altere. Mutassa meg, hogy egy $v \in E_n$ vektor pontosan akkor ortogonális az M altér minden vektorára, ha ortogonális az M altér egy bázisának minden vektorára.

6. Határozza meg az E_4 euklideszi tér v pontjának az M alterétől való távolságát, ha v koordináta vektora az $\mathcal{X}$ ortonormált bázisban
$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \quad \text{és} \quad M = \{x \in E_4 \mid [x, s]_{\mathcal{X}} = 0\}\,, \quad \text{ahol } \mathbf{s} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

7. Határozza meg az α paraméter értékét ha tudjuk, hogy a $p(t) = 1 + t + t^2$ és $q(t) = \alpha t - t^2$ polinomok ortogonálisak a
$$\langle p(t), q(t) \rangle \stackrel{\text{def}}{=} \int_{-1}^{1} p(t)q(t)\, dt$$
egyenlőséggel értelmezett skaláris szorzat mellett.

4.2 A transzponált lineáris leképezés

A mátrixaritmetikai bevezetőben megismerkedtünk a transzponált mátrix fogalmával. Egy $\mathbf{A}$ mátrix transzponáltja azt az $\mathbf{A}^*$ mátrixot jelentette, amelyet az $\mathbf{A}$ sorainak és oszlopainak felcserélésével kapunk. A mátrixok és a lineáris leképezések között szoros a kapcsolat, így aligha meglepő, hogy a lineáris leképezések transzponáltja is definiált fogalom. Itt mi csupán a fogalom bevezetésére és néhány — a későbbiek szempontjából nélkülözhetetlen — eredmény ismertetésére szorítkozunk.

Amint az jól ismert, véges dimenziós valós vektorterekben mindig értelmezhető skaláris szorzat, ezért az ilyen vektorterek lineáris leképezéseinek vizsgálatakor is célszerű ezt figyelembe venni. Ha tehát $A \in L(V, W)$ lineáris leképezés a valós véges dimenziós V vektortérből az ugyancsak valós véges dimenziós W vektortérbe, akkor mind V-t, mind W-t euklideszi tereknek tekinthetjük, amelyekben a skaláris szorzat értelmezve van.

4.2.1 Definíció. *Legyenek V és W euklideszi terek és $A \in L(V, W)$ lineáris leképezés. Az A transzponáltjának nevezzük azt az $A^* : W \to V$ leképezést, amelyre minden $v \in V$, $w \in W$ vektorra*
$$\langle A(v), w \rangle_W = \langle v, A^*(w) \rangle_V$$
teljesül.

A definícióval kapcsolatban hangsúlyoznunk kell, hogy az egyenlőség baloldalán W-beli vektorok skaláris szorzata, míg a jobboldalon V-beli vektorok skaláris szorzata áll. Erre utalnak a skaláris szorzat jele mellett feltüntetett indexek. A definícióból nem látszik közvetlenül, ezért külön igazoljuk, hogy

4.2.2 Tétel. *Az $A \in L(V, W)$ lineáris leképezés transzponáltja egyértelműen meghatározott és $A^* \in L(W, V)$.*

Bizonyítás. *Unicitás:* Tegyük fel, hogy A^* és $\bar{A}^*$ is transzponáltja A-nak. Akkor minden $v \in V$, $w \in W$ vektorokra
$$\langle v, A^*(w) \rangle_V = \langle A(v), w \rangle_W = \langle v, \bar{A}^*(w) \rangle_V\,,$$

azaz
$$\langle v, A^*(w)\rangle_V - \langle v, \bar{A}^*(w)\rangle = \langle v, A^*(w) - \bar{A}^*(w)\rangle_V = 0\,.$$
Az utóbbi egyenlőség azt jelenti, hogy $A^*(w) - \bar{A}^*(w)$ V minden vektorára ortogonális, ami csak úgy lehet, ha $A^*(w) - \bar{A}^*(w) = 0$. Mivel ez minden $w \in W$-re teljesül, következik, hogy $A^* = \bar{A}^*$.
A transzponált lineáris leképezés: tetszőleges $w_1, w_2 \in W$-re és $\alpha, \beta \in \mathbb{R}$-re és minden $v \in V$-re
$$\langle v, A^*(\alpha w_1 + \beta w_2)\rangle_V = \langle A(v), \alpha w_1 + \beta w_2\rangle_W =$$
$$\alpha \langle A(v), w_1\rangle_W + \beta \langle A(v), w_2\rangle_W =$$
$$\alpha \langle v, A^*(w_1)\rangle_V + \beta \langle v, A^*(w_2)\rangle_V = \langle v, \alpha A^*(w_1) + \beta A^*(w_2)\rangle_V\ .$$
Az egyenlőségsorozat bal- és jobboldalát összehasonlítva, az unicitás bizonyításánál látott érveléssel adódik, hogy
$$A^*(\alpha w_1 + \beta w_2) = \alpha A^*(w_1) + \beta A^*(w_2)\,,$$
azaz A^* lineáris leképezés. □

Vizsgáljuk meg, hogy milyen kapcsolat van az A lineáris leképezés mátrixa és A^* transzponáltjának mátrixa között. Ehhez rögzítenünk kell mind a V-ben egy $\mathcal{X} = \{x_1, \ldots, x_n\}$, mind a W-ben egy $\mathcal{Y} = \{y_1, \ldots, y_m\}$ bázist. Mint tudjuk, A mátrixa az $\mathcal{X} - \mathcal{Y}$ bázispárra vonatkozóan: $\mathbf{A} = [\mathbf{A(x_1)}_{\mathcal{Y}}, \ldots, \mathbf{A(x_n)}_{\mathcal{Y}}]$, azaz a mátrix i-edik sorának j-edik eleme az $A(x_j)$ vektor y_i-re vonatkozó koordinátája. Amennyiben az $\mathcal{Y}$ bázis ortonormált, akkor az
$$\langle A(x_j), y_i\rangle_W$$
skaláris szorzat éppen ezt a koordinátát szolgáltatja, tehát
$$\mathbf{A} = \begin{bmatrix} \langle A(x_1), y_1\rangle_W\,, & \ldots, & \langle A(x_n), y_1\rangle_W \\ \langle A(x_1), y_2\rangle_W\,, & \ldots, & \langle A(x_n), y_2\rangle_W \\ \vdots & & \vdots \\ \langle A(x_1), y_m\rangle_W\,, & \ldots, & \langle A(x_n), y_m\rangle_W \end{bmatrix}$$
alakú. Ha az $\mathcal{X}$ bázis is ortonormált V-ben, akkor az A^* mátrixa az $\mathcal{Y}-\mathcal{X}$ bázispárra vonatkozóan
$$\mathbf{A}^* = [\mathbf{A}^*(\mathbf{y_1})_{\mathcal{X}}, \ldots, \mathbf{A}^*(\mathbf{y_m})_{\mathcal{X}}] = \begin{bmatrix} \langle x_1, A^*(y_1)\rangle_V\,, & \ldots, & \langle x_1, A^*(y_m)\rangle_V \\ \langle x_2, A^*(y_1)\rangle_V\,, & \ldots, & \langle x_2, A^*(y_m)\rangle_V \\ \vdots & & \vdots \\ \langle x_n, A^*(y_1)\rangle_V\,, & \ldots, & \langle x_n, A^*(y_m)\rangle_V \end{bmatrix}$$
alakú. Látható, hogy $\mathbf{A}^*$ j-edik sorának i-edik eleme
$$\langle x_j, A^*(y_i)\rangle_V = \langle A(x_j), y_i\rangle_W$$
megegyezik az $\mathbf{A}$ mátrix j-edik oszlopának i-edik elemével, tehát A^* mátrixa éppen A mátrixának transzponáltja.

A lineáris leképezések és transzponáltjaik közötti kapcsolat további jellemzésére bebizonyítjuk az alábbi tételt.

4.2.3 Tétel. *Legyenek V és W euklideszi terek és $A \in L(V,W)$ transzponáltja $A^* \in L(W,V)$. Akkor*

$$(\operatorname{im} A)^\perp = \ker A^* \qquad \textit{és} \qquad (\operatorname{im} A^*)^\perp = \ker A\,,$$

következésképpen

$$V = \ker A \oplus \operatorname{im} A^* \qquad \textit{és} \qquad W = \ker A^* \oplus \operatorname{im} A\,.$$

Bizonyítás. Minden $v \in V$-re és $w \in \ker A^*$-ra

$$\langle A(v), w\rangle_W = \langle v, A^*(w)\rangle_V = \langle v, 0\rangle_V = 0 \Longrightarrow w \in (\operatorname{im} A)^\perp\,.$$

Másrészt, ha minden $v \in V$-re

$$0 = \langle A(v), w\rangle_W = \langle v, A^*(w)\rangle_V \Longrightarrow A^*(w) \in V^\perp = \{0\}\,,$$

azaz $w \in \ker A^*$. Ez igazolja az első állítást, amelynek a második állítás analogonja. Ezek ismeretében a $V = \ker A \oplus \operatorname{im} A^*$ és a $W = \ker A^* \oplus \operatorname{im} A$ állítások azonnal adódnak a projekciós tételből. □

4.2.4 Definíció. *Legyen V euklideszi tér. Az $A \in L(V)$ lineáris transzformációt szimmetrikusnak nevezzük, ha megegyezik transzponáltjával. Egy reguláris $B \in R(V)$ lineáris transzformációt ortogonálisnak mondjuk, ha transzponáltja az inverzével egyenlő.*

Az előző tétel szimmetrikus lineáris transzformációkra tehát így hangzik:

4.2.5 Következmény. *Ha A szimmetrikus lineáris transzformációja a V euklideszi térnek, akkor*

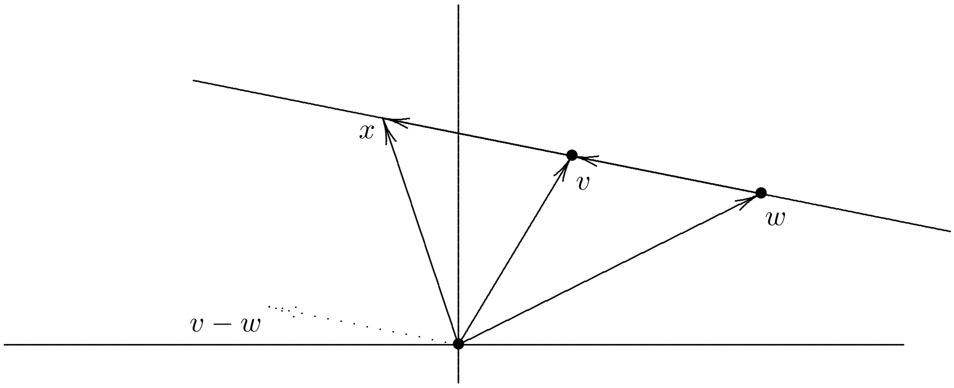
$$(\operatorname{im} A)^\perp = \ker A \qquad \textit{és} \qquad V = \ker A \oplus \operatorname{im} A\,.$$

4.3 Geometriai fogalmak általánosítása

Ebben a pontban geometriai fogalmak tetszőleges euklideszi terekre való kiterjesztésével foglalkozunk. Persze nem vállalkozhatunk minden, két és 3-dimenziós geometriai fogalom tetszőleges n-dimenziós euklideszi terekben való általánosításának megfogalmazására, csak a legfontosabbakra térünk ki.

A 2- és 3-dimenziós tér bármely két v és w pontja (vektora) meghatározott egy egyenest, amelynek bármely x pontja (vektora) megkapható, ha a w vektorhoz hozzáadjuk a $v - w$ vektor valamilyen λ skalárszorosát. Tehát a v és w pontokon átmenő egyenes minden pontja

$$x = w + \lambda(v - w) = \lambda v + (1 - \lambda)w\,, (\lambda \in \mathbb{R})$$



4.1. ábra: Egyenes a síkban

alakú lineáris kombinációja a v és w pontoknak. Egész pontosan: ha λ végigfut az összes valós számon, akkor x végigfut az egyenes pontjain. Az alábbi 4.1 ábra mutatja a kétdimenziós tér egy egyenesének ilyen megadását.

Ez sugallja, hogy tetszőleges E euklideszi tér v és w pontján átmenő *egyenesén* az

$$\overleftrightarrow{vw} = \{x \in E_n \mid x = \lambda v + (1-\lambda)w \ \ \lambda \in \mathbb{R}\} = w + \mathbb{R}(v-w)$$

ponthalmazt értsük.
A 2-dimenziós térben a szakaszok az egyeneseknek korlátos zárt részhalmazai, ha csak a v és w pontokat összekötő szakasz pontjait akarjuk megkapni, akkor a λ skalárt a valós számok $[0,1]$ zárt intervallumán kell végigfuttatnunk. Ennek megfelelően az n-dimenziós euklideszi tér v és w pontjait összekötő *szakaszán* a

$$\overline{v,w} = \{x \in E_n \mid x = \lambda v + (1-\lambda)w, \quad 0 \le \lambda \le 1\}$$

ponthalmazt fogjuk érteni.

A szakasz pontjai a végpontok olyan lineáris kombinációja, hogy az együtthatók nemnegatívak és összegük 1-gyel egyenlő. Az ilyen lineáris kombinációkat *konvex lineáris kombinációknak* nevezzük. Általánosan, ha $\lambda_i \ge 0\,, \quad (i = 1,\ldots,n)$ valós számok és $\sum_{i=1}^n \lambda_i = 1\,,$ akkor a $v_i\,, \quad (i = 1,\ldots,n)$ vektorok $\sum_{i=1}^n \lambda_i v_i$ lineáris kombinációját konvex lineáris kombinációnak mondjuk.

4.3.1 Definíció. *Az E euklideszi tér olyan K részhalmazát, amely bármely két pontjának konvex lineáris kombinációit is tartalmazza, konvex halmaznak hívjuk.*

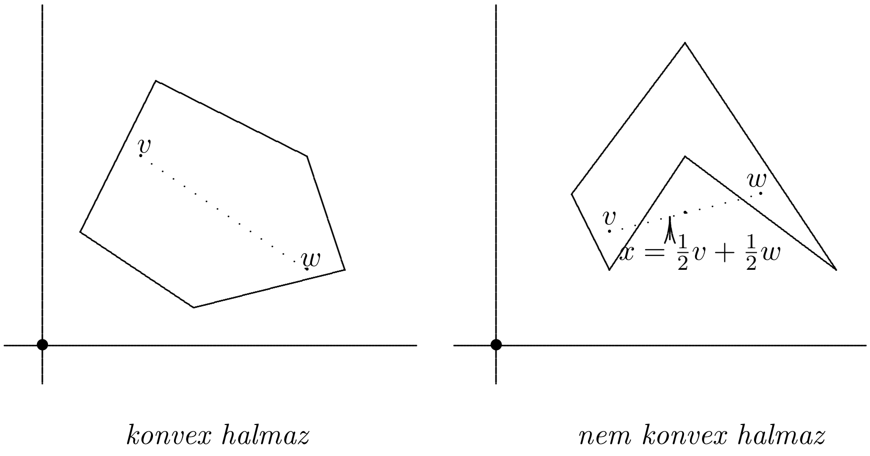
Más szavakkal, azt mondhatjuk, hogy K pontosan akkor konvex halmaz, ha bármely két pontját összekötő szakasz részhalmaza K-nak.

Persze, ha két pont egybeesik, akkor az azokat "összekötő szakasz" egyetlen pontra zsugorodik, ezért az egyetlen pontból álló halmazok is konvexek, sőt megállapodunk abban, hogy az üres halmazt is konvexnek mondjuk.

A sík egy-egy konvex és nem konvex részhalmazát mutatja az 4.2 ábra.

A konvex halmazok persze nemcsak bármely két pontjuknak, de bármely véges sok pontjuk konvex lineáris kombinációit is tartalmazzák. Ez akár a konvex halmazok definíciója is lehetne. Ezt mondja a következő:

4.3.2 Állítás. *Az euklideszi tér egy K részhalmaza pontosan akkor konvex, ha bármely véges sok pontjának konvex lineáris kombinációit is tartalmazza.*



4.2. ábra: Konvex és nem konvex halmazok

Bizonyítás. Az elegendőség nyilvánvaló.
A szükségességet a tekintett pontok száma szerinti teljes indukcióval igazojuk. Két pontjának konvex lineáris kombinációit konvex halmaz tartalmazza, hiszen ezzel értelmeztük egy halmaz konvexitását. Tegyük fel, hogy a K konvex halmaz legfeljebb $n-1 \geq 2$ pontjának minden konvex lineáris kombinációját tartalmazza, és legyenek $v_1, \ldots, v_{n-1}, v_n$ K-beli pontok, továbbá $\lambda_1, \ldots, \lambda_{n-1}, \lambda_n$ olyan nemnegatív skalárok, hogy $\sum_{i=1}^n \lambda_i = 1$. Feltehetjük, hogy ezek a skalárok mindegyike pozitív, mert különben a $\sum_i^n \lambda_i v_i$ ténylegesen n-nél kevesebb pont konvex lineáris kombinációja lenne, és az indukciós feltevés alapján azonnal adódna, hogy eleme K-nak. Igy a

$$\frac{\lambda_i}{1-\lambda_n}\,, \quad (i = 1, \ldots, n-1)$$

skalárok olyan nemnegatív számok, amelyeknek az összege 1-gyel egyenlő, következésképpen a

$$\sum_{i=1}^{n-1} \frac{\lambda_i}{1-\lambda_n} v_i \in K\,.$$

Akkor viszont a

$$\lambda_n v_n + (1-\lambda_n)\left(\sum_{i=1}^{n-1} \frac{\lambda_i}{1-\lambda_n} v_i\right) = \sum_{i=1}^{n} \lambda_i v_i$$

vektor is K-ban van, és ezt kellett igazolnunk. □

A következő állítás segítségünkre lesz tetszőleges részhalmaz konvex lezártjának értelmezésénél.

4.3.3 Állítás. *Konvex halmazok metszete is konvex.*

Bizonyítás. Legyenek K_γ , $(\gamma \in \Gamma)$ konvex halmazok, és legyen $v, w \in \bigcap_{\gamma \in \Gamma} K_\gamma$. (Feltehetjük, hogy a metszetnek van legalább két különböző eleme, mert különben a metszet biztos konvex.) Akkor $v, w \in K_\gamma$ minden $\gamma \in \Gamma$-ra, és a K_γ halamazok konvexitása miatt v és w minden

$$x = \lambda v + (1-\lambda) w \quad 0 \leq \lambda \leq 1$$

konvex lineáris kombinációja is minden K_γ halmazban benne van. Akkor viszont $x \in \cap_{\gamma\in\Gamma} K_\gamma$ is teljesül. Ezzel a bizonyítást befejeztük. □

4.3.4 Definíció. *Legyen H az E euklideszi tér tetszőleges részhalmaza. Azt a legszűkebb konvex halmazt, amely H-t tartalmazza, a H konvex burkának nevezzük és* $\mathrm{co}\,(H)$*-val jelöljük.*

Nyilván $\mathrm{co}\,(H)$ az összes H-t tartalmazó konvex halmaz közös része. Mivel az egész euklideszi tér konvex, ez mindig létezik. Reméljük, hogy most az olvasó a vektorterek részhalmazai lineáris burkára asszociál, már csak azért is, mert itt is érvényes az analóg állítás, hogy

4.3.5 Állítás. *Egy H halmaz* $\mathrm{co}\,(H)$ *konvex burka megegyezik a H halmaz elemei összes konvex lineáris kombinációinak halmazával.*

Bizonyítás. Jelöljük a H halmaz elemei összes konvex lineáris kombinációinak halmazát H'-vel. Akkor H' konvex, mert ha

$$\sum_{i=1}^{p} \alpha_i v_i \quad \text{és} \quad \sum_{j=1}^{q} \beta_j w_j$$

a H halmaz elemeinek konvex lineáris kombinációja, akkor minden $\lambda\,,\ \ (0 \leq \lambda \leq 1)$ valós számra a

$$\lambda \left(\sum_{i=1}^{p} \alpha_i v_i \right) + (1-\lambda) \left(\sum_{j=1}^{q} \beta_j w_j \right)$$

is konvex lineáris kombinációja H elemeinek. Mivel $H \subseteq H'$ kapjuk, hogy $H' \subseteq \mathrm{co}\,(H)$.
Az is igaz viszont, hogy $H \subseteq \mathrm{co}\,(H)$ és konvex halmaz zárt az elemeinek konvex lineáris kombináció képzésére (lásd az (4.3.2) állítást), ezért $\mathrm{co}\,(H) \subseteq H'$ is fennáll. Akkor $\mathrm{co}\,(H) = H'$ és ezt kellett igazolnunk. □

Véve két pontot az euklideszi térben és képezve ennek a kételemű részhalmaznak a konvex burkát, a már megismert szakaszt kapjuk. Ezért a továbbiakban a v és w pontokat összekötő szakasz jelölésére a kifejezőbb $\mathrm{co}\,(v, w)$-t használjuk. Ha három olyan pontot veszünk, melyek közül egyik sem konvex lineáris kombinációja a másik kettőnek, akkor konvex burkuk háromszöget ad. Fel kell hívjuk az olvasó figyelmét, hogy háromszöget kaphatunk lineárisan összefüggő rendszert alkotó három pont konvex burkaként is, csupán azt kellett megkövetelnünk, hogy konvex lineáris kombinációval ne lehessen előállítani egyik pontot sem a másik kettőből, ami azt jelenti, hogy egyik pont sem legyen eleme a másik két pont által meghatározott szakasznak. Hasonló konstrukcióval kaphatjuk a sík összes konvex sokszögét és a térbeli konvex testeket is. Persze, már a kétdimenziós euklideszi terekben is vannak olyan konvex részhalmazok, amelyek nem kaphatók meg véges részhalmaz konvex burkaként, csak gondoljunk egy körre, amely nyilván ilyen tulajdonságú. Az euklideszi tér véges részhalmazának konvex burkát *poliédernek* nevezzük.

Ha K konvex halmaz és $p \in K$ olyan pontja, amely nem belső pontja egyetlen K-beli szakasznak sem, akkor p-t a K halmaz *extremális pontjának*, vagy más szóval *csúcspontjának* hívjuk.

4.3.6 Tétel. *Legyen $H = \{p_1, \ldots, p_k\}$ az euklideszi tér egy véges részhalmaza. Akkor* $\mathrm{co}\,(H)$ *minden extremális pontja H-beli, és $p_\ell \in H$ pontosan akkor extremális pontja* $\mathrm{co}\,(H)$*-nak, ha nem állítható elő a $H \setminus \{p_\ell\}$ halmaz pontjainak konvex lineáris kombinációjaként.*

Bizonyítás. Legyen $c \in \mathrm{co}\,(H)$ a H halmazbeli elemek

$$c = \sum_{i=1}^{k} \alpha_i p_i\,, \quad (\alpha_i \geq 0\,,\ \sum_{i=1}^{k} \alpha_i = 1)$$

konvex lineáris kombinációja. Ha $c \notin H$, akkor biztosan létezik olyan j, hogy $0 < \alpha_j < 1$. Akkor a

$$p = \sum_{\substack{i=1 \\ i \neq j}}^{k} \frac{\alpha_i}{1 - \alpha_j} p_i$$

pont eleme $\mathrm{co}\,(H)$-nak, hiszen H-beli pontok konvex lineáris kombinációja, és

$$c = \alpha_j p_j + (1 - \alpha_j) p$$

belső pontja a $\overline{p_j, p}$ szakasznak, igazolva ezzel, hogy $\mathrm{co}\,(H)$ extremális pontjai H-ban kell legyenek.
A második állítás igazolása végett lássuk be azt, hogy ha $p_\ell \in H$ előállítható $H \setminus \{p_\ell\}$-beli pontok konvex lineáris kombinációjaként, akkor $\mathrm{co}\,(H) = \mathrm{co}\,(H \setminus \{p_\ell\}$. Valóban, ha

$$p_\ell = \sum_{\substack{i=1 \\ i \neq \ell}}^{k} \beta_i p_i\,, \quad \beta_i \geq 0, \quad \sum_{\substack{i=1 \\ i \neq \ell}}^{k} \beta_i = 1\,,$$

akkor $\mathrm{co}\,(H)$ bármely $\sum_{i=1}^{k} \gamma_i p_i$ eleme előállítható

$$\sum_{\substack{i=1 \\ i \neq \ell}}^{k} \gamma_i p_i + \gamma_\ell \sum_{\substack{i=1 \\ i \neq \ell}}^{k} \beta_i p_i = \sum_{\substack{i=1 \\ i \neq \ell}}^{k} (\gamma_i + \gamma_\ell \beta_i) p_i$$

alakban, ahol nyilván $\gamma_i + \gamma_\ell \beta_i \geq 0$ minden $i \neq \ell (= 1, \ldots, k)$-ra és

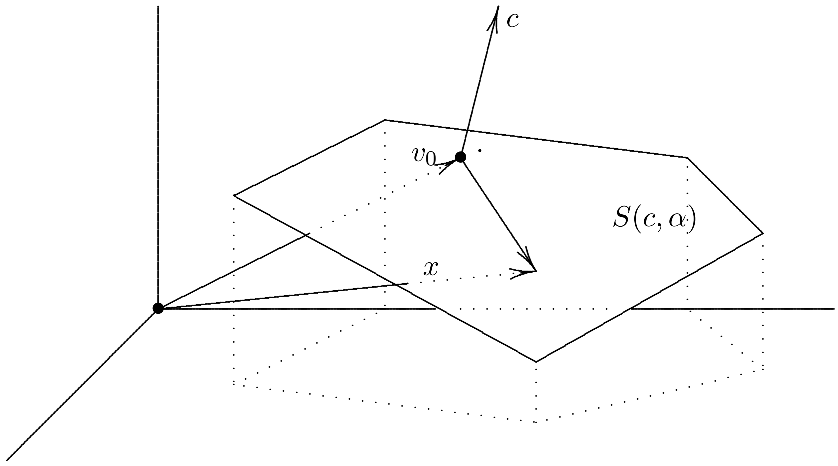
$$\sum_{\substack{i=1 \\ i \neq \ell}}^{k} (\gamma_i + \gamma_\ell \beta_i) = 1\,.$$

Ezzel igazoltuk, hogy $\mathrm{co}\,(H) \subseteq \mathrm{co}\,(H \setminus \{p_\ell\}$, és mivel a fordított irányú tartalmazás nyilvánvalóan fennáll, azt is, hogy $\mathrm{co}\,(H) = \mathrm{co}\,(H \setminus \{p_\ell\}$. Az első, már igazolt állítás felhasználásával, mostmár kapjuk, hogy $\mathrm{co}\,(H)$ extremális pontjai a H halmaz olyan minimális $\bar{H}$ részhalmazának a pontjai, amelyre $\mathrm{co}\,(H) = \mathrm{co}\,(\bar{H})$ teljesül. □

A fenti tétel azt jelenti, hogy az euklideszi terek poliéderei esetében a csúcspontok ugyanolyan szerepet játszanak a poliéder "generálásában", mint a bázisvektorok a tér generálásában.

A 3-dimenziós euklideszi tér pontjainak és egyeneseinek megfelelőit tetszőleges euklideszi terekben már értelmeztük, de nem definiáltuk még a síkok megfelelőit.

A síkok egyik megfogható tulajdonsága, hogy ha egy pontjukba a síkra merőleges vektort állítunk, az a sikban fekvő minden vektorra merőleges lesz. Az 4.3 ábra azt igyekszik illusztrálni, hogy ha ezt a c-vel jelzett vektort a sík v_0 pontjába toljuk, akkor a sík bármely más x pontja (pontjába mutató vektor) azzal jellemezhető, hogy a síkban fekvő $x - v_0$ vektor merőleges c-re.



4.3. ábra: Sík a 3-dimenziós térben

A merőlegesség, ortogonalitás kifejezhető minden euklideszi térben a skaláris szorzat segítségével, nevezetesen az egymásra merőleges vektorok skaláris szorzata nulla. Éppen ez szolgált a 3-dimenziós térben a síkok

$$\langle c, x - v_0 \rangle = 0$$

egyenletének megadására. Ezt az egyenletet kicsit más alakban szokás megadni, bevezetve a $\langle c, v_0 \rangle$ skalárra az α jelölést, azt mondhatjuk, hogy pontosan azok az x pontok alkotják a szóbanforgó síkot, amelyek kielégítik az

$$\langle c, x \rangle = \alpha$$

egyenletet.

Ez az értelmezés tetszőleges euklideszi térben lehetséges. Azt mondjuk, hogy az E euklideszi tér *hipersíkja* a tér

$$S(c, \alpha) = \{x \in E \mid \langle c, x \rangle = \alpha\}$$

részhalmaza, ahol $c \in E$ és $\alpha \in \mathbb{R}$ egy adott vektor, illetve skalár. A c vektort a hipersík normálvektorának is szokás nevezni, annak kifejezésére, hogy c a síkban fekvő minden vektorra ortogonális. Valóban, ha $x_1,\ x_2 \in S(c, \alpha)$, akkor

$$\langle c, x_1 - x_2 \rangle = \langle c, x_1 \rangle - \langle c, x_2 \rangle = \alpha - \alpha = 0\,.$$

Megjegyzés: Emlékezve arra a megjegyzésre, hogy amennyiben a skaláris szorzat valamelyik változóját rögzítjük, akkor az a másik változónak lineáris függvénye, a hipersíkok interpretálhatók úgy is, mint nemnulla lineáris függvények nívóhalmaza.

Valóban a $\langle c, \cdot \rangle : E \to \mathbb{R}$ lineáris funkcionál $\alpha \in \mathbb{R}$ skalárhoz tartozó nívóhalmaza éppen $S(c, \alpha)$. Megfordítva, ha $f \in E^*$ egy tetszőleges nemnulla lineáris funkcionál és $\alpha \in \mathbb{R}$ tetszőleges skalár, akkor a $\{x \mid f(x) = \alpha\}$ ponthalmaz egy hipersík. Véges dimenziós esetben a kétféle megadás ekvivalenciája könnyen igazolható, míg végtelen dimenziós esetben a hipersík nemnulla lineáris funkcionál nivóhalmazaként való értelmezése általánosabb.

4.3.7 Állítás. *A hipersíkok, egyenesek, szakaszok a térnek konvex részhalmazai, ha a hipersík, vagy egyenes az origót (a nullvektort) is tartalmazza, akkor az euklideszi térnek alterei is.*

Bizonyítás. Legyen az $S(c, \alpha)$ hipersíknak x_1 és x_2 két tetszőleges pontja és $\lambda, (0 \le \lambda \le 1)$ valamely skalár. Akkor

$$\langle c, \lambda x_1 + (1-\lambda)x_2 \rangle = \lambda \langle c, x_1 \rangle + (1-\lambda) \langle c, x_2 \rangle = \lambda\alpha + (1-\lambda)\alpha = \alpha\,,$$

igazolva, hogy a hipersík zárt a konvex lineáris kombináció képzésére nézve, azaz konvex. Az origó pontosan akkor eleme a hipersíknak, ha az $\alpha = 0$. Akkor viszont bármely x_1, $x_2 \in S(c, 0)$ pontjára ortogonális a c normálvektor, így azok tetszőleges $\beta_1 x_1 + \beta_2 x_2$ lineáris kombinációjára is, igazolva, hogy az origón átmenő hipersíkok alterek. Az egyenesek és szakaszok konvexitásának, illetve az origón átmenő egyenesek altér voltának igazolását az olvasóra bizzuk. □

Az n-dimenziós euklideszi terek egyenesei egydimenziósak, abban az értelemben, hogy azok vagy egydimenziós alterek, vagy egydimenziós alterek eltolásával kaphatók. Az origón átmenő hipersíkok $n-1$-dimenziós alterek. (Ez következik az (4.1.16) tételből, és abból, hogy egy $S(c, \alpha)$ hipersík az 1-dimenziós $\mathrm{lin}\,(c)$ ortogonális komplementere.) Általában pedig a hipersíkok $n-1$-dimenziós alterek eltolásával kaphatók. (Eltoláson azt értjük, hogy az altér minden pontjához ugyanazt a vektort hozzáadva kapjuk az egyenes, illetve hipersík pontjait.) A hipersík normálvektora persze nem egyértelmű, egy c normálvektor tetszőleges nemzéró skalárszorosai is normálvektorok, és persze a normálvektortól függ az α skalár. A 3-dimenziós térben három nem egy egyenesre eső pont egyértelműen meghatározta a rájuk illeszkedő síkot. Ehhez az n-dimenziós euklideszi térben n pontra van szükség, és ezek közül legalább $n-1$ lineárisan független kell legyen. Az alábbi példában bemutatjuk, hogy 4 pont megadásával hogyan határozható meg a 4-dimenziós euklideszi tér egy hipersíkja.

5 Példa. *Legyen az E_4 euklideszi tér egy ortonormált bázisa az $\mathcal{X} = \{v_1, v_2, v_3, v_4\}$ és*

$$\begin{array}{rcrrrr} a_1 &=& v_1 & -v_2 & +v_3 & -v_4\,, \\ a_2 &=& 2v_1 & & +v_3 & -2v_4\,, \\ a_3 &=& & 2v_2 & -v_3\,, & \\ a_4 &=& -v_1 & +v_2 & +2v_3 & +v_4\,. \end{array}$$

Határozzuk meg az a_1, a_2, a_3, a_4 pontokat tartalmazó hipersíkot.

Jelöljük a keresett hipersík egy normálvektorát c-vel, és legyen ennek a koordináta vektora az $\mathcal{X}$ ortonormált bázisban $\mathbf{c} = [\gamma_1, \gamma_2, \gamma_3, \gamma_4]$. Mivel a_1, a_2, a_3, a_4 a hipersík pontjai, az $a_1 - a_2$, $a_1 - a_3$, $a_1 - a_4$ vektorok a hipersíkra illeszkednek, ennélfogva ortogonálisak a c vektorra, azzal való skaláris szorzatuk zéró. A skaláris szorzatukat most egyszerűen az $\mathcal{X}$ ortonormált bázisra vonatkozó belső szorzatként kapjuk az (4.1.20) tétel szerint. Mivel

$$\mathbf{a}_1 - \mathbf{a}_2 = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{a}_1 - \mathbf{a}_3 = \begin{bmatrix} 1 \\ -3 \\ 2 \\ -1 \end{bmatrix} \quad \mathbf{a}_1 - \mathbf{a}_4 = \begin{bmatrix} 2 \\ -2 \\ -1 \\ -2 \end{bmatrix},$$

teljesülni kell az alábbi egyenleteknek:

$$\begin{array}{rrrrcl} -\gamma_1 & -\gamma_2 & & +\gamma_4 & = & 0 \\ \gamma_1 & -3\gamma_2 & +2\gamma_3 & -\gamma_4 & = & 0 \\ 2\gamma_1 & -2\gamma_2 & -\gamma_3 & -2\gamma_4 & = & 0 \end{array}$$

Ennek a homogén lineáris egyenletrendszernek a nemnulla megoldásai a lehetséges normálvektorok. Az elemi bázistranszformációs módszerrel a megoldás:

$$\begin{array}{c|rrrr} & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ \hline & \boxed{-1} & -1 & 0 & 1 \\ & 1 & -3 & 2 & -1 \\ & 2 & -2 & -1 & -2 \\ \hline \end{array} \rightarrow \begin{array}{c|rrr} & \gamma_2 & \gamma_3 & \gamma_4 \\ \hline \gamma_1 & 1 & 0 & -1 \\ & -4 & 2 & 0 \\ & -4 & \boxed{-1} & 0 \\ \hline \end{array} \rightarrow$$

$$\rightarrow \begin{array}{c|rr} & \gamma_2 & \gamma_4 \\ \hline \gamma_1 & 1 & -1 \\ & \boxed{-12} & 0 \\ \gamma_3 & 4 & 0 \\ \hline \end{array} \rightarrow \begin{array}{c|r} & \gamma_4 \\ \hline \gamma_1 & -1 \\ \gamma_2 & 0 \\ \gamma_3 & 0 \\ \hline \end{array}$$

Az utolsó táblázat alapján az általános megoldás:

$$\mathbf{c} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \cdot \tau\,, \quad \text{ahol } \tau \text{ nemnulla valós.}$$

A $\tau = 1$ választással kapott $\mathbf{c} = [1,\ 0,\ 0,\ 1]$ koordináta vektorú $c = v_1 + v_4$ vektor egy lehetséges normálvektor. Ezzel a normálvektorral a hipersíkhoz tartozó skalár a $\langle c, a_1 \rangle$ skaláris szorzat, ami esetünkben $[c, a_1]_{\mathcal{X}} = 0$. Igy a keresett hipersík:

$$S(c, 0) = \{x \in E_4 \mid \langle c, x \rangle \, (= [c, x]_{\mathcal{X}}) = 0\}\,.$$

Tehát a $\mathbf{x} = [\xi_1, \xi_2, \xi_3, -\xi_1]$ koordináta vektorú $x = \xi_1 v_1 + \xi_2 v_2 + \xi_3 v_3 - \xi_1 v_4$ pontok alkotják a hipersíkot, ahol ξ_i, $(i = 1, 2, 3)$ tetszőleges valós számok. Amint az könnyen látható a kapott $S(c, 0)$ hipersík most 3-dimenziós altér. □

Az euklideszi terek további nevezetes konvex részhalmazai a félterek: Legyen $S(c, \alpha)$ az E euklideszi tér valamely hipersíkja. A

$$H_1(c, \alpha) = \{x \in E \mid \langle c, x \rangle < \alpha\}\,,$$

$$H_2(c,\alpha) = \{x \in E \mid \langle c, x\rangle > \alpha\}\,,$$
$$H_3(c,\alpha) = \{x \in E \mid \langle c, x\rangle \le \alpha\}\,,$$
$$H_4(c,\alpha) = \{x \in E \mid \langle c, x\rangle \ge \alpha\}$$

részhalmazokat *féltereknek* nevezzük. Ezen belül, mivel a $H_1(c,\alpha)$ és $H_2(c,\alpha)$ halmazok nyíltak, azokat *nyílt féltereknek* mondjuk, és minthogy a $H_3(c,\alpha)$ és $H_4(c,\alpha)$ halmazok zártak, ezeket *zárt féltereknek* hívjuk.

Könnyen belátható, hogy az egész E euklideszi tér felbontható az $S(c,\alpha)$, a $H_1(c,\alpha)$ és $H_2(c,\alpha)$ részhalmazainak diszjunkt (közös pont nélküli) egyesítésére.

Az előzőekből kiderült, hogy a geometriai fogalmak általánosítása szempontjából nagyon hasznosnak bizonyúlt a konvex lineáris kombináció fogalma. Az optimumszámításban azonban szükség van további geometriai fogalmakra is, amelyek értelmezéséhez a lineáris kombináció egy másik speciális esetének megadása szükséges.

4.3.8 Definíció. *Az E euklideszi tér $v_1, v_2, \ldots, v_n$ pontjainak $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ lineáris kombinációját kúpkombinációnak nevezzük, ha minden $i(= 1, 2, \ldots, n)$-re $\alpha_i \ge 0$.*

Amint láttuk az euklideszi tér konvex részhalmazai azok, amelyek zártak a konvex lineáris kombináció képzésre nézve. Ennek megfelelően azt mondjuk, hogy

4.3.9 Definíció. *Az E euklideszi tér egy nemüres C részhalmazát kúpnak nevezzük, ha zárt elemeinek kúpkombinációjára nézve.*

Tekintettel arra, hogy kúpok közös része is kúp, lehetőség van az E tetszőleges H részhalmaza által generált kúp definiálására; nevezetesen ezen az összes H-t tartalmazó kúp közös részét kell értenünk. Könnyen bizonyítható, hogy ez a kúp megegyezik a H halmaz elemeinek összes kúpkombinációját tartalmazó $\mathrm{con}\,(H)$-val jelölt kúppal.

Egy C kúpot végesen generált kúpnak, vagy röviden *véges kúpnak* nevezünk, ha van olyan véges elemszámú H részhalmaza, hogy $C = \mathrm{con}\,(H)$.

A lineáris programozási problémák megoldásánál betöltött szerepük miatt meg kell még említenünk a poliedrikus halmaz fogalmát. Az euklideszi tér egy részhalmazát akkor nevezzük *poliedrikus halmaznak,* ha az véges sok zárt féltér közös része.

Meg kell említenünk, hogy a poliedrikus halmaz fogalma másképpen is definiálható, ehhez azonban előbb meg kell mondjuk, hogy hogyan kell képezni az euklideszi tér részhalmazainak algebrai összegét. Ha L_1 és L_2 az E euklideszi tér két részhalmaza, akkor algebrai összegük az

$$L_1 + L_2 = \{\ell_1 + \ell_2 \mid \ell_1 \in L_1, \ell_2 \in L_2\}$$

halmaz.

A részhalmazok algebrai összegének értelmezése után a poliedrikus halmazok fogalma így is definiálható:

4.3.10 Definíció. *Az E euklideszi tér egy P részhalmaza poliedrikus halmaz, ha van olyan K poliéder és C véges kúp, hogy $P = K + C$.*

A poliedrikus halmaz két definíciójának ekvivalenciája nem nyilvánvaló, bizonyítása olyan tételekre támaszkodik, amelyeket itt nem közölhetünk a terjedelmi korlátok miatt.

4.3.1 Térelemek távolsága*

A projekciós tételre támaszkodva megmutattuk, hogy hogyan számítható ki az euklideszi tér valamely pontjának egy altértől való távolsága. Azóta megismertük néhány geometriai fogalom euklideszi terekre való általánosítását, értelmeztük az egyenes, illetve hipersík fogalmát. Az alábbiakban megmutatjuk, hogyan számítható ki pontnak egyenestől, hipersíktól, illetve hipersík hipersíktól való távolsága.

Pont és egyenes távolsága

Határozzuk meg az E euklideszi tér v pontjának az $\overleftrightarrow{x_1x_2}$ egyenesétől való távolságát. A Pythagorasz tételből következik, hogy az egyenes olyan z pontja lesz v-hez legközelebb, amelyre a $v-z$ vektor ortogonális az egyenesre. Ezért

$$d(v,\overleftrightarrow{x_1x_2}) = d(v,z) = \|v-z\|\,,$$

ahol $z \in \overleftrightarrow{x_1x_2}$ és $v-z \perp \overleftrightarrow{x_1x_2}$. Első lépésben tehát a z pont megkeresése a feladat.

$$z = \lambda x_1 + (1-\lambda)x_2 = x_2 + \lambda(x_1 - x_2)\,,$$

ahol λ a

$$\langle v - z, x_1 - x_2\rangle = 0$$

ortogonalitást kifejező egyenletből kapható:

$$\langle v - x_2 - \lambda(x_1 - x_2), x_1 - x_2\rangle = \langle v - x_2, x_1 - x_2\rangle - \lambda\|x_1 - x_2\|^2 = 0\,,$$

azaz,

$$\lambda = \frac{\langle v - x_2, x_1 - x_2\rangle}{\|x_1 - x_2\|^2}.$$

A λ ismeretében az egyenes keresett z pontja:

$$z = x_2 + \frac{\langle v - x_2, x_1 - x_2\rangle}{\|x_1 - x_2\|^2}(x_1 - x_2)\,.$$

Végül pedig a v és z pontok távolsága:

$$\|v-z\| = [\langle v - x_2 - \lambda(x_1 - x_2), v - x_2 - \lambda(x_1 - x_2)\rangle]^{\frac{1}{2}} =$$

$$\left[\|v - x_2\|^2 - 2\lambda\langle v - x_2, x_1 - x_2\rangle + \lambda^2\|x_1 - x_2\|^2\right]^{\frac{1}{2}} =$$

$$\sqrt{\|v - x_2\|^2 - \frac{\langle v - x_2, x_1 - x_2\rangle^2}{\|x_1 - x_2\|^2}}\,.$$

Pont és hipersík távolsága

Számítsuk ki az E euklideszi tér v pontjának az $S(c,\alpha)$ hipersíktól való távolságát. Hívatkozva a Pythagoras-tételre, adódik, hogy

$$d(v, S(c,\alpha)) = d(v,z) = \|v-z\|\,,$$

ahol $z \in S(c,\alpha)$ és $v - z \perp S(c,\alpha)$, vagy ami ezzel ekvivalens, $v - z = \lambda c$, hiszen c a hipersík normálvektora, ortogonális a hipersíkra. Az utóbbi egyenlőségből kapjuk, hogy $z = v - \lambda c$, és mert z a hipersík pontja

$$\langle c, z\rangle = \langle c, v - \lambda c\rangle = \langle c, v\rangle - \lambda \|c\|^2 = \alpha\,.$$

Ebből λ kifejezhető:

$$\lambda = \frac{\langle c, v\rangle - \alpha}{\|c\|^2}\,,$$

és íggy a keresett távolság:

$$\|v - z\| = \|\lambda c\| = \frac{|\,\langle c, v\rangle - \alpha|}{\|c\|}\,.$$

Párhuzamos hipersíkok távolsága

Két hipersík pontosan akkor párhuzamos egymással, ha normálvektoraik egymásnak nemnulla skalárszorosai. Mivel azonban egy hipersík normálvektora csak skalárszorzótól eltekintve egyértelműen meghatározott, a párhuzamos hipersíkoknak választható egyenlő normálvektor. Legyen tehát $S(c,\alpha)$ és $S(c,\beta)$ a két egymással párhuzamos hipersík, amelyek távolságát akarjuk kiszámítani. Nyilvánvalóan

$$d(S(c,\alpha), S(c,\beta)) = d(v,w) = \|v - w\|\,,$$

ahol $v \in S(c,\alpha)\,, w \in S(c,\beta$ és $v - w$ ortogonális a hipersíkokra, azaz

$$v - w = \lambda c\,.$$

Ez utóbbi egyenlőségből kifejezve v-t és kihasználva, hogy $v \in S(c,\alpha)$ és $w \in S(c,\beta$, kapjuk, hogy

$$\alpha = \langle v, c\rangle = \langle w + \lambda c, c\rangle = \langle w, c\rangle + \lambda\|c\|^2 = \beta + \lambda\|c\|^2\,.$$

Az egyenlőségsorozat bal- és jobboldalát összehasonlítva kifejezhető λ, azaz

$$\lambda = \frac{\alpha - \beta}{\|c\|^2}\,,$$

és így a keresett távolság:

$$\|v - w\| = \|\lambda c\| = \frac{|\alpha - \beta|}{\|c\|}\,.$$

2. Gyakorlatok, feladatok

1. Legyen az E_n euklideszi tér egy véges nemüres részhalmaza M. Mutassuk meg, hogy $\mathrm{con}\,(M)$ korlátos halmaz.

2. Igazoljuk, hogy a valós számtest feletti lineáris egyenletrendszer megoldáshalmaza konvex, és ha van legalább két eleme, akkor nem korlátos.

3. Bizonyítsuk be, hogy egy poliedrikus halmaznak véges sok extremális pontja van.

4. Legyen az E_4 euklideszi tér x_1, x_2, x_3 pontjának koordináta vektora egy $\mathcal{X}$ ortonormált bázisban

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ -2 \\ 1 \\ -1 \end{bmatrix} \quad \text{és} \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}$$

a. Határozzuk meg az x_1 pontnak az $\overleftrightarrow{x_2x_3}$ egyenestől való távolságát.
b. Számítsuk ki az x_1, x_2, x_3 pontok által meghatározott háromszög területét.

5. Mutassuk meg, hogy az E_n euklideszi tér háromszögeinek súlypontja is a csúcspontok számtani közepe.

6. Határozzuk meg az E_4 euklideszi tér v pontjának az $S(c, 0)$ hipersíktól való távolságát, ha egy ortonormált bázisban $\mathbf{v} = [1, 0, -1, 2]$ és $\mathbf{c} = [2, -1, 0, 3]$.

7. Számítsuk ki az $S(c, 0)$ és az $S(c, 3)$ hipersíkok távolságát, ha $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$ ortonormált bázis és $c = x_1 + 2x_2 - x_3 + x_4$.

4.4 Unitér terek*

Az unitér terek a komplex vektorterekből ugyanúgy kaphatók, mint ahogy a valós vektorterekből az euklideszi tereket kaptuk. Ezért ebben a pontban csupán a komplex terekben értelmezett skaláris szorzat megadására szorítkozunk, és azt mutatjuk meg, hogy minden véges dimenziós komplex vektortérben értelmezhető skaláris szorzat, tehát unitér térré tehető.

4.4.1 Definíció. *Legyen V komplex vektortér és*

$$\langle , \rangle : V \oplus V \to \mathbb{C}$$

olyan függvény, amely minden $v, w \in V$ vektorpárhoz egy $\langle v, w\rangle$-vel jelölt komplex számot rendel. A $\langle , \rangle$ függvényt komplex skaláris szorzatnak nevezzük, ha eleget tesz az alábbi feltételeknek: bármely $v, w, z \in V$ vektorokra és tetszőleges λ komplex számra

(1) $\langle v, w\rangle = \overline{\langle w, v\rangle}$,

(2) $\langle \lambda v, w\rangle = \bar{\lambda}\,\langle v, w\rangle$,

(3) $\langle v + w, z\rangle = \langle v, z\rangle + \langle w, z\rangle$,

(4) $\langle v, v\rangle \geq 0 \quad \text{és} \quad \langle v, v\rangle = 0 \Longleftrightarrow v = 0.$

4.4.2 Definíció. *Egy véges dimenziós komplex vektorteret unitér térnek nevezünk, ha komplex skaláris szorzat van benne értelmezve.*

A véges dimenziós valós vektorterekben valamely bázisra vonatkozó, úgynevezett belső szorzat egyszersmind skaláris szorzat, sőt azt is láttuk, hogy minden skaláris szorzat a tér ortonormált bázisára vonatkozó belső szorzat. Hasonló igaz véges dimenziós komplex vektorterekben is, amennyiben a komplex belső szorzatot az alábbi módon definiáljuk.

4.4.3 Definíció. *Legyen $\mathcal{X} = \{x_1, \ldots, x_n\}$ a V komplex vektortér egy bázisa és legyenek*

$$v = \alpha_1 x_1 + \cdots + \alpha_n x_n \quad \textit{és} \quad w = \beta_1 x_1 + \cdots + \beta_n x_n$$

tetszőleges V-beli vektorok. A

$$[v, w]_{\mathcal{X}} = \sum_{i=1}^{n} \bar{\alpha_i} \beta_i$$

komplex számot a v és w vektorok $\mathcal{X}$ bázisra vonatkozó komplex belső szorzatának nevezzük.

Az olvasó könnyen ellenőrizheti, hogy a komplex belső szorzat rendelkezik a komplex skaláris szorzattól megkövetelt négy tulajdonsággal, tehát minden véges dimenziós komplex vektortér unitér térré tehető.

A komplex skaláris szorzatos terek is normált terek a

$$\|z\| \overset{\text{def}}{=} \sqrt{\langle z, z\rangle}$$

norma-függvénnyel, de a Cauchy-Schwarz-egyenlőtlenség bizonyítása kicsit bonyolultabb a valós esetben látottól, ezért azt részletezzük. A komplex skaláris szorzatos V tér bármely két s, t vektorára és bármely λ komplex számra fennáll az

$$\langle s + \lambda t, s + \lambda t\rangle = \langle s, s\rangle + \lambda \langle s, t\rangle + \bar{\lambda} \langle t, s\rangle + \bar{\lambda}\lambda \langle t, t\rangle \geq 0$$

egyenlőtlenség, ami kicsit tömörebben az

$$\|s\|^2 + \lambda \langle s, t\rangle + \bar{\lambda} \langle t, s\rangle + |\lambda|^2 \|t\|^2 \geq 0 \tag{1}$$

alakban is írható. Az (1) egyenlőtlenséget szorozva a pozitív $\|s\|^2$-tel, kapjuk, hogy

$$\|s\|^4 + \lambda \|s\|^2 \langle s, t\rangle + \bar{\lambda} \|s\|^2 \langle t, s\rangle + |\lambda|^2 \|s\|^2 \|t\|^2 =$$

$$= \left(\|s\|^2 + \lambda \langle s, t\rangle\right) \left(\|s\|^2 + \bar{\lambda} \langle t, s\rangle\right) - |\lambda|^2 \langle s, t\rangle \langle t, s\rangle + |\lambda|^2 \|s\|^2 \|t\|^2 \geq 0\,. \tag{2}$$

Mivel a (2) egyenlőtlenség minden komplex λ számra teljesül, fennáll a

$$\lambda = -\frac{\langle s, t\rangle}{\|s\|^2}$$

számra is, amikor is (2) a

$$-|\lambda|^2 \langle s, t\rangle \langle t, s\rangle + |\lambda|^2 \|s\|^2 \|t\|^2 \geq 0 \tag{3}$$

egyenlőtlenségre redukálódik. Végül (3)-at átrendezve, a pozitív $|\lambda|^2$-tel elosztva és figyelembe véve, hogy

$$\langle s,t\rangle \langle t,s\rangle = \langle s,t\rangle \overline{\langle s,t\rangle} = |\langle s,t\rangle|^2 \ ,$$

kapjuk, hogy

$$|\langle s,t\rangle|^2 \le \|s\|^2\|t\|^2 \, ,$$

amiből négyzetgyököt vonva adódik a kívánt

$$|\langle s,t\rangle| \le \|s\| \cdot \|t\|$$

Cauchy-Schwarz-egyenlőtlenség.

Mivel a komplex vektorterek is normált terek, azok is metrikus térré tehetők a

$$d(s,t) \stackrel{\text{def}}{=} \|s-t\|$$

metrikával, és a Cauchy-Schwarz-egyenlőtlenség birtokában a komplex vektorok ortogonalitása is ugyanúgy értelmezhető, mint a valós skaláris szorzatos terekben.

Igy unitér terekben is vannak ortonormált bázisok, és a komplex skaláris szorzat is a vektorok ortonormált bázisra vonatkozó komplex belső szorzatára redukálódik.

Az euklideszi geometria fogalmai kiterjeszthetők az unitér terekre is, ezekkel azonban itt nem foglalkozunk.

5. Fejezet

Invariáns alterek

Ennek a fejezetnek az a célja, hogy a véges dimenziós vektor terek lineáris transzformációit jellemezzük. Milyen jellemzésre gondolunk? A 2-dimenziós valós térbeli lineáris transzformációk, mint a nyújtás, tükrözés, elforgatás, vetítés és párhuzamos affinitás, hatásáról jól kialakult elképzelésünk van. Nagyobb dimenziójú terekben persze nem várhatjuk, hogy egy lineáris transzformáció hatását ugyanúgy "lássuk" mint a síkon, de arra törekedhetünk, hogy meg tudjuk mondani azok hatását a tér szemléltethető alterein. Tehát a tér altereit használjuk a vizsgált lineáris transzformációk jellemzésére. Persze ehhez csak olyan alterek használhatók, amelyeknek vektorai a transzformáció során az altérben maradnak. Ezek az úgynevezett invariáns alterek. Ha az egész vektorteret fel tudjuk bontani olyan invariáns alterek direktöszegére, amelyeken már tudjuk a lineáris transzformáció hatását, akkor a transzformációt ismertnek mondhatjuk. Kiderül, hogy szoros kapcsolat van a lineáris transzformációk polinomjainak faktorizációi és a tér invariáns altereinek direktösszegére való felbontása között. Ez szükségessé teszi, hogy ismertessünk néhány a polinomok faktorizációjával kapcsolatos nélkülözhetetlen eredményt. Ezeket külön alpontban gyűjtöttük össze, amit a területtel ismerős olvasó nyugodtan kihagyhat.

5.1 Invariáns alterek, transzformációk polinomjai

Mint azt a fejezet bevezetőjében említettük a lineáris transzformációkat úgy kívánjuk jellemezni, hogy felbontjuk a teret olyan alterek direkt összegére, amelyen a lineáris transzformáció hatásáról már jó elképzelésünk van. Az ilyen altereknek persze a lineáris transzformációra nézve "zártnak" kell lennie. A lineáris transzformációra vonatkozó zártság fogalmát az alábbi definícióban rögzítettük.

5.1.1 Definíció. *Legyen V egy $\mathbb{F}$ test feletti vektortér és A lineáris transzformációja V-nek. A V vektortér egy M alterét A-ra nézve invariánsnak mondjuk, ha minden $v \in M$-re az $A(v)$ képvektor is M-ben van.*

Hangsúlyoznunk kell, hogy azt nem követeltük meg a definícióban, hogy M minden eleme előálljon valamely V-beli vektor képeként, és azt sem, hogy ha egy w vektor $A(w)$ képe M-ben van, akkor w-nek is M-ben kell lennie, csupán azt, hogy az M-beli vektorok képe M-ben kell maradjon. A rövidség kedvéért sokszor az A-ra nézve invariáns alterekre az *A-invariáns* jelzővel hívatkozunk. Időnként nem igazán logikusan ugyan, de röviden a vektortér A-invariáns altere kifejezés helyett azt mondjuk, hogy az A *invariáns altere.*

Tetszőleges A transzformációnak nyilván invariáns altere a zérusaltér, $\ker(A)$, $\mathrm{im}\,(A)$ és maga a V vektortér.

Kevésbé triviális A-invariáns altereket kaphatunk a következő konstrukcióval: kiválasztunk egy tetszőleges $v \in V$ vektort, majd képezzük a $\{v, A(v), A^2(v), \ldots\}$ vektorok által generált alteret, amelyet $\mathrm{lin}\,(v, A)$-val fogunk jelölni. Ha V n-dimenziós, akkor a $\{v, A(v), A^2(v), \ldots A^n(v)\}$ vektorrendszer biztosan lineárisan összefüggő, hiszen $n+1$ vektort tartalmaz, és még az is igaz, hogy ha $A^k(v)$ $(1 \leq k \leq n)$ kifejezhető a $v, A(v), \ldots, A^{k-1}(v)$ vektorok lineáris kombinációjaként, akkor minden pozitív egész m-re az $A^{k+m}(v)$ is előállítható, mint a $v, A(v), \ldots, A^{k-1}(v)$ vektorok lineáris kombinációja. Az állítás m szerinti teljes indukcióval kőnnyen igazolható. Valóban, ha

$$A^k(v) = \sum_{i=0}^{k-1} \alpha_i A^i(v)\,,$$

akkor

$$A^{k+1}(v) = \sum_{i=0}^{k-1} \alpha_i A^{i+1}(v) = \sum_{i=0}^{k-2} \alpha_i A^{i+1}(v) + \alpha_{k-1}\left(\sum_{i=0}^{k-1} \alpha_i A^i(v)\right) =$$

$$= \alpha_{k-1}\alpha_0 v + \sum_{i=1}^{k-1} (\alpha_{i-1} + \alpha_{k-1}\alpha_i) A^i(v)\,,$$

tehát $m = 1$-re igaz az állítás. Ha $m - 1 \geq 1$-re

$$A^{k+m-1}(v) = \sum_{i=0}^{k-1} \beta_i A^i(v)\,,$$

akkor a

$$A^{k+m}(v) = \sum_{i=0}^{k-1} \beta_i A^{i+1}(v) = \sum_{i=0}^{k-2} \beta_i A^{i+1}(v) + \beta_{k-1}\left(\sum_{i=0}^{k-1} \beta_i A^i(v)\right) =$$

$$= \beta_{k-1}\beta_0 v + \sum_{i=1}^{k-1} (\beta_{i-1} + \beta_{k-1}\beta_i) A^i(v)\,,$$

számolással kapjuk, hogy $A^{k+m}(v)$ is kifejezhető a $v, A(v), \ldots, A^{k-1}(v)$ vektorok lineáris kombinációjaként. Így $\mathrm{lin}\,(v, A) = lin\,(v, A(v), \ldots, A^{n-1}(v))$. Tulajdonképpen ezért használjuk a $\mathrm{lin}\,(v, A)$ jelölést. Nem nehéz belátnunk, hogy $\mathrm{lin}\,(v, A)$ az a legszűkebb A-invariáns altér, amely a v vektort tartalmazza.

Tekintve, hogy egy vektortér lineáris transzformációi nemcsak vektorteret alkotnak, de lineáris transzformációk szorzata is lineáris transzformáció, és így egy lineáris transzformáció nemnegatív egész kitevős hatványainak lineáris kombinációja is a tér lineáris transzformációja, tetszőleges $A \in L(V)$ esetén a

$$p(A) = \alpha_0 A^0 + \alpha_1 A + \cdots + \alpha_k A^k$$

polinom is lineáris transzformációja V-nek. Könnyen igazolható, hogy az A és $p(A)$ transzformációk szorzata független a tényezők sorrendjétől. Ezt ismerve, meg tudjuk mutatni, hogy a $p(A)$ transzformáció magtere és képtere is A-invariáns altér. Valóban, ha $v \in \ker(p(A))$, azaz $p(A)(v) = 0$, akkor $p(A)(A(v)) = A(p(A)(v)) = A(0) = 0$, tehát $A(v) \in \ker(p(A))$. Hasonlóan, ha $v \in \operatorname{im}(p(A))$, azaz létezik olyan $w \in V$, amelyre $p(A)(w) = v$, akkor $p(A)(A(w)) = A(p(A)(w)) = A(v)$, igazolva, hogy $A(v) \in \operatorname{im}(p(A))$ is teljesül.

Utóbbi példánk sugallja, hogy szükségünk lesz a lineáris transzformációk polinomjaira, azok invariáns altereinek konstruálásakor. Ezért most egy kis kitérőt teszünk, és összefoglaljuk a polinomokkal kapcsolatos azon fogalmakat és állításokat, amelyekre az alábbiakban szükségünk lesz.

5.1.1 Polinomok

Az előző fejezetekben, már szerepeltek a valós együtthatós polinomok, így kicsit meglepő lehet, hogy az azokra vonatkozó fogalmak és eredmények összefoglalására csak most kerül sor. Ennek az a magyarázata, hogy eddig a polinomokat mint egy vektortér elemeit vizsgáltuk, és ezért a polinomok összeadásával és skalárral való szorzásával kapcsolatos tulajdonságaik játszottak elsődleges szerepet. Az alábbiakban a polinomok szorzásával kapcsolatos fogalmakat és eredményeket gyűjtöttük össze. Mindenekelőtt fogalmazzuk meg újra, hogy mit kell értenünk egy tetszőleges $\mathbb{F}$ test feletti polinomon, majd meg kell ismerkednünk a polinomok maradékos osztásával.

Rögzítünk egy változónak nevezett szimbólumot, legyen ez mondjuk t, és az $\mathbb{F}$ test feletti t változós polinomok $\mathbb{F}[t]$ halmazát alkossák a

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

alakú formális kifejezések, ahol az $\alpha_0, \alpha_1, \ldots, \alpha_n$ az $\mathbb{F}$ test elemei, n pedig nemnegatív egész szám. A $p(t)$ polinomot n-edfokúnak mondjuk, ha a legnagyobb kitevőjű nem nulla együtthatójú t-hatvány a t^n. Jelekkel: $\deg p(t) = n$. Az $\alpha_n \neq 0$ skalárt a polinom *főegyütthatójának* nevezzük és $p(t)$-t *normáltnak* hívjuk, ha főegyütthatója 1. Ugyanúgy, mint azt a valós együtthatós polinomok esetében láttuk, tetszőleges $\mathbb{F}$ test feletti polinomok halmazán is értelmezhető összeadás és az $\mathbb{F}$ test elemeivel, mint skalárokkal való szorzás. ($\mathbb{F}$ operátortartománya $\mathbb{F}[t]$-nek.) Könnyen ellenőrizhető, hogy $\mathbb{F}[t]$ ezekkel a műveletekkel az $\mathbb{F}$ test feletti vektortér. De az összeadáson és a skalárokkal való szorzáson kívül értelmezhető a polinomok szorzása is. Egy

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

és egy

$$q(t) = \beta_0 + \beta_1 t + \cdots + \beta_m t^m$$

polinom szorzatán értve a

$$p(t) \cdot q(t) = \alpha_0\beta_0 + (\alpha_0\beta_1 + \alpha_1\beta_0)\, t + \cdots + \alpha_n\beta_m t^{n+m}$$

polinomot. A szorzatban a t^k-t tartalmazó tag együtthatója a

$$\sum_{i+j=k} \alpha_i\beta_j \ \ (0 \le i \le n; 0 \le j \le m)\,.$$

Nem nehéz ellenőrizni, hogy a polinomok szorzása kommutatív, asszociatív, az összeadásra nézve disztributív művelet. A szorzás definíciójából következik, hogy $\deg p(t)q(t) = \deg p(t) + \deg q(t)$. Azt mondjuk, hogy a $q(t) \in \mathbb{F}[t]$ polinom osztója a $p(t) \in \mathbb{F}[t]$ polinomnak, jelekkel $q(t) \mid p(t)$, ha létezik olyan $s(t) \in \mathbb{F}[t]$, amelyre $p(t) = s(t) \cdot q(t)$ teljesül. Egy $p(t) \in \mathbb{F}[t]$ polinomot *irreducibilisnek* nevezünk, ha nincs olyan $q(t) \in \mathbb{F}[t]$ osztója, amelyre $0 < \deg q(t) < \deg p(t)$ teljesül. Bár a polinomok szorzásra vonatkozó inverze általában nem létezik, de végezhető, úgynevezett maradékos osztás. Konkretizálva, igaz a következő állítás.

5.1.2 Állítás. *Ha $p(t), q(t)(\neq 0) \in \mathbb{F}[t]$ tetszőleges polinomok, akkor léteznek olyan egyértelműen meghatározott $h(t), r(t) \in \mathbb{F}[t]$ polinomok, hogy*

$$p(t) = h(t) \cdot q(t) + r(t)\,,$$

és ha $\deg q(t) \neq 0$, akkor $0 \leq \deg r(t) < \deg q(t)$.

Bizonyítás. Legyenek

$$p(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_n t^n \quad \alpha_n \neq 0$$

és

$$q(t) = \beta_0 + \beta_1 t + \cdots + \beta_m t^m \quad \beta_m \neq 0\,.$$

Ha $n < m$, akkor a $h(t) = 0$ és $r(t) = p(t)$ polinomok eleget tesznek a kívánt feltételnek. Ha $n \geq m$, akkor képezzük a

$$p_1(t) = p(t) - \frac{\alpha_n}{\beta_m} t^{n-m} q(t) =$$

$$\alpha_{10} + \alpha_{11} t + \cdots + \alpha_{n_1} t^{n_1} \quad \alpha_{n_1} \neq 0; \; (0 \leq n_1 \leq n-1)$$

n_1-edfokú polinomot. Ha $n_1 < m$, akkor legyen

$$h(t) = \frac{\alpha_n}{\beta_m} t^{n-m} \quad \text{és} \quad r(t) = p_1(t).$$

Ellenkező esetben képezzük a

$$p_2(t) = p_1(t) - \frac{\alpha_{n_1}}{\beta_m} t^{n_1-m} q(t) =$$

$$= \alpha_{20} + \alpha_{21} t + \cdots + \alpha_{n_2} t^{n_2} \quad \alpha_{n_2} \neq 0; \; (0 \leq n_2 \leq n_1)$$

n_2-edfokú polinomot. Ha $n_2 < m$, akkor legyen

$$h(t) = \frac{\alpha_n}{\beta_m} t^{n-m} + \frac{\alpha_{n_1}}{\beta_m} t^{n_1-m} \quad \text{és} \quad r(t) = p_2(t)\,.$$

Ha $p_2(t)$ foka még nem kisebb $q(t)$ fokánál, akkor hasonlóan $p_1(t)$ és $p_2(t)$ kiszámításához, képezzük a

$$p_3(t) = p_2(t) - \frac{\alpha_{n_2}}{\beta_m} t^{n_2-m} q(t)$$

n_3-adfokú polinomot, és így tovább. Mivel a $p(t), p_1(t), p_2(t), \ldots$ polinomok fokszámai monoton csökkenő $n > n_1 > n_2 > \ldots$ sorozatot alkotnak, véges számú lépés után elérünk egy

$$p_k(t) = p_{k-1}(t) - \frac{\alpha_{n_{k-1}}}{\beta_m} t^{n_{k-1}-m} q(t)$$

polinomhoz, amelynek n_k foka már kisebb m-nél, és

$$p_k(t) = p(t) - \left(\frac{\alpha_n}{\beta_m} t^{n-m} + \frac{\alpha_{n_1}}{\beta_m} t^{n_1-m} + \cdots + \frac{\alpha_{n_{k-1}}}{\beta_m} t^{n_k-m} \right) q(t).$$

Ebből azonnal adódik, hogy a

$$h(t) = \frac{\alpha_n}{\beta_m} t^{n-m} + \frac{\alpha_{n_1}}{\beta_m} t^{n_1-m} + \cdots + \frac{\alpha_{n_{k-1}}}{\beta_m} t^{n_k-m} \quad \text{és} \quad r(t) = p_k(t)$$

polinomok felhasználásával $p(t)$ a kívánt alakban áll elő.
Az egyértelműség igazolása végett tegyük fel, hogy a $h'(t)$ és $r'(t)$ polinomokra is teljesül, hogy

$$p(t) = h'(t)q(t) + r'(t) \qquad \text{és} \qquad 0 \leq \deg r'(t) < \deg q(t).$$

Akkor

$$\left(h(t) - h'(t)\right) q(t) = r'(t) - r(t)\,.$$

Az egyenlőség jobboldalán lévő polinom fokszáma kisebb mint $q(t)$ foka, míg a baloldalán lévő polinom foka akkor és csak akkor kisebb $q(t)$ fokánál, ha $h(t) - h'(t) = 0$, azaz $h(t) = h'(t)$. Akkor viszont $r(t) = r'(t)$ is teljesül.
Ezzel bizonyításunk teljes. □

Két polinom $p(t), q(t) \in \mathbb{F}[t]$ legnagyobb közös osztóján olyan $d(t) \in \mathbb{F}[t]$ normált polinomot értünk, amely mind $p(t)$-nek, mind $q(t)$-nek osztója és ha $c(t)$ is osztója $p(t)$-nek is és $q(t)$-nek is, akkor $c(t) \mid d(t)$. Azt mondjuk, hogy a $p(t)$ és $q(t)$ polinomok *relatív prímek*, ha legnagyobb közös osztójuk a konstans 1 polinom.

Két polinom $p(t), q(t) \in \mathbb{F}[t]$ legkisebb közös többszöröse pedig olyan $k(t) \in \mathbb{F}[t]$ normált polinom, amelynek mind $p(t)$, mind $q(t)$ osztója, és ugyanakkor osztója $p(t)$ és $q(t)$ minden közös többszörösének. Megmutatható, hogy két polinom legkisebb közös többszöröse egyenlő a polinomok szorzatának és legnagyobb közös osztójuk hányadosával.

A maradékos osztás birtokában, adhatunk olyan algoritmust, amellyel tetszőleges két $p(t), q(t) \in \mathbb{F}[t]$ polinom legnagyobb közös osztója meghatározható. Az eljárás a következő:
Elvégezzük az alábbi maradékos osztások sorozatát, mindaddig, amíg a maradék nulla nem lesz. Mivel a maradék polinomok fokszámai szigorúan csökkenő sorozatot alkotnak véges számú lépés után a maradék zéró lesz.

$$\begin{aligned}
p(t) &= h_1(t)q(t) + r_1(t) \;\text{ ahol }\; 0 \leq \deg r_1(t) < \deg q(t) \qquad (5.1)\\
q(t) &= h_2(t)r_1(t) + r_2(t) \;\text{ ahol }\; 0 \leq \deg r_2(t) < \deg r_1(t)\\
r_1(t) &= h_3(t)r_2(t) + r_3(t) \;\text{ ahol }\; 0 \leq \deg r_3(t) < \deg r_2(t)\\
&\;\;\vdots\\
r_{n-2}(t) &= h_n(t)r_{n-1}(t) + r_n(t) \;\text{ ahol }\; 0 \leq \deg r_n(t) < \deg r_{n-2}(t)\\
r_{n-1}(t) &= h_{n+1}(t)r_n(t)
\end{aligned}$$

Az utolsó nemzéró $r_n(t)$ polinomot szükség esetén normáljuk, megszorozva főegyütthatója reciprokával. Jelölje az így kapott polinomot $d(t)$. Azt állítjuk, hogy $d(t)$ a $p(t)$ és $q(t)$ polinomok legnagyobb közös osztója. $d(t) \mid r_n(t)$, és az utolsó egyenlőség miatt $r_n(t) \mid r_{n-1}(t)$ akkor viszont az utolsó előtti egyenlőség miatt $r_n(t) \mid r_{n-2}(t)$ és így tovább haladva kapjuk, hogy $r_n(t)$ osztója $q(t)$ és $p(t)$-nek is. Másrészt ha $c(t)$ osztója mind $p(t)$-nek mind $q(t)$-nek, akkor az első egyenlet alapján $c(t) \mid r_1(t)$, amiből a második egyenlőség alapján $c(t) \mid r_2(t)$ és így tovább egyenlőségről egyenlőségre haladva végül kapjuk, hogy $c(t) \mid r_n(t)$. Minthogy $d(t)$ és $r_n(t)$ csak konstans szorzóban különböznek $r_n(t) \mid d(t)$ is telejesül, ahonnan már a $c(t) \mid d(t)$ is következik.

A polinomok legnagyobb közös osztójának meghatározását szolgáló algoritmus alkalmas arra, hogy megmutassuk, hogy amennyiben a $p(t)$ és $q(t)$ polinomok legnagyobb közös osztója $d(t)$, akkor találhatók olyan $f(t)$ és $g(t)$ polinomok, hogy

$$d(t) = f(t)p(t) + g(t)q(t)\,. \tag{5.2}$$

Ezt könnyen beláthatjuk, ha figyelembe vesszük a (5.1) algoritmus egyenlőségeinek átrendezésével kapott

$$\begin{aligned}
r_1(t) &= p(t) - h_1(t)q(t)\\
r_2(t) &= q(t) - h_2(t)r_1(t)\\
r_3(t) &= r_1(t) - h_3(t)r_2(t)\\
&\;\vdots\\
r_{n-1}(t) &= r_{n-3}(t) - h_{n-1}(t)r_{n-2}(t)\\
r_n(t) &= r_{n-2}(t) - h_n(t)r_{n-1}(t)
\end{aligned}$$

egyenleteket. Ebből láthatjuk, hogy $r_1(t)$ a $p(t)$ és $q(t)$ polinomszorosainak összege. Ezt helyettesítve a második egyenlőségbe

$$r_2(t) = q(t) - h_2(t)(p(t) - h_1(t)q(t)) = h_2(t)p(t) + (1 + h_1(t)h_2(t))q(t)\,,$$

tehát $r_2(t)$ is a $p(t)$ és $q(t)$ polinomszorosainak összege. Tegyük fel, hogy már igazoltuk az $r_1(t), \ldots, r_{n-2}(t), r_{n-1}(t)$ polinomok mindegyikéről, hogy kifejezhetők a $p(t)$ és $q(t)$ polinomszorosainak összegeként. Akkor az utolsó egyenlet mutatja, hogy ez lehetséges $r_n(t)$-re is. Minthogy $d(t)$ az $r_n(t)$ polinom skalárszorosa, így valóban léteznek olyan $f(t)$ és $g(t)$ polinomok, hogy $d(t) = f(t)p(t) + g(t)q(t)$.

1 Példa. *Határozzuk meg a $p(t) = t^4 - 5t^2 + 4$ és a $q(t) = t^3 + 3t^2 - t - 3$ polinomok legnagyobb közös osztóját és keressük meg azokat az $f(t)$ és $g(t)$ polinomokat, amelyekkel a legnagyobb közös osztó felírható $f(t)p(t) + g(t)q(t)$ alakban!*

$$\begin{array}{rrrrrrrrrl}
t^4 & & & - & 5t^2 & & & + & 4 & \div\; t^3+3t^2-t-3 \;=\; t-3\,,\\
t^4 & + & 3t^3 & - & t^2 & - & 3t & & & \\
\hline
 & - & 3t^3 & - & 4t^2 & + & 3t & + & 4 & \\
 & - & 3t^3 & - & 9t^2 & + & 3t & + & 9 & \\
\hline
 & & & & 5t^2 & & & - & 5 &
\end{array}$$

amiből

$$t^4 - 5t^2 + 4 = (t-3)(t^3 + 3t^2 - t - 3) + 5t^2 - 5\,.$$

$$\begin{array}{rrrrrrrrrrr}
t^3 & + & 3t^2 & - & t & - & 3 & \div & 5t^2 - 5 & = & \frac{1}{5}t + \frac{3}{5}\,, \\
t^3 & & & - & t & & & & & & \\
\hline
 & & 3t^2 & & & - & 3 & & & & \\
 & & 3t^2 & & & - & 3 & & & & \\
\hline
 & & & & & & 0 & & & &
\end{array}$$

következésképpen

$$t^3 + 3t^2 - t - 3 = (\frac{1}{5}t + \frac{3}{5})(5t^2 - 5)\,.$$

Mivel az $5t^2 - 5$ polinom az utolsó nemzéró maradék, a $t^2 - 1$ normált polinom a legnagyobb közös osztó.

$$t^2 - 1 = \frac{1}{5}(t^4 - 5t^2 + 4) + (-\frac{1}{5}t + \frac{3}{5})(t^3 + 3t^2 - t - 3)\,,$$

tehát $f(t) = \frac{1}{5}$ és $g(t) = -\frac{1}{5}t + \frac{3}{5}$. □

Amint azt analízis tanulmányaink során láttuk, egy $p(t) \in \mathbb{F}[t]$ polinom segítségével értelmezhetünk $\mathbb{F} \longrightarrow \mathbb{F}$ alakú függvényt, az $\alpha \longrightarrow p(\alpha)$ hozzárendeléssel. Az $\alpha \in \mathbb{F}$ elemet a $p(t)$ polinom *gyökének* nevezzük, ha $p(\alpha) = 0$. Persze nem biztos, hogy egy $\mathbb{F}$ feletti polinomnak van gyöke $\mathbb{F}$-ben, de ha α gyöke $p(t)$-nek, akkor az elsőfokú $t - \alpha$ polinom osztója $p(t)$-nek. A $p(t)$-nek $t - \alpha$-val való maradékos osztása

$$p(t) = h(t)(t - \alpha) + r(t) \quad \text{és} \quad 0 \le \deg r(t) < 1\,,$$

alakú, tehát $r(t) = \beta$ konstans polinom, α-t helyettesítve kapjuk, hogy

$$p(\alpha) = h(\alpha)(\alpha - \alpha) + \beta = \beta\,.$$

Ebből következik, hogy $t - \alpha$ pontosan akkor osztója $p(t)$-nek, ha $p(\alpha) = 0$. Ha $\mathbb{F}[t]$ minden polinomjának van gyöke $\mathbb{F}$-ben, akkor az $\mathbb{F}$ testet *algebrailag zártnak* nevezzük. A valós számok teste, mint tudjuk algebrailag nem zárt, de a komplex számok teste már igen. A komplex számhalmaz azonban tartalmazza a valós számokat is, azok a komplex számtest úgynevezett résztestét alkotják. Bizonyított az az — algebra alaptételével lényegében ekvivalens — állítás, hogy minden $\mathbb{F}$ test részteste valamely algebrailag zárt testnek. Egy algebrailag zárt testben tehát, minden irreducibilis polinom elsőfokú. A polinomok szorzási szabálya alapján nyilvánvaló, hogy akkor egy n-edfokú polinom n darab elsőfokú irreducibilis polinom szorzatára bontható, megengedve, hogy ugyanaz az irreducibilis faktor többször is szerepeljen a szorzatban. Ha $(t - \alpha)$ m-szer szerepel egy $p(t)$ polinom ilyen faktorizációjában, akkor azt mondjuk, hogy az α m *multiplicitású* gyöke $p(t)$-nek. Egy algebrailag zárt test feletti n-edfokú polinomnak tehát n gyöke van, ha mindegyiket annyiszor vesszük figyelembe, mint amennyi a multiplicitása. Ha $\mathbb{F}$ tetszőleges test, akkor egy $p(t) \in \mathbb{F}[t]$ polinom faktorizálható, nem feltétlenül elsőfokú, de irreducibilis polinomok szorzatára. Persze ebben az esetben is lehet ugyanaz az irreducibilis polinom többszörös multiplicitású tényezője $p(t)$-nek. Így az általános esetben $p(t)$ irreducibilis polinomok szorzatára való felbontása

$$p(t) = (p_1(t))^{m_1} \cdot \cdots \cdot (p_r(t))^{m_r}$$

alakú, ahol $p_i(t) \in \mathbb{F}[t] \quad (i = 1, \ldots, r)$ irreducibilis polinomok. Felhívjuk az olvasó figyelmét a polinomok irreducibilis polinomok hatványainak szorzatára való felbontása és az egész számok prímhatványok szorzataként való előállítása közötti analógiára.

5.1.2 Lineáris transzformációk és mátrixaik polinomjai

Egy $p(t) \in \mathbb{F}[t]$ polinom nemcsak $\mathbb{F} \to \mathbb{F}$ alakú függvények értelmezéséhez használható, de minden olyan algebrai struktúra önmagába való leképezéseinek definiálásához is, amelyben szorzás van értelmezve és amelynek az $\mathbb{F}$ test operátortartománya. Ilyen egy $\mathbb{F}$ feletti V vektortér lineáris transzformációinak $L(V)$ tere is és persze az $\mathbb{F}$ elemeiből képzett $\dim V \times \dim V$ tipusú mátrixok $\mathbb{F}_{\dim V \times \dim V}$ mátrixok tere is, amely mint jól tudjuk izomorf $L(V)$-vel, és a mátrixok szorzását úgy értelmeztük, hogy minden

$$\Phi : L(V) \to \mathbb{F}_{\dim V \times \dim V}$$

izomorf leképezés felcserélhető a szorzásművelettel is. Minden $A \in L(V)$ lineáris transzformációhoz a

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

polinom segítségével hozzárendelhetjük a V vektortér

$$p(A) = \alpha_0 A^0 + \alpha_1 A + \cdots + \alpha_n A^n$$

lineáris transzformációját és hasonlóan minden $\mathbf{A} \in \mathbb{F}_{\dim V \times \dim V}$ mátrixhoz egy

$$p(\mathbf{A}) = \alpha_0 \mathbf{A}^0 + \alpha_1 \mathbf{A} + \cdots + \alpha_n \mathbf{A}^n \in \mathbb{F}_{\dim V \times \dim V}$$

mátrixot. Látható, hogy a $p(A)$ lineáris transzformáció az A lineáris transzformáció hatványainak lineáris kombinációja, és hasonlóan a $p(\mathbf{A})$ mátrix a $\mathbf{A}$ mátrix hatványainak lineáris kombinációja. A lineáris transzformációk és mátrixaik vektorterének izomorfiája biztosítja, hogy egy A lineáris transzformáció bármely $p(t)$ polinomjának mátrixa valamely rögzített bázisban megegyezik az A transzformáció $\mathbf{A}$ mátrixának $p(\mathbf{A})$ polinomjával.
Nem nehéz belátni, hogy egy lineáris transzformáció (mátrix) polinomjaira igazak a következő számolási szabályok: ha

$$p(t) + q(t) = r(t) \qquad \text{és} \qquad p(t) \cdot q(t) = s(t)\,,$$

akkor

$$p(A) + q(A) = r(A) \qquad \text{és} \qquad p(A) \cdot q(A) = s(A)\,,$$

illetve

$$p(\mathbf{A}) + q(\mathbf{A}) = r(\mathbf{A}) \qquad \text{és} \qquad p(\mathbf{A}) \cdot q(\mathbf{A}) = s(\mathbf{A}).$$

Mivel a polinomok szorzása kommutatív, egy lineáris transzformáció (egy mátrix) két polinomjának szorzata is független a tényezők sorrendjétől. Azt mondjuk, hogy az A lineáris transzformáció (egy $\mathbf{A}$ mátrix) *gyöke* a $p(t)$ polinomnak, ha $p(A) = 0$ ($p(\mathbf{A}) = \mathbf{0}$). Egy lineáris transzformáció pontosan akkor gyöke egy $p(t)$ polinomnak, ha mátrixa gyöke annak. Minthogy egy lineáris transzformációnak különböző bázisokban különböző, egymáshoz hasonló mátrixai vannak azonnal kapjuk, hogy hasonló mátrixok ugyanazoknak a polinomoknak a gyökei. Ezt az eredményt persze közvetlenül is megkaphatjuk, figyelembe véve, hogy ha $\mathbf{B}$ invertálható mátrix, akkor minden $p(t)$ polinomra

$$p(\mathbf{B}^{-1}\mathbf{A}\mathbf{B}) = \mathbf{B}^{-1}p(\mathbf{A})\mathbf{B}\,.$$

Ezek az észrevételek lehetővé teszik, hogy feladatok numerikus megoldásakor lineáris transzformációk polinomjai helyett a transzformációk valamely bázisra vonatkozó mátrixainak polinomjaival számolhassunk.

A következőkben állításainkat csak a lineáris transzformációk polinomjaira mondjuk ki, azzal az előrebocsájtott megjegyzéssel, hogy azok átfogalmazhatók a mátrixok polinomjaira is.

Ha V n-dimenziós F feletti vektortér, akkor minden $A \in L(V)$ gyöke valamely $\mathbb{F}[t]$-beli nemzéró polinomnak, hiszen $\dim L(V) = n^2$ lévén az

$$A^0 = I, A, \ldots, A^{n^2}$$

$n^2 + 1$ elemű lineáris transzformáció rendszer lineárisan összefüggő, így van olyan nemtriviális lineáris kombinációjuk, hogy

$$\alpha_0 A^0 + \alpha_1 A + \cdots + \alpha_{n^2} A^{n^2} = 0\,.$$

Akkor a

$$p(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_{n^2} t^{n^2}$$

nemzéró polinom olyan, amelynek A gyöke. Nyilvánvaló, hogy ha A gyöke egy $p(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_m t^m$ m-edfokú polinomnak, akkor gyöke az $1/\alpha_m \cdot p(t)$ ugyancsak m-edfokú normált polinomnak is.

5.1.3 Definíció. *Azt a legkisebb fokszámú normált polinomot, amelynek A gyöke, az A transzformáció minimálpolinomjának nevezzük.*

Minden lineáris transzformáció minimálpolinomja egyértelműen meghatározott, hiszen ha $m(t)$ és $m'(t)$ is minimálpolinomja egy A transzformációnak, akkor A gyöke az $m(t) - m'(t)$ polinomnak is, holott $m(t) - m'(t)$ foka kisebb mint $m(t)$ fokszáma, ellentmondva a $m(t)$ értelemezésének.

5.1.4 Állítás. *Az A lineáris transzformáció minimálpolinomja osztója minden olyan polinomnak, amelynek A gyöke.*

Bizonyítás. Legyen $f(t)$ tetszőleges olyan polinom, amelynek az A transzformáció gyöke, azaz $f(A) = 0$. Végezzünk maradékos osztást,

$$f(t) = h(t)m(t) + r(t) \;\text{ ahol }\; 0 \le \deg r(t) < \deg m(t).$$

Helyettesítve az A transzformációt, kapjuk, hogy

$$0 = f(A) = h(A)m(A) + r(A) = r(A)\,,$$

ami csak akkor lehet igaz, ha $r(t) \equiv 0$, mert A nem lehet gyöke egyetlen $m(t)$ fokszámánál alacsonyabb fokú nemzéró polinomnak sem. Ezzel állításunkat igazoltuk. $\square$

Az alábbi tétel egy lineáris transzformáció invariáns altereinek dimenziójáról nyújt felvilágosítást.

5.1.5 Tétel.

(1) *Ha az $A \in L(V)$ lineáris transzformáció minimálpolinomjának van k-adfokú irreducibilis faktora, akkor van a V vektortérben k-dimenziós A-invariáns altér.*

(2) *Ha $p(t)$ egy k-adfokú irreducibilis faktora az A minimálpolinomjának, akkor* $\ker(p(A))$ *k-dimenziós A-invariáns altér, vagy k-dimenziós A-invariáns alterek direktösszege.*

Bizonyítás. (1) Jelölje $m(t)$ az A minimálpolinomját és legyen az $m(t) = p(t)q(t)$ szorzatként való előállításban $p(t)$ k-adfokú irreducibilis tényező. A $\ker(p(A))$ altér invariáns A-ra nézve. Legyen $v \neq 0$ egy tetszőleges vektora $\ker(p(A))$-nak (ilyen van, mert különben már a $q(t)$ polinomnak gyöke lenne A). Meg fogjuk mutatni, hogy $\mathrm{lin}\,(v, A)$ k-dimenziós A-invariáns altere $\ker(p(A))$-nak, következésképpen V-nek is. Tekintsük a $\{v, A(v), \dots A^n(v)\}$ vektorrendszert, ahol $n = \dim V$. Ez nyilvánvalóan lineárisan összefüggő rendszer. Legyen $m(\leq n)$ az a legnagyobb pozitív egész, amelyre a $\{v, A(v), \dots, A^{m-1}(v)\}$ vektorrendszer még lineárisan független, azaz $A^m(v)$ már kifejezhető a $v, A(v), \dots, A^{m-1}(v)$ vektorok

$$A^m(v) = \varepsilon_0 v + \varepsilon_1 A(v) + \cdots + \varepsilon_{m-1} A^{m-1}(v)$$

lineáris kombinációjaként. Az

$$s(t) = -\varepsilon_0 - \varepsilon_1 t - \cdots - \varepsilon_{m-1} t^{m-1} + t^m$$

polinomra tehát teljesül, hogy $s(A)(v) = 0$, de az m számra tett kikötésünk alapján állíthatjuk, hogy m-nél alacsonyabb fokú nemzéró polinomja A-nak a v vektort nem transzformálja a nullvektorba. Mivel v eleme volt $\ker(p(A))$-nak, tehát $p(A)(v) = 0$, ezért

$$m = \deg s(t) \leq \deg p(t) = k\,.$$

Maradékos osztást végezve a $p(t)$ és $s(t)$ polinomokkal, kapjuk, hogy

$$p(t) = h(t)s(t) + r(t) \qquad \text{és} \qquad 0 \leq \deg r(t) < \deg s(t) = m\,,$$

és így

$$p(A) = h(A)s(A) + r(A)\,.$$

Akkor

$$0 = p(A)(v) = h(A)s(A)(v) + r(A)(v) = r(A)(v)\,,$$

és ebből következik, hogy $r(t) \equiv 0$ és $p(t) = h(t)s(t)$. Mivel feltevésünk szerint $p(t)$ irreducibilis, a $h(t)$ polinom csak konstans polinom lehet, ezért $s(t)$ és $p(t)$ fokszáma egyenlő, tehát $k = m$. A $\mathrm{lin}\,\left(v, A(v), \dots, A^{m-1}(v)\right) = \mathrm{lin}\,(v, A)$ altér tehát k-dimenziós és A-invariáns.

(2) A második állítás igazolása: $\ker(p(A))$ két különböző nemzéró v és w vektorára $\mathrm{lin}\,(v, A)$ és $\mathrm{lin}\,(w, A)$ vagy diszjunktak (csak a nullvektor a közös elemük), vagy egyenlők, mert ha $x \in \mathrm{lin}\,(v, A) \bigcap \mathrm{lin}\,(w, A)$ nemzéró vektor, akkor az $\{x, A(x), \dots, A^{k-1}(x)\}$ vektorrendszer lineárisan független, és mind $\mathrm{lin}\,(v, A)$-nak, mind $\mathrm{lin}\,(w, A)$-nak részrendszere azok A-invarianciája miatt. Mivel

$$\dim \mathrm{lin}\,(v, A) = \dim \mathrm{lin}\,(w, A) = k$$

azt is kapjuk, hogy az $\{x, A(x), \ldots, A^{k-1}(x)\}$ vektorrendszer mind $\mathrm{lin}\,(v, A)$-nak, mind $\mathrm{lin}\,(w, A)$-nak bázisa, és így

$$\mathrm{lin}\,(v, A) = \mathrm{lin}\,(w, A)\,.$$

Már most $\ker\,(p(A))$ direktösszegre való felbontása a következőképpen történhet: választunk egy nemzéró $v_1 \in \ker\,(p(A))$ vektort és képezzük a $\mathrm{lin}\,(v_1, A)$ alterét. Ha $\ker\,(p(A)) \setminus \mathrm{lin}\,(v_1, A)$ nem üres, akkor választunk belőle egy v_2 vektort és képezzük a $\mathrm{lin}\,(v_1, A) \oplus \mathrm{lin}\,(v_2, A)$ $2 \cdot k$-dimenziós alterét $\ker\,(p(A))$-nak. Ha ez még valódi altér, akkor választhatunk

$$v_3 \in \ker\,(p(A)) \setminus \mathrm{lin}\,(v_1, A) \oplus \mathrm{lin}\,(v_2, A)$$

vektort és képezhetjük a

$$\mathrm{lin}\,(v_1, A) \oplus \mathrm{lin}\,(v_2, A) \oplus \mathrm{lin}\,(v_3, A)$$

alteret, és így tovább. $\ker\,(p(A))$ véges dimenziós volta biztosítja, hogy létezik olyan $r \geq 1$ egész, hogy

$$\ker\,(p(A)) \setminus \mathrm{lin}\,(v_1, A) \oplus \cdots \oplus \mathrm{lin}\,(v_r, A) = \emptyset\,,$$

és akkor az eljárás befejeződik. Hangsúlyoznunk kell, hogy erősen kihasználtuk, hogy amennyiben egy $x \neq 0$ vektor eleme valamely $\mathrm{lin}\,(y, A)$ altérnek, akkor $\mathrm{lin}\,(x, A) = \mathrm{lin}\,(y, A)$, amiből már következik, hogy ha

$$v_i \notin \mathrm{lin}\,(v_1, A) \oplus \cdots \oplus \mathrm{lin}\,(v_{i-1}, A)\,,$$

akkor

$$\mathrm{lin}\,(v_i, A) \bigcap \mathrm{lin}\,(v_1, A) \oplus \cdots \oplus \mathrm{lin}\,(v_{i-1}, A) = \{0\}\,.$$

Ezzel a bizonyítást befejeztük. □

Ismert, hogy minden valós együtthatós irreducibilis polinom legfeljebb másodfokú, míg minden komplex együtthatós irreducibilis polinom elsőfokú. Így az előző tétel értelmében a véges dimenziós valós vektorterek minden lineáris transzformációjának van legfeljebb 2-dimenziós, a komplex vektorterek lineáris transzformációinak pedig 1-dimenziós invariáns altere.

5.2 Sajátvektorok és sajátértékek

Ebben a pontban a vektorterek olyan lineáris transzformációit tanulmányozzuk, amelyekre nézve létezik a térnek 1-dimenziós invariáns altere.

Legyen az $\mathbb{F}$ test feletti V vektortérnek $A \in L(V)$ lineáris transzformációja. Ha az A $m(t)$ minimálpolinomjának $\lambda \in \mathbb{F}$ gyöke, akkor a (5.1.5) tétel szerint a $\ker\,(A-\lambda I)$ 1-dimenziós A-invariáns altér, vagy 1-dimenziós A-invariáns alterek direktösszege. Az A lináris transzformáció minimálpolinomjának az $\mathbb{F}$ testben lévő gyökeit, az A *sajátértékeinek* nevezzük. Ha λ sajátértéke A-nak, akkor a $\ker\,(A - \lambda I)$ altér minden nemzéró s vektorát az A lineáris transzformáció *sajátvektorának* hívjuk. Magára a $\ker(A - \lambda I)$ altérre gyakran a λ-hoz tartozó *sajátaltér* néven hívatkozunk. A sajátaltér tetszőleges s vektorára

$$A(s) = (A - \lambda I)(s) + \lambda s = \lambda s$$

teljesül.

Másrészt, ha $A - \lambda I$ szinguláris transzformációja V-nek, akkor van nemzéró v vektora $\ker(A - \lambda I)$-nek. Az $m(t)$ minimálpolinomnak a $t - \lambda \in \mathbb{F}[t]$ elsőfokú polinommal való maradékos osztását végezve

$$m(t) = q(t)(t - \lambda) + \gamma$$

adódik, hiszen a maradék csak 0-adfokú, azaz skalár lehet. Akkor az

$$m(A) = q(A)(A - \lambda I) + \gamma I$$

transzformációval

$$0 = m(A)(v) = q(A)(A - \lambda I)(v) + \gamma I(v) = \gamma v\,,$$

amiből kapjuk, hogy $\gamma = 0$, azaz $t - \lambda$ osztója $m(t)$-nek. De akkor $m(\lambda) = 0$, azaz λ gyöke a minimálpolinomnak.

A (5.1.5) tétel kiegészítve a fenti észrevétellel a következő állítást verifikálja:

5.2.1 Tétel. *Ha V az $\mathbb{F}$ test feletti vektortér és A lineáris transzformációja, akkor az A minimálpolinomjának a $\lambda \in \mathbb{F}$ skalár pontosan akkor gyöke, azaz λ pontosan akkor sajátértéke A-nak, ha az $A - \lambda I \in L(V)$ lineáris transzformáció szinguláris.*

A (5.2.1) tételből azonnal adódik:

5.2.2 Következmény. *Egy lineáris transzformációnak akkor és csak akkor sajátértéke a 0, ha a transzformáció szinguláris.*

Sok esetben használhatók a következő észrevételek:

5.2.3 Állítás. (a) *Az $A \in L(V)$ (V-ben értelmezve van skaláris szorzat) lineáris transzformációnak λ akkor és csak akkor sajátértéke, ha transzponáltjának sajátértéke.*
(b) *Ha az $A \in L(V)$ lineáris transzformációnak λ sajátértéke, akkor tetszőleges $p(t) \in \mathbb{F}[t]$ polinomra a $p(A)$ lineáris transzformációnak sajátértéke $p(\lambda)$, és ha A invertálható, akkor az A^{-1} lineáris transzformációnak sajátértéke $\frac{1}{\lambda}$.*

Bizonyítás. Az (a) állítás annak a következménye, hogy a lineáris transzformáció és transzponáltjának egyenlő a rangja és mivel

$$(A - \lambda I)^* = A^* - \lambda I\,,$$

az $A - \lambda I$ lineáris transzformáció pontosan akkor szinguláris, ha az $A^* - \lambda I$ lineáris transzformáció szinguláris.
(b) Az A lineáris transzformációnak λ pontosan akkor sajátértéke, a (5.2.1) tétel szerint, ha van olyan $s(\neq 0) \in V$ vektor, hogy $A(s) = \lambda s$. Akkor $A^2(s) = A(A(s)) = A(\lambda s) = \lambda A(s) = \lambda^2 s$, és általánosan minden k természetes számra $A^k(s) = \lambda^k s$ teljesül. Ebből már következik, hogy bármely $p(t) \in \mathbb{F}[t]$ polinomra $p(A)(s) = p(\lambda)s$,

igazolva első állításunkat, hogy $p(\lambda)$ sajátértéke a $p(A)$ transzformációnak.
Az utolsó állítás következik az

$$s = A^{-1}(A(s)) = A^{-1}(\lambda s) = \lambda A^{-1}(s)$$

egyenlőségből, mert a (5.2) következmény alapján $\lambda \neq 0$ és így

$$A^{-1}(s) = \frac{1}{\lambda} s\,.$$

$\square$

Az alábbi példa azt mutatja be, hogy a (5.2.1) tétel hogyan használható egy vektortér sajátértékeinek és sajátvektorainak meghatározására.

2 Példa. *Határozzuk meg a 3-dimenziós valós V vektortér azon A lineáris transzformációjának sajátértékeit és sajátvektorait, amely a tér egy $\mathcal{V} = \{v_1, v_2, v_3\}$ bázisának vektorait rendre az $A(v_1) = v_2 + v_3$, $A(v_2) = v_1 + v_3$, $A(v_3) = v_1 + v_2$ vektorokba viszi!*

Az A, illetve az $A - \lambda I$ transzformációk $\mathcal{V}$ bázisra vonatkozó mátrixai

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad \text{és} \qquad \mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{bmatrix}$$

Állapítsuk meg, hogy a λ paraméter mely értékei esetén lesz a $\mathbf{A} - \lambda\mathbf{I}$ mátrix szinguláris. Az elemi bázistranszformációs technika most is alkalmazható, hiszen kérdésünk úgy is feltehető, hogy a λ praméter milyen értékei mellett van nemtriviális megoldása az $\mathbf{A} - \lambda\mathbf{E}$ együtthatómátrixú homogén lineáris egyenletrendszernek. Az ismeretlenekre bevezetve a ξ_1, ξ_2, ξ_3 jelölést a

$$\begin{array}{c|ccc|} & \xi_1 & \xi_2 & \xi_3 \\ \hline v_1 & -\lambda & 1 & 1 \\ v_2 & \boxed{1} & -\lambda & 1 \\ v_3 & 1 & 1 & -\lambda \\ \hline \end{array} \quad \longrightarrow \quad \begin{array}{c|cc|} & \xi_2 & \xi_3 \\ \hline v_1 & 1-\lambda^2 & 1+\lambda \\ \xi_1 & -\lambda & 1 \\ v_3 & \boxed{1+\lambda} & -\lambda-1 \\ \hline \end{array}$$

táblázatokból leolvashatjuk, hogy $\lambda = -1$ esetén a

$$\begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \tau_2 \\ \tau_3 \end{bmatrix}$$

koordináta vektorú vektor nemtriviális megoldás, ha a τ_2, τ_3 szabadon választható skalárok közül legalább az egyik nem nulla. Például az

$$\mathbf{s}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad \text{és} \qquad \mathbf{s}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

koordináta vektorú lineárisan független vektorok a transzformáció $\lambda = -1$ sajátértékéhez tartozó sajátvektorai. Visszatérve a táblázathoz, látható, hogy ha $\lambda \neq -1$, akkor további báziscsere hajtható végre és kapjuk a

	ξ_3
v_1	$2+\lambda-\lambda^2$
ξ_1	$1-\lambda$
ξ_2	-1

táblázatot, miszerint $\lambda = 2$ esetén is van nemtriviális megoldás, amelynek alakja

$$\begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \tau_3 \;\; (\tau_3 \neq 0)\,.$$

Igy a $\lambda = 2$ sajátérték, és egy ehhez tartozó sajátvektor koordináta vektora az

$$\mathbf{s}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

A feladathoz nem tartozik ugyan, de érdemes meghatározni az A transzformáció mátrixát a sajátvektorok alkotta

$$\mathcal{S} = \{s_1 = -v_1 + v_2,\; s_2 = v_1 + v_3,\; s_3 = v_1 + v_2 + v_3\}$$

bázisban. Mivel $A(s_1) = -s_1,\; A(s_2) = -s_2$ és $A(s_3) = 2 \cdot s_3$, kapjuk, hogy

$$\mathbf{A}_{\mathcal{S}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

diagonális mátrix, amelynek diagonálisában éppen az A sajátértékei vannak. □

A fenti példában azt láttuk, hogy a vizsgált lineáris transzformáció sajátvektorai bázisát alkották a tárgyvektortérnek, és ebben a bázisban a lineáris transzformáció mátrixa diagonális mátrix lett. Ez általánosan is igaz, nevezetesen:

5.2.4 Tétel. *Ha az n-dimenziós $\mathbb{F}$ test feletti V vektortér $A \in L(V)$ lineáris transzformációjának sajátvektoraiból álló $\mathcal{S} = \{s_1, \ldots, s_n\}$ vektorrendszere bázis, akkor az A mátrixa az $\mathcal{S}$ bázisban diagonális mátrix, melynek diagonálisában éppen az egyes sajátvektorokhoz tartozó sajátértékek vannak.*

Bizonyítás. Jelöljék az egyes sajátvektorokhoz tartozó sajátértékeket rendre $\lambda_1, \ldots, \lambda_n$. Akkor, tekintve, hogy

$$A(s_i) = \lambda_i s_i \qquad (i = 1, \ldots, n)\,,$$

kapjuk, hogy

$$\mathbf{A}(s_i)_{\mathcal{S}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

ahol éppen az i-edik koordináta a λ_i, a többi pedig zéró. Mivel az A lineáris transzformáció $\mathbf{A}_{\mathcal{S}}$ mátrixának i-edik oszlopa éppen $\mathbf{A}(s_i)_{\mathcal{S}}$ minden $(i = 1, \ldots, n)$-re,

$$\mathbf{A}_{\mathcal{S}} = \mathbf{B} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

amint állítottuk. □

Egy vektortér olyan lineáris transzformációját, amelynek sajátvektorai generálják a teret, tehát mátrixa diagonalizálható, *egyszerű*, vagy más elnevezéssel *diagonalizálható* lineáris transzformációnak hívjuk. Megjegyzendő, hogy az egyáltalán nem biztos, hogy a különböző sajátvektorok különböző sajátértékekhez tartoznak, amint az (2) példa is mutatta, (a -1 sajátértékhez két egymástól lineárisan független sajátvektor tartozott). Másrészt viszont igaz, hogy különböző sajátértékekhez tartozó sajátvektorok lineárisan független rendszert alkotnak. Ezt az állítást fogalmaztuk meg a következő tételben.

5.2.5 Tétel. *Ha a V vektortér A lineáris transzformációjának $\lambda_1, \ldots, \lambda_k$ különböző sajátértékei, akkor a hozzájuk tartozó $s_1, \ldots, s_k$ sajátvektorok lineárisan független rendszert alkotnak.*

Bizonyítás. A különböző sajátértékek száma szerinti teljes indukcióval igazoljuk az állítást. Ha $k = 1$, akkor a sajátvektor nem nullvektor lévén lineárisan független egyelemű rendszer. Tegyük fel, hogy a $\lambda_1, \ldots, \lambda_{k-1}$ különböző sajátértékekhez tartozó sajátvektorok $\{s_1, \ldots, s_{k-1}\}$ rendszere lineárisan független és legyen λ_k az eddigiektől különböző sajátérték és s_k a hozzá tartozó sajátvektor. Indirekt tegyük fel, hogy

$$(a) \qquad s_k = \sum_{i=1}^{k-1} \alpha_i s_i\,,$$

azaz, hogy az $\{s_1, \ldots, s_{k-1}, s_k\}$ vektorrendszer már lineárisan összefüggő. Ha alkalmazzuk az A transzformációt az (a) egyenlettel adott s_k vektorra, akkor azt kapjuk, hogy

$$\lambda_k s_k = A(s_k) = \sum_{i=1}^{k-1} \alpha_i A(s_i) =$$

$$(b) \qquad = \sum_{i=1}^{k-1} \alpha_i \lambda_i s_i \,.$$

Igy az (a) egyenlet λ_k-szorosát kivonva a (b) egyenletből a

$$0 = \sum_{i=1}^{k-1} \alpha_i (\lambda_i - \lambda_k) s_i$$

adódik. Ez lehetetlen, hiszen $s_k \neq 0$ (mert sajátvektor), így az α_i skalárok valamelyike nemnulla. Másrészt a sajátértékek különbözősége miatt a $\lambda_i - \lambda_k$, $(i = 1, \ldots, k-1)$ skalárok is különböznek nullától és az indukciós feltevés szerint az $\{s_1, \ldots, s_{k-1}\}$ rendszer lineárisan független volt. Ez az ellentmondás abból eredt, hogy feltételeztük, hogy az $\{s_1, \ldots, s_{k-1}, s_k\}$ vektorrendszer lineárisan összefüggő, tehát lineárisan független kell legyen. □

Ha egy n-dimenziós V vektortér A lineáris transzformációjának n különböző sajátértéke van, akkor A egyszerű. Ez a feltétel elegendő, de nem szükséges, amint azt a (2) példa is mutatta.

5.3 A sík elemi lineáris transzformációi

Minden kétdimenziós valós vektortér izomorf a sík egy rögzített pontjából kiinduló helyvektorok terével. Mivel a a véges dimenziós valós terek lineáris transzformációit azzal kívánjuk jellemezni, hogy hogyan viselkednek a tér szemléltethető alterein, célszerűnek látszik áttekinteni, hogy melyek a sik helyvektorainak elemi lineáris transzformációi. Az elemi jelzővel itt arra akarunk utalni, hogy csupán azokat a lineáris transzformációkat kívánjuk felsorolni, amelyek szorzataként a sík minden lineáris transzformációja előállítható. Ahhoz, hogy egy lineáris transzformáció hatását, geometriai jelentését jól érzékelhessük, általában speciális bázist (koordinátarendszert) kell választanunk. Az alábbiakban a középiskolából már jól ismert egységnyi hosszúságú és egymásra merőleges $\{i, j\}$ helyvektorok alkotta bázisban jellemezzük a sík elemi lineáris transzformációit:

(1) Nyújtás/zsugorítás: Kétféle nyújtásról/zsugorításról beszélhetünk,

- **(a) középpontos** nyújtásról/zsugorításról, amikor a sík minden helyvektorát a transzformáció pozitív λ-szorosába viszi, illetve
- **(b) tengelyes** nyújtásról, amikor csak az egyik tengely irányába eső összetevő nyúlik meg/zsugorodik össze, azaz egy $\varrho i + \theta j$ vektor képe a $\lambda \varrho i + \theta j$ vektor $(\lambda > 0)$.

(2) Tükrözés: A tükrözés is lehet

- **(a) középpontos**, amikor minden vektor az ellentettjébe transzformálódik, vagy
- **(b) tengelyes** tükrözés esetén, pontosabban a j irányú tengelyre való tükrözés esetén a $\varrho i + \theta j$ vektor képe a $-\varrho i + \theta j$ vektor.

(3) Vetítés:

(a) **középpontra** való vetítésen értjük, azt, amikor minden vektort az origóra, azaz a nullvektorba transzformálunk, és

(b) **tengelyre** való vetítés az amikor minden vektort az egyik, mondjuk az i irányú tengellyel párhuzamosan a j irányú tengelyre vetítünk. Ez a lineáris transzformáció a $\varrho i + \theta j$ vektort a θj vektorba viszi.

(4) Párhuzamos affinitás: az egyik mondjuk az i irányú tengely pontjainak j-vel párhuzamos eltolása a két bázisvektor szögfelező egyenesébe. Ez a transzformáció a $\varrho i + \theta j$ vektort a $\varrho i + (\varrho + \theta)j$ vektorba transzformálja.

(5) Forgatás: minden vektort forgassunk el ϕ szöggel (az i vektort j felé mozgatva). Ekkor az i vektor képe $\cos\phi i + \sin\phi j$, míg a j vektor képe $-\sin\phi i + \cos\phi j$ lesz, így egy tetszőleges $\varrho i + \theta j$ vektor a $(\varrho\cos\phi - \theta\sin\phi)i + (\varrho\sin\phi + \theta\cos\phi)j$ vektorba transzformálódik.

Az olvasó könnyen beláthatja, hogy a sík felsorolt transzformációi valóban *lineáris* transzformációk. Javasoljuk, hogy mindegyik transzformációnak állapítsa meg a mátrixát az i, j bázisban.

Következő feladatunk annak igazolása, hogy valóban a sík minden lineáris transzformációja a fentiekben felsorolt transzformációk szorzatára bontható. Ennek érdekében a sík transzformációit annak megfelelően osztályozzuk, hogy hány különböző irányú sajátvektoruk van.

(a) **Egyetlen sajátértékhez tartozó két lineárisan független sajátvektor** esetén a sík minden nemzéró vektora ugyanazon sajátértékhez tartozó sajátvektor. Valóban, ha e és f a sík A lineáris transzformációjának két lineárisan független sajátvektora és mindkettő a λ sajátértékhez tartoznak, azaz $A(e) = \lambda e$ és $A(f) = \lambda f$, akkor a sík minden v vektora kifejezhető

$$v = \alpha e + \beta f$$

alakban, és akkor

$$A(v) = A(\alpha e + \beta f) = \alpha A(e) + \beta A(f) = \alpha\lambda e + \beta\lambda f = \lambda(\alpha e + \beta f) = \lambda v\,,$$

alátámasztva kijelentésünket.
Akkor az A transzformáció nyújtás ($\lambda \geq 1$), vagy zsugorítás ($0 < \lambda < 1$)), vagy középpontos vetítés ($\lambda = 0$), vagy középpontos tükrözés és nyújtás/zsugorítás szorzata ($\lambda < 0$).

(b) **Ha két különböző sajátérték van.** Jelölje megint A a lineáris transzformációt és e a λ-hoz tartozó, míg f a μ sajátértékhez tartozó sajátvektorokat. Az e, f vektorok most is bázist alkotnak, hiszen lineárisan függetlenek a (5.2.5) tétel szerint. Bontsuk fel A-t a $B(e) = \lambda e, B(f) = f$ és a $C(e) = e, C(f) = \mu f$ egyenlőségekkel meghatározott lineáris transzformációk szorzatára. Akkor mind B, mind C az e, illetve f irányú tengely irányába történő nyújtás/zsugorítás, vagy a rájuk merőleges tengelyre való tükrözés és a fentiek szorzata, esetleg valamelyikük középpontos vetítés, és A mindezek szorzata.

(c) **Egy sajátvektor van.** Jelölje e a sajátvektort és legyen t erre merőleges vektora a síknak, akkor e, t bázis. Ha λ a megfelelő sajátérték, akkor $A(e) = \lambda e$ és $A(t) =$

$\alpha e + \beta t$ és $\alpha \neq 0$, mert különben t is sajátvektor lenne. Megmutatjuk, hogy $\beta = \lambda$. Tekintsük a $\alpha e + (\beta - \lambda)t$ vektort.

$$A(\alpha e + (\beta - \lambda)t) = \alpha\lambda e + (\beta - \lambda)(\alpha e + \beta t) = \beta\alpha e + (\beta - \lambda)t),$$

ami mutatja, hogy $\beta - \lambda = 0$, mert különben az $\alpha e + (\beta - \lambda)t$ vektor sajátvektor lenne, és nem lenne e irányú. Azt kaptuk tehát, hogy $A(t) = \alpha e + \lambda t$.
(1) ha $\lambda = 0$ akkor $A = B \cdot C \cdot D$, lineáris transzformációk szorzata, ahol $B(e) = \alpha e$, $B(t) = \alpha t$ nyújtás/zsugorítás középpontos tükrözés, vagy ezek szorzata, $C(e) = -t$, $C(t) = e$ $\frac{\pi}{2}$ szöggel való forgatás, és $D(e) = 0$, $D(t) = t$ az e irányú tengellyel párhuzamos vetítés.
(2) ha $\lambda \neq 0$, akkor az $u = \frac{\alpha}{\lambda}e, t$ ortogonális bázisban $A(u) = \lambda u$, $A(t) = \lambda u + \lambda t$ felbontható $A = B \cdot C$ szorzatra, ahol $B(u) = u$, $B(t) = u + t$ párhuzamos affinitás és $C(u) = \lambda u$, $C(t) = \lambda t$ középpontos nyújtás/zsugorítás, vagy tükrözés és az előbbi szorzata.
(d) **Ha nincs sajátvektor,** akkor az azt jelenti, hogy az A transzformáció minimálpolinomjának gyökei komplex számok. Ebben az esetben A azonosítása, mint elemi lineáris transzformációk szorzata a valós térben maradva igen nehézkes, pontosabban annak a bázisnak a megkeresése, amelyben ez az azonosítás bemutatható nem könnyű. Ezért felhasználjuk, hogy az sík A lineáris transzformációjához tartozik a 2-dimenziós komplex tér egy $\widehat{A}$ lineáris transzformációja, amelynek minimálpolinomja megegyezik A minimálpolinomjával. Így, ha A minimálpolinomjának komplex gyökei $\lambda = \alpha + i\beta$ és $\bar{\lambda} = \alpha - i\beta$ konjugált komplex számok, akkor azok a $\widehat{A}$ lineáris transzformációnak is sajátértékei és a hozzá tartozó komplex sajátvektorok $x = a+ib$ és $\bar{x} = a - ib$ alakúak. Ez abból következik, hogy

$$(\widehat{A} - \lambda)(a + ib) = A(a) + iA(b) - \lambda(a + ib) = 0,$$

és akkor konjugálva mindkét oldalt, kapjuk, hogy

$$0 = A(a) - iA(b) - \bar{\lambda}(a - ib) = (\widehat{A} - \bar{\lambda})(a - ib).$$

Összehasonlítva a valós és képzetes részeket bármelyik egyenlőségből

$$A(a) = \alpha a - \beta b \quad \text{és} \quad A(b) = \beta a + \alpha b$$

adódik, ahol $\alpha, \beta \in \mathcal{R}$ és $a, b \in V$ és $\beta \neq 0$, továbbá a és b lineárisan függetlenek, mert A-nak nincs valós sajátértéke. Az $\alpha^2 + \beta^2 = \delta^2$ jelöléssel az $\frac{\alpha}{\delta}$ és $\frac{\beta}{\delta}$ egy egyértelműen meghatározott ϕ szög koszinusza, illetve szinusza és az A lineáris transzformáció egy forgatás és nyújtás/zsugorítás egyidejü végrehajtását jelenti.

5.4 Eklideszi terek lineáris transzformációi

Az euklideszi terek lineáris transzformációról azt lehet tudni a (5.1.5) tétel alapján, hogy van a térnek 2-dimenziós, a transzformációra nézve invariáns altere. Sajnos ha M az A transzformációnak invariáns altere és az $M^\perp$ az M-nek ortogonális kiegészítője, akkor az általános esetben $M^\perp$ nem feltétlen A-invariáns altér. Ezért az euklideszi tereknek vannak olyan lineáris transzformációi, hogy a tér nem bontható fel a tekintett transzformációra invariáns 2-dimenziós altereinek direktösszegére. Ez azt jelenti, hogy kénytelenek vagyunk lemondani arról a törekvésünkről, hogy minden lineáris transzformációt a tér szemléltethető invariáns alterein való hatásával jellemezzünk, ebből a szempontból a tekintett lineáris transzformációk körét szűkítenünk kell. Annak érdekében, hogy megállapíthassuk, hogy melyek azok a lineáris transzformációk, amelyeket lehet alterein való hatásukkal ”vizuálissá” tenni, bizonyítjuk a következő tételt:

5.4.1 Tétel. *Ha az E euklideszi tér M altere az $A \in L(E)$ lineáris transzformációra nézve invariáns, akkor $M^\perp$, ortogonális kiegészítője invariáns A^*-ra nézve, ahol A^* az A transzponáltját jelöli.*

Bizonyítás. Legyen w tetszőleges vektora $M^\perp$-nak és v bármelyik vektora M-nek. Akkor

$$\langle A^*(w), v\rangle = \langle w, A(v)\rangle = 0\,,$$

hiszen $A(v) \in M$, igazolva, hogy $A^*(w)$ ortogonális bármely $v \in M$ vektorra, azaz $A^*(w) \in M^\perp$. □

A fenti eredmény ismeretében vizsgálatainkat olyan lineáris transzformációkra korlátozzuk, amelyekre nézve invariáns alterek egyszersmind transzponáltjaikra nézve is invariánsak. Ha az A lineáris transzformációja az E euklideszi térnek ilyen, akkor a tér felbontható ”szemléltethető” A-invariáns alterei direktösszegére. Ezt igazoljuk az alábbi tételben.

5.4.2 Tétel. *Ha az E euklideszi térnek az A lineáris transzformációja rendelkezik azzal a tulajdonsággal, hogy minden invariáns altere transzponáltjára nézve is invariáns, akkor E legfeljebb 2-dimenziós, A-invariáns, páronként ortogonális altereinek direktösszege.*

Bizonyítás. Az E dimenziója szerinti teljes indukcióval bizonyítjuk az állítást. Ha E dimenziója nem nagyobb mint kettő, akkor nyilvánvalóan igaz rá a tétel. Tegyük fel, hogy n-nél alacsonyabb dimenziójú euklideszi terekre már tudjuk, hogy felbonthatók a feltételezett tulajdonságú transzformációra nézve invariáns, páronként ortogonális, legfeljebb 2-dimenziós alterei direktösszegére. Ezekután legyen E n-dimenziós. Az A transzformációnak a (5.1.5) tétel szerint, van legfeljebb 2-dimenziós invariáns $M \neq \{0\}$ altere E-ben, és ez A^*-ra nézve is invariáns. Alkalmazva a (5.4.1) tételt A helyett A^*-ra, kapjuk, hogy $M^\perp$ A-invariáns altér, és dimenziója legfeljebb $n-1$. Akkor az indukciós feltevés szerint $M^\perp$ felbomlik páronként ortogonális legfeljebb 2-dimenziós A-invariáns alterei direktösszegére és ezek az alterek mind ortogonálisak M-re is. Így az M altérrel együtt az E kívánt direkt felbontását szolgáltatják. □

Az euklideszi terek két olyan lineáris transzformáció tipusával foglalkozunk, amelyek kielégítik a fenti tétel feltételét. Ilyenek nyilvánvalóan a szimmetrikus transzformációk ($A^* = A$) és az ortogonális transzformációk ($A^* = A^{-1}$).

5.4.1 Szimmetrikus lineáris transzformációk

Megmutatjuk, hogy a szimmetrikus lineáris transzformációk diagonalizálhatók. Ehhez a fenti (5.4.2) tétel szerint elegendő azt igazolnunk, hogy ha A szimmetrikus és M az euklideszi tér kétdimenziós A-invariáns altere, akkor van M-ben A-nak sajátvektora. Vegyük észre, hogy ha $s \in M$ sajátvektora A-nak és $t \in M$ ortogonális

s-re, akkor

$$\langle A(t), s\rangle = \langle t, A(s)\rangle = \langle t, \lambda s\rangle = \lambda \langle t, s\rangle = 0\,,$$

ami csak úgy lehet (mert M 2-dimenziós), ha $A(t)$ a t vektor skalárszorosa, azaz t is sajátvektora A-nak. A sajátvektor létezését bizonyítandó, legyen u és v a két bázisvektora M-nek. Akkor

$$A(u) = \alpha u + \beta v \qquad \text{és} \qquad A(v) = \beta u + \gamma v$$

alakú, mert A szimmetrikus. Határozzuk meg az A M-re való leszűkítésének minimálpolinomját.

$$A^2(u) = \alpha A(u) + \beta A(v) = (\alpha^2+\beta^2)u + (\alpha\beta + \beta\gamma)v$$

és

$$A^2(v) = \beta A(u) + \gamma A(v) = (\alpha\beta+\beta\gamma)u + (\beta^2+\gamma^2)v\,,$$

(1)
$$\begin{array}{c|cc} & A(u) & A^2(u) \\ \hline u & \alpha & \alpha^2+\beta^2 \\ v & \boxed{\beta} & \beta(\alpha+\gamma) \\ \hline \end{array} \;\rightarrow\; \begin{array}{c|c} & A^2(u) \\ \hline u & \beta^2-\alpha\gamma \\ A(u) & \alpha+\gamma \\ \hline \end{array}$$

(2)
$$\begin{array}{c|cc} & A(v) & A^2(v) \\ \hline u & \boxed{\beta} & \beta(\alpha+\gamma) \\ v & \gamma & \alpha^2+\beta^2 \\ \hline \end{array} \;\rightarrow\; \begin{array}{c|c} & A^2(u) \\ \hline A(v) & \alpha+\gamma \\ v & \beta^2-\alpha\gamma \\ \hline \end{array}$$

(1)-ből és (2)-ből következik, hogy

$$A^2(u) - (\alpha+\gamma)A(u) - (\beta^2-\alpha\gamma)u = 0 \qquad \text{és}$$

$$A^2(v) - (\alpha+\gamma)A(v) - (\beta^2-\alpha\gamma)v = 0\,,$$

de akkor tetszőleges M-beli $m = \mu u + \nu v$ vektorra

$$A^2(m) - (\alpha+\gamma)A(m) - (\beta^2-\alpha\gamma)m =$$

$$\mu(A^2(u) - (\alpha+\gamma)A(u) - (\beta^2-\alpha\gamma)u) + \nu(A^2(v) - (\alpha+\gamma)A(v) - (\beta^2-\alpha\gamma)v) = 0\,,$$

azaz $A^2 - (\alpha+\gamma)A + (\alpha\gamma-\beta^2)I$ lineáris transzformáció az M-en a zérótranszformáció. Ez azt jelenti, hogy A M-re való leszűkítésének minimálpolinomja – ami az A minimálpolinomjának nyilván osztója –

$$m_M(t) = t^2 - (\alpha+\gamma)t + \alpha\gamma - \beta^2\,.$$

Ennek a polinomnak a gyökei valósak, mert diszkriminánsa

$$(\alpha+\gamma)^2 - 4(\alpha\gamma-\beta^2) = (\alpha-\gamma)^2 + 4\beta^2 \geq 0\,.$$

Ez biztosítja, hogy A-nak van sajátvektora M-ben. Akkor viszont, mint azt előbb beláttuk, az M erre a sajátvektorra ortogonális nemzéró vektora is sajátvektora A-nak, és M két ortogonális egydimenziós altér direktösszege. A (5.4.2) tétellel összevetve a most kapott eredményt, állíthatjuk, hogy az euklideszi terek szimmetrikus lineáris transzformációinak vannak olyan sajátvektorai, amelyek a tér ortonormált bázisát alkotják. Másrészt, ha egy lineáris transzformációnak a sajátvektoraiból összerakható az euklideszi tér egy ortonormált bázisa, akkor abban a bázisban a transzformáció mátrixa diagonális mátrix, amiből következik, hogy a transzformáció szimmetrikus. Gyűjtsük össze a kapott eredményeket az alábbi tételben:

5.4.3 Tétel. *Az E euklideszi tér A lineáris transzformációja pontosan akkor szimmetrikus, ha vannak olyan sajátvektorai, amelyek a tér ortonormált bázisát alkotják.*

5.4.2 Ortogonális lineáris transzformációk

Már említettük, hogy ha A ortogonális lineáris transzformációja az E euklideszi térnek, E akkor is felbomlik legfeljebb 2-dimenziós A-invariáns alterei direktösszegére.

Az ortogonális lineáris transzformációk karakterisztikus tulajdonsága, hogy a skaláris szorzatot változatlanul hagyják. Ha A ortogonális lineáris transzformáció, akkor tetszőleges $v, w \in E$ vektorokra

$$\langle A(v), A(w)\rangle = \langle A^* A(v), w\rangle = \langle v, w\rangle \ .$$

Másrészt ha A változatlanul hagyja a skaláris szorzatot, akkor

$$\langle v, w\rangle = \langle A(v), A(w)\rangle = \langle (A^* A)(v), w\rangle \ ,$$

ami tetszőleges vektorpár estén csak úgy teljesülhet, ha $A^* = A^{-1}$. Ez ekvivalens azzal a feltétellel, hogy egy lineáris transzformáció pontosan akkor ortogonális, ha a tér ortonormált bázisát ortonormált bázisba viszi. Ezért bármely ortonormált bázisra vonatkozó mátrixa i-edik és j-edik oszlopának belső szorzata δ_{ij}.

Ha M a tér 2-dimenziós A-invariáns altere, akkor most nem feltétlenül van A-nak sajátértéke M-ben, de ha $s \in M$ sajátvektor, akkor az M s-re ortogonális nemzéró vektorai is sajátvektorok, mint azt az

$$\langle s, A(t)\rangle = \left\langle A^{-1}(s), t\right\rangle = \frac{1}{\lambda}\langle s, t\rangle = 0$$

egyenlőség mutatja. (λ-val az s-hez tartozó sajátértéket jelöltük.) Feltehetjük, hogy s egységnyi normájú. Akkor

$$1 = \langle s, s\rangle = \langle A(s), A(s)\rangle = \lambda^2 \,,$$

tehát ortogonális lineáris transzformáció sajátértéke csak 1 vagy -1 lehet. Ha nincs A-nak M-ben sajátvektora, akkor az u és v ortonormált bázisvektorokra

$$A(u) = \alpha u + \beta v \qquad \text{és} \qquad A(v) = \gamma u + \delta v \quad (\beta \neq 0 \neq \gamma) \,,$$

ahol az A ortogonalitása miatt a skaláregyütthatók ki kell elégítsék az

$$(a) \qquad \alpha^2 + \beta^2 = 1 \,,$$

$$(b) \qquad \alpha\gamma + \beta\delta = 0 \,,$$

$$(c) \qquad \gamma^2 + \delta^2 = 1$$

egyenleteket. Jelölje ϕ az $A(u)$ vektor u vektorral bezárt szögét, akkor $\cos\phi = \langle u, A(u)\rangle = \alpha$ és az (a) egyenlet szerint $\beta = \sin\phi$, vagy $\beta = -\sin\phi$. A (b) egyenletből átrendezéssel kapjuk, hogy $-\delta = \alpha\gamma/\beta$, amit a (c) egyenletbe helyettesítve

$$\gamma^2 + \frac{\alpha^2\gamma^2}{\beta^2} = \frac{\gamma^2(\beta^2+\alpha^2)}{\beta^2} = \frac{\gamma^2}{\beta^2} = 1\,.$$

Akkor ebből vagy $\gamma = \beta$ és a (b) egyenlet alapján $\delta = -\alpha$, vagy $\gamma = -\beta$ és $\delta = \alpha$ következik. Mivel A nem szimmetrikus, hiszen akkor lenne sajátvektora, csak a második eset állhat fenn. Így az A ortogonális transzformáció M-re való leszűkítésének mátrixa az u, v ortonormált bázisban az alábbiak egyike:

$$\mathbf{A} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \quad \text{vagy} \quad \mathbf{A} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix},$$

ami pozitív irányú, vagy negatív irányú ϕ szöggel való forgatást jelent.

5.5 Kvadratikus alakok

A többváltozós függvények lokális szélsőértékeinek meghatározásakor beleütközünk a következő problémába. Meg kell állapítanunk, hogy egy

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \xi_i\xi_j\alpha_{ij}$$

alakú másodfokú függvény a $\xi_i \ (i = 1,\ldots,n)$ változók minden értéke mellett azonos előjelű-e. Ez a kérdés motiválja a véges dimenziós valós vektor tereken értelmezett kvadratikus alakok vizsgálatát.

5.5.1 Definíció. *Legyen V n-dimenziós valós vektortér és $\mathcal{X} = \{x_1,\ldots,x_n\}$ egy bázisa. Tetszőleges*

$$v = \epsilon_1 x_1 + \cdots + \epsilon_n x_n$$

vektorhoz rendeljük hozzá a

$$\mathcal{Q}(v) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}\epsilon_i\epsilon_j$$

valós számot, ahol az $\alpha_{ij} \ (i,j = 1,\ldots,n)$ adott valós skalárok. Az így kapott

$$\mathcal{Q} : V \to \mathcal{R}$$

függvényt kvadratikus alaknak nevezzük.

Az elnevezést az indokolja, hogy a kvadratikus alak a vektorváltozó koordinátáinak másodfokú függvénye. Ilyen függvényekkel adhatók meg a sík úgynevezett másodrendű görbéi (kör, elipszis, parabola, hiperbola...).

Mátrixaritmetikai eszközöket használva, a kvadratikus alakok egyszerűbb formában is megadhatók. Gyűjtsük össze az együtthatókat a $\mathbf{A} = [\alpha_{ij}]$ $n \times n$ tipusú mátrixba, és akkor

$$\mathcal{Q}(v) = \mathbf{v}^* \mathbf{A} \mathbf{v}$$

mátrix szorzatként is megadható, ahol $\mathbf{v}$ a v vektor $\mathcal{X}$ bázisra vonatkozó koordinátavektora, és $\mathbf{v}^*$ az ebből készített sormátrix.

Jelölje $A \in L(V)$ azt a lineáris transzformációt, amelynek $\mathcal{X}$ bázisra vonatkozó mátrixa éppen $\mathbf{A}$, azaz legyen minden $j(= 1, \ldots, n)$-re

$$A(x_j) = \alpha_{1j} x_1 + \cdots + \alpha_{ij} x_i + \cdots + \alpha_{nj} x_n \,.$$

Akkor a kvadratikus alak az $\mathcal{X}$ bázis által meghatározott

$$\mathcal{Q}(v) = [v, A(v)]_{\mathcal{X}}$$

belső szorzatként is definiálható.

Amint azt az euklideszi terek tárgyalásakor láttuk, tetszőleges véges dimenziós valós vektortérben értelmezhető skaláris szorzat valamely $\mathcal{X}$ bázisra vonatkozó belső szorzattal, és akkor nyilvánvalóan $\mathcal{X}$ erre a skaláris szorzatra nézve ortonormált bázis lesz. Azt is igazoltuk, hogy fordítva is igaz, véges dimenziós euklideszi térben a skaláris szorzat, valamely ortonormált bázis szerinti belső szorzat.

A fentiekben leírt megfontolások a következő tételt igazolják:

5.5.2 Tétel. *Az V euklideszi téren értelmezett*

$$\mathcal{Q} : V \to \mathcal{R}$$

függvény pontosan akkor kvadratikus alak, ha létezik olyan $A \in L(V)$ lineáris transzformáció, hogy minden $v \in V$-re

$$\mathcal{Q}(v) = \langle v, A(v) \rangle$$

teljesül.

A kvadratikus alakok általunk adott definíciója felveti a kérdést, hogy vajon a kvadratikus alakhoz rendelt $\mathbf{A}$ mátrix és így a hozzárendelt A lineáris transzformáció egyértelmű-e. Könnyen látható, hogy a válasz nem. Ugyanis, ha β_{ij} $(i, j = 1, \ldots, n)$ valós skalároknak olyan rendszere, hogy minden i, j párra az

$$\alpha_{ij} + \alpha_{ji} = \beta_{ij} + \beta_{ji}$$

egyenlőségek teljesülnek, akkor az ϵ_i $(i = 1, \ldots, n)$ változók minden értékére

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} \epsilon_i \epsilon_j = \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \epsilon_i \epsilon_j \,.$$

Mátrixokkal megfogalmazva ugyanez: ha

$$\mathbf{B} + \mathbf{B}^* = \mathbf{A} + \mathbf{A}^* \,,$$

akkor minden $\mathbf{v}$-re

$$\mathbf{v}^*\mathbf{B}\mathbf{v} = \mathbf{v}^*\mathbf{A}\mathbf{v}\,.$$

Mivel rögzitett bázis mellett, kölcsönösen egyértelmű megfeleltetés létezik a V tér lineáris transzformációi és a $\dim V \times \dim V$ tipusú mátrixok között, azt is állíthatjuk, hogy $A, B \in L(V)$-re

$$B + B^* = A + A^* \Longrightarrow (\forall v \in V) : \langle v, B(v)\rangle = \langle v, A(v)\rangle\ .$$

Igaz viszont a következő tétel:

5.5.3 Tétel. *A V euklideszi téren értelmezett minden $\mathcal{Q} : V \to \mathcal{R}$ kvadratikus alakhoz létezik pontosan egy olyan $A \in L(V)$ szimmetrikus lineáris transzformáció, hogy minden $v \in V$-re*

$$\mathcal{Q}(v) = \langle v, A(v)\rangle\ .$$

Bizonyítás. Tegyük fel, hogy a $\mathcal{Q}$ kvadratikus alakhoz tartozó lineáris transzformáció $B \in L(V)$, de B nem szimmetrikus. Legyen $A = \frac{B+B^*}{2}$. Akkor A szimmetrikus lineáris transzformációja V-nek és

$$\mathcal{Q}(v) = \langle v, B(v)\rangle = \frac{1}{2}(\langle v, B(v)\rangle + \langle v, B(v)\rangle) =$$

$$\frac{1}{2}(\langle v, B(v)\rangle + \langle B^*(v), v\rangle) = \frac{1}{2}(\langle v, B(v)\rangle + \langle v, B^*(v)\rangle) = \langle v, A(v)\rangle\,.$$

Az unicitási állítás igazolása végett tegyük fel, hogy $\mathcal{X} = \{x_1, \ldots, x_n\}$ ortonormált bázisa V-nek. Ha A szimmetrikus, akkor minden $x_i, x_j \in \mathcal{X}$ vektorpárra

$$\langle x_i, A(x_j)\rangle = \langle A^*(x_i), x_j\rangle = \langle A(x_i), x_j\rangle = \langle x_j, A(x_i)\rangle\ ,$$

és így

$$\mathcal{Q}(x_i + x_j) = \langle x_i + x_j, A(x_i + x_j)\rangle = \langle x_i + x_j, A(x_i) + A(x_j)\rangle =$$

$$\langle x_i, A(x_i)\rangle\, 2\,\langle x_i, A(x_j)\rangle + \langle x_j, A(x_j)\rangle = \mathcal{Q}(x_i) + 2\,\langle x_i, A(x_j)\rangle + \mathcal{Q}(x_j)\,,$$

amiből

$$\langle x_i, A(x_j)\rangle = \frac{\mathcal{Q}(x_i + x_j) - \mathcal{Q}(x_i) - \mathcal{Q}(x_j)}{2}\,,$$

azaz az $A(x_j)$ vektor x_i-re vonatkozó koordinátája $\mathcal{Q}$ értékeinek ismeretében meghatározható. Akkor viszont, kihasználva, hogy minden $j(= 1, \ldots, n)$-re

$$A(x_j) = \sum_{i=1}^{n} \langle x_i, A(x_j)\rangle\, x_i\,,$$

az A lineáris transzformáció a bázisvektorokon, és akkor már az egész V téren is, meghatározott a $\mathcal{Q}$ kvadratikus alak által felvett értékek által. □

Az euklideszi tereken értelmezett kvadratikus alakokat osztályozhatjuk az általuk felvehető értékek szerint a következőképpen:

5.5.4 Definíció. *Azt mondjuk, hogy az V euklideszi téren értelmezett $\mathcal{Q} : V \to \mathcal{R}$ kvadratikus alak*

(1) pozitív definit*, ha $\mathcal{Q}(v) > 0$ minden nemzéró $v \in V$-re,*

(2) negatív definit*, ha $\mathcal{Q}(v) < 0$ minden nemzéró $v \in V$-re,*

(3) pozitiv szemidefinit*, ha minden $v \in V$-re $\mathcal{Q}(v) \geq 0$,*

(4) negatív szemidefinit*, ha minden $v \in V$-re $\mathcal{Q}(v) \leq 0$,*

(5) indefinit*, ha $\mathcal{Q}$ mind pozitív, mind negatív értéket felvesz.*

Ezeket az elnevezéseket használjuk euklideszi tereken értelmezett szimmetrikus lineáris transzformációkra is és valós szimmetrikus mátrixokra is. Például pozitív definitnek mondunk egy valós szimmetrikus mátrixot, ha az pozitív definit kvadratikus alak mátrixa valamely bázisban. Hasonló értelemben beszélünk pozitív szemidefinit, negatív definit, negatív szemidefinit és indefinit szimmetrikus mátrixokról is.

Egy általános alakú kvadratikus alakról nem könnyű eldönteni, hogy melyik definitségi osztályba tartozik. Ezt és a probléma lehetséges megoldását mutatja a következő egyszerű példa.

Tekintsük a $\mathcal{Q}(v) = \xi_1^2 - 6\xi_1\xi_2 + 2\xi_2^2$ kvadratikus alakot, ahol ξ_1, ξ_2 a v vektor koordinátái a 2-dimenziós euklideszi tér valamely $\mathcal{X} = \{x_1, x_2\}$ bázisában. Ahhoz, hogy eldönthessük, hogy $\mathcal{Q}$ melyik definitségi osztályba tartozik, célszerű átírni a $\mathcal{Q}(v) = (\xi_1 - 3\xi_2)^2 - 7\xi_2^2$ formába. Ha az $\mathcal{X}$ bázis helyett áttérünk az $\mathcal{Y} = \{y_1, y_2\}$ bázisra, ahol $y_1 = x_1$ és $y_2 = 3x_1 + x_2$, akkor a v vektor $\mathcal{Y}$-ra vonatkozó koordinátái a

$$\begin{array}{c|cc|c} & y_1 & y_2 & v \\ \hline x_1 & \boxed{1} & 3 & \xi_1 \\ x_2 & 0 & 1 & \xi_2 \\ \hline \end{array} \rightarrow \begin{array}{c|c|c} & y_2 & v \\ \hline y_1 & 3 & \xi_1 \\ x_2 & \boxed{1} & \xi_2 \\ \hline \end{array} \rightarrow \begin{array}{c|r} & v \\ \hline y_1 & \xi_1 - 3\xi_2 \\ y_2 & \xi_2 \\ \hline \end{array}$$

számolás alapján $\eta_1 = \xi_1 - 3\xi_2$ és $\eta_2 = \xi_2$ és $\mathcal{Q}(v) = \eta_1^2 - 7\eta_2^2$, amiből már nyilvánvaló, hogy $\mathcal{Q}$ indefinit. Az $\{x_1, x_2\}$ bázisban a kvadratikus alak mátrixa az

$$\mathbf{A} = \begin{bmatrix} 1 & -3 \\ -3 & 2 \end{bmatrix}, \text{ míg az új } \mathcal{Y} = \{y_1, y_2\} \text{ bázisban az } \mathbf{A}' = \begin{bmatrix} 1 & 0 \\ 0 & -7 \end{bmatrix}$$

diagonális mátrix. Vegyük észre, hogy $\mathcal{Q}(v)$ értéke az $\mathcal{Y} = \{y_1, y_2\}$ bázisban éppen v koordinátái négyzetének a diagonálisban lévő skalárokkal képzett lineáris kombinációja. $\square$

A példa tehát azt sugallja, hogy a kvadratikus alak definitségének megállapítása úgy történhet, hogy áttérünk olyan bázisra (koordináta rendszerre), amelyben a kvadratikus alak mátrixa diagonális, és megvizsgáljuk, hogy a diagonálisbeli skalárok mindegyike azonos előjelű-e, vagy vannak mind pozitív mind negatív előjelű elemek. Tetszőleges n-dimenziós V euklideszi tereken értelmezett kvadratikus alakok esetén is, ha a kvadratikus alak mátrixa valamely bázisban diagonális mátrix, akkor bármely $v \in V$ helyen értéke megkapható, ha vesszük v e bázisra vonatkozó koordinátái négyzetének a diagonálisbeli megfelelő skalárokkal vett lineáris kombinációját. Valóban, ha V-nek $\mathcal{X} = \{x_1, \ldots, x_n\}$ olyan bázisa, amelyre a kvadratikus alak mátrixa az

$$\mathbf{A} = \begin{bmatrix} \alpha_{11} & 0 & \dots & 0 \\ 0 & \alpha_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_{nn} \end{bmatrix}$$

diagonális mátrix, akkor az

$$\mathbf{x} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$$

koordinátavektorú x vektornál a kvadratikus alak értéke

$$\mathcal{Q}(x) = \mathbf{x}^*\mathbf{A}\mathbf{x} = \alpha_{11}\xi_1^2 + \cdots + \alpha_{nn}\xi_n^2 \,.$$

Euklideszi téren értelmezett kvadratikus alak ilyen négyzetösszegre redukált formájából azonnal leolvasható, hogy milyen értékeket vehet fel, hiszen valós szám négyzete nem lehet negatív, így a kvadratikus alak pozitív definit, ha a diagonális mátrixa diagonálisában minden elem pozitív, pozitív szemidefinit, ha ezek az elemek nemnegatívak, de esetleg van közöttük zérus is, a kvadratikus alak negatív definit, ha a diagonális mátrixa diagonálisában minden elem negatív, negatív szemidefinit, ha ezek az elemek nempozitívak, de esetleg van közöttük zérus is, és indefinit, ha a diagonálisban van pozitív és negatív elem egyaránt.

A kérdés mostmár csupán az, hogy minden kvadratikus alakhoz található-e olyan bázisa a térnek, amelyben a mátrixa diagonális mátrix. Meg fogjuk mutatni, hogy erre a kérdésre igen a válasz.

5.5.5 Tétel. *Bármely, a V euklideszi téren értelmezett*

$$\mathcal{Q} : V \to \mathcal{R}$$

kvadratikus alakhoz lehet találni a térnek olyan bázisát, amelyben a kvadratikus alak mátrixa diagonális mátrix.

Bizonyítás. A fentiekben beláttuk, hogy $\mathcal{Q}$-hoz tartozik a térnek pontosan egy $A \in L(V)$ szimmetrikus lineáris transzformációja, hogy minden $v \in V$-re

$$\mathcal{Q}(v) = \langle v, A(v) \rangle \,.$$

A 5.4.3 tétel szerint van A sajátvektoraiból álló $\mathcal{S} = \{s_1, \dots, s_n\}$ ortonormált bázisa a V euklideszi térnek, amelyben A mátrixa

$$\mathbf{A}_{\mathcal{S}} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

alakú diagonális mátrix, ahol λ_i az s_i sajátvektorhoz tartozó sajátérték. Már csak azt kell belátnunk, hogy a $\mathcal{Q}$ kvadratikus alak mátrixa minden ortonormált bázisban ugyanaz, mint az A lineáris transzformáció mátrixa. Legyen ezért most $\mathcal{X} = \{x_1, \dots, x_n\}$ egy tetszőleges ortonormált bázisa V-nek, és $v = \sum_{i=1}^n \epsilon_i x_i \in V$. Akkor

$$\mathcal{Q}(v) = \left\langle \sum_{i=1}^n \epsilon_i x_i, A(\sum_{i=1}^n \epsilon_i x_i) \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \epsilon_i \epsilon_j \langle x_i, A(x_j) \rangle \ ,$$

azaz $\mathcal{Q}$ mátrixa

$$\mathbf{A} = \begin{bmatrix} \langle x_1, A(x_1) \rangle & \langle x_1, A(x_2) \rangle & \dots & \langle x_1, A(x_n) \rangle \\ \langle x_2, A(x_1) \rangle & \langle x_2, A(x_2) \rangle & \dots & \langle x_2, A(x_n) \rangle \\ \vdots & \vdots & \dots & \vdots \\ \langle x_n, A(x_1) \rangle & \langle x_n, A(x_2) \rangle & \dots & \langle x_n, A(x_n) \rangle \end{bmatrix}$$

ugyanaz, mint A-nak az $\mathcal{X}$ bázisra vonatkozó mátrixa, mert ortonormált bázisban a skaláris szorzat belső szorzat, ennélfogva minden i, j $(i, j = 1, \dots n)$ indexpárra az $A(x_j)$ vektor x_i-re vonatkozó koordinátája éppen $\langle x_i, A(x_j) \rangle$. □

A fenti tétel szerint egy kvadratikus alak definitségét könnyen eldönthetjük a hozzá tartozó szimmetrikus lineáris transzformáció sajátértékeinek ismeretében. Ha a lineáris transzformáció minden sajátértéke pozitív, akkor a kvadratikus alak pozitív definit, ha mindegyik negatív, akkor negatív definit, ha mind pozitív, mind negatív sajátértékei vannak a transzformációnak, akkor a kvadratikus alak indefinit.

3 Példa. *Állapítsuk meg, hogy a $\mathcal{Q} : V \to \mathcal{R}$ kvadratikus alak melyik definitségi osztályba tartozik, ha minden $v \in V$-re*

$$\mathcal{Q}(v) = 2\xi_1^2 + 2\xi_1\xi_2 - 2\xi_1\xi_3 + 2\xi_2\xi_3 \,,$$

ahol ξ_1, ξ_2, ξ_3 a v vektor x_1, x_2, x_3 bázisvektorokra vonatkozó koordinátái!

A kvadratikus alak szimmetrikus mátrixa:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix},$$

tehát a kvadratikus alakhoz tartozó A lineáris transzformáció a bázisvektorokat az

$$A(x_1) = 2x_1 + x_2 - x_3,\ A(x_2) = x_1 + x_3,\ A(x_3) = -x_1 + x_2$$

vektorokba viszi. Az A sajátértékeit kellene meghatároznunk, azok előjeléből, már megállapíthatjuk $\mathcal{Q}$ definitségét. Az A minimálpolinomjának meghatározása érdekében kiszámítjuk mind az $\mathbf{A}^2$, mind az $\mathbf{A}^3$ mátrixot, ezek

$$\mathbf{A}^2 = \begin{bmatrix} 6 & 1 & -1 \\ 1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \qquad \text{és} \qquad \mathbf{A}^3 = \begin{bmatrix} 14 & 5 & -5 \\ 5 & 0 & 1 \\ -5 & 1 & 0 \end{bmatrix},$$

majd az A transzformációnak olyan $p_i(A)$ $(i = 1, 2, 3)$ polinomjait, amelyekre $p_i(A)(x_i) = 0$, ugyanis a $p_1(t), p_2(t), p_3(t)$ polinomok legkisebb közös többszöröse az A minimálpolinomja. Ha valamely $p_i(t)$ polinom harmadfokú, akkor a többit már nem is kell meghatároznunk, mert igazolható, hogy a minimálpolinom fokszáma nme lehet nagyobb a tér dimenziójánál. A $p_i(t)$ polinomok meghatározása történhet elemi bázistranszformációval:

$$\begin{array}{c|ccc} & A(x_1) & A^2(x_1) & A^3(x_1) \\ \hline x_1 & 2 & 6 & 14 \\ x_2 & \boxed{1} & 1 & 5 \\ x_3 & -1 & -1 & -5 \\ \hline \end{array} \rightarrow \begin{array}{r|cc} & A^2(x_1) & A^3(x_1) \\ \hline x_1 & 4 & 4 \\ A(x_1) & 1 & 5 \\ x_3 & 0 & 0 \\ \hline \end{array}$$

Az utóbbi táblázatból kiolvasható, hogy $A^2(x_1) = A(x_1)+4x_1$ azaz $p_1(t) = t^2-t-4$. Mivel $p_1(t)$ gyökei, $\frac{1\pm\sqrt{17}}{2}$, amelyek A minimálpolinomjának is gyökei, különböző előjelűek, már most megállapíthatjuk, hogy $\mathcal{Q}$ diagonális mátrixában a diagonálisbeli elemek között van pozitív és negatív skalár is, ezért $\mathcal{Q}$ indefinit. A $\mathcal{Q}$ diagonális mátrixának meghatározásához azonban meg kell keresni A harmadik sajátértékét is. Ezért meghatározzuk a $p_2(t)$ polinomot is. A

$$\begin{array}{c|ccc} & A(x_2) & A^2(x_2) & A^3(x_2) \\ \hline x_1 & \boxed{1} & 1 & 5 \\ x_2 & 0 & 2 & 0 \\ x_3 & 1 & -1 & 1 \\ \hline \end{array} \rightarrow \begin{array}{r|cc} & A^2(x_2) & A^3(x_2) \\ \hline A(x_2) & 1 & 5 \\ x_2 & 2 & 0 \\ x_3 & \boxed{-2} & -4 \\ \hline \end{array} \rightarrow$$

$$\begin{array}{r|c} & A^3(x_2) \\ \hline A(x_2) & 3 \\ x_2 & -4 \\ A^2(x_2) & 2 \\ \hline \end{array}$$

számítások szerint $A^3(x_2) = 2A^2(x_2)+3A(x_2)-4x_2$, azaz $p_2(t) = t^3-2t^2-3t+4$. Mivel $p_2(t) = (t-1)p_1(t)$, harmadfokú, az A minimálpolinomja meg kell egyezzen $p_2(t)$-vel és az A harmadik sajátértéke 1. Az A sajátvektorai alkotta ortonormált bázisban a $\mathcal{Q}$ mátrixa

$$\mathbf{A}' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1+\sqrt{17}}{2} & 0 \\ 0 & 0 & \frac{1-\sqrt{17}}{2} \end{bmatrix}.$$

6. Fejezet

Differenciálszámítás

Ez a fejezet az eddig tanult lineáris algebra tananyag alkalmazásaként megmutatja, hogy hogyan vihető át a derivált fogalma többváltozós függvényekre. Látni fogjuk, hogy a derivált tulajdonképpen az első félévben megismert érintő approximáció fogalmának természetes kiterjesztése a lineáris algebra eszközeivel. Tárgyalni fogjuk a derivált legfontosabb tulajdonságait, majd rátérünk a többváltozós függvények szélsőértékeinek meghatározására.

6.1 Mátrixok normája

Ebben a bevezető jellegű szakaszban a lineáris leképezések, illetve a mátrixok normájával, és azok legfontosabb tulajdonságaival ismerkedünk meg. Amint azt látni fogjuk, ezzel a normával ellátva a leképezések vektortere ugyanolyan normált teret alkot, amilyenre már számos példát láttunk az analízis tanulmányaink során. Igy lehetőségünk nyílik a mátrixok terének topológiai jellegű vizsgálatára, amelyre az alkalmazások (például Neumann-sorok) szempontjából is nagy szükségünk lesz.

Lényeges szempont a továbbiakban, hogy az euklideszi terek egy ortonormált bázisát rögzítettnek tekintjük, és nem teszünk különbséget egy lineáris leképezés, illetve annak az adott bázisban vett mátrixa között. Ugyanarra gondolunk tehát, ha akár leképezésről, akár mátrixról beszélünk. Ez elvi problémát sem okozhat, hiszen izomorf vektorterek azonosításáról van szó. Ha valamikor a bázis megváltoztatása kerülne szóba, akkor erre külön felhívjuk a figyelmet. Egyébként, hacsak mást nem mondunk, vektortéren mindig valós test feletti vektorteret értünk.

Legyenek tehát a továbbiakban X és Y euklideszi terek, és $\dim X = p$, illetve $\dim Y = q$. Továbbra is használjuk az $L(X,Y)$ jelölést az X téren értelmezett, Y térbe képező lineáris leképezések vektorterére. Ha történetesen $X = Y$, akkor a rövidebb $L(X)$ jelölésmóddal élünk.

Tekintsünk egy $A \in L(X,Y)$ lineáris leképezést.

6.1.1 Definíció. *Az A leképezés* normáján *az egységgömb felszínén felvett értékei abszolút értékeinek felső határát értjük, azaz*

$$\|A\| = \sup_{\|x\|=1} \|Ax\| \, .$$

Világos, hogy a definícióban supremum helyett maximum is írható, hiszen egy folytonos függvény egy kompakt halmazon Weierstrass tétele értelmében felveszi a legnagyobb értékét. (Lásd az 1. gyakorlatot.)

Első pillantásra nem világos, hogy a normát miért éppen így értelmezzük. Amint azt látni fogjuk, ez a definíció valóban normát definiál, de ezt nagyon sok más módon is meg lehetne tenni. Mondhatnánk például azt, hogy a normát definiáljuk a legnagyobb abszolút értékű oszlop abszolút értékével, azaz

$$\|A\| = \max_{1\le j\le p}\left(\sum_{i=1}^{q} a_{ij}^2\right)^{1/2} , \tag{6.1}$$

vagy éppen vehetnénk normának az elemek abszolút értékeinek maximumát is, tehát

$$\|A\| = \max_{1\le i\le q, 1\le j\le p} |a_{ij}| . \tag{6.2}$$

Nem nehéz belátni, hogy a (6.1) és (6.2) relációk tényleg kielégítik a norma axiómáit (lásd a 2. gyakorlatot). Az általunk bevezetett definíció mellett az szól, hogy, amint azt látni fogjuk, igen praktikus tulajdonságai vannak, valamint szimmetrikus mátrixokra nagyon szép algebrai jelentése is van. További érv az, hogy a fenti két reláció a normát a mátrix elemeinek segítségével értelmezi, így a norma függhet a bázis megválasztásától. A 6.1.1 Definíció azonban a leképezés normáját vezeti be, amely nem változik új bázisra történő áttéréskor, ha a skaláris szorzatot már rögzítettük. Térjünk tehát rá az általunk értelmezett norma tulajdonságainak összefoglalására.

6.1.2 Állítás. *Az $L(X,Y)$ vektortér a 6.1.1 Definícióban bevezetett normával normált teret alkot, azaz*

$$\|A\| \ge 0 , \quad \text{és} \quad \|A\| = 0 \quad \text{akkor és csak akkor, ha} \quad A = 0 ,$$

továbbá

$$\|A+B\| \le \|A\| + \|B\| , \tag{6.3}$$

illetve

$$\|\lambda A\| = |\lambda| \cdot \|A\|$$

bármely $A, B \in L(X,Y)$, és λ skalár mellett.

Bizonyítás. A (6.3) egyenlőtlenség abból adódik, hogy

$$\sup_{\|x\|=1} \|Ax + Bx\| \le \sup_{\|x\|=1} \|Ax\| + \sup_{\|x\|=1} \|Bx\| ,$$

míg a másik két reláció a definíció nyilvánvaló következménye. □

Megjegyezzük, hogy a (6.3) egyenlőtlenséget a szokásoknak megfelelően háromszögegyenlőtlenségnek nevezzük.

6.1.3 Állítás. *Minden $x \in X$ esetén*

$$\|Ax\| \le \|A\| \cdot \|x\| .$$

Bizonyítás. Az állítás triviális ha $x = 0$. Ha $x \neq 0$, akkor, minthogy $x/\|x\|$ egységnyi normájú vektor, a definíció alapján azt kapjuk, hogy

$$\|A\| \geq \left\| A\left(\frac{x}{\|x\|}\right)\right\| ,$$

ami az állításunkat igazolja. □

Az is könnyen belátható a definíció alapján, hogy $\|A\|$ éppen azzal a legkisebb λ nemnegatív számmal egyezik meg, amelyre minden x mellett érvényes az $\|Ax\| \leq \lambda\|x\|$ egyenlőtlenség (lásd a 3. gyakorlatot).

6.1.4 Állítás. *Ha A és B olyan leképezések, hogy a BA szorzat értelmes, akkor*

$$\|BA\| \leq \|B\| \cdot \|A\| .$$

Bizonyítás. Valóban, bármely x vektor mellett

$$\|(BA)x\| = \|B(Ax)\| \leq \|B\| \cdot \|Ax\| \leq \|B\| \cdot \|A\| \cdot \|x\|$$

az előző állításunk alapján. Innen azonnal adódik az állítás. □

A norma néhány praktikus tulajdonságának megismerése után térjünk rá annak vizsgálatára, hogy vajon milyen algebrai jelentést hordoz egy mátrix normája. Amint azt látni fogjuk, sok esetben könnyebb a norma meghatározása az algebrai jelentése, mint közvetlenül a definíció alapján.

6.1.5 Állítás. *Tekintsünk egy $q \times p$ méretű A mátrixot. Akkor $\|A\|$ megegyezik az A^*A mátrix legnagyobb sajátértékének négyzetgyökével.*

Bizonyítás. A lineáris algebrából jól ismert, hogy A^*A szimmetrikus pozitív szemidefinit mátrix, tehát a sajátértékei nemnegatív valós számok. Legyen most $v_1, \ldots, v_p$ az X vektortérnek egy olyan ortonormált bázisa, amelyben A^*A diagonális alakú. A $v_1, \ldots, v_p$ vektorok az A^*A mátrix sajátvektorai, a megfelelő sajátértékeket jelölje $\lambda_1, \ldots, \lambda_p$. A sajátértékek között azonosak is előfordulhatnak, mindegyiket annyiszor írtuk ki, amennyi a multiplicitása. Tegyük fel, hogy a sajátértékek között a legnagyobb éppen λ_k. Azt kell igazolnunk, hogy $\|A\| = \sqrt{\lambda_k}$.

Tekintsünk egy tetszőleges $x \in X$ vektort, amely egységnyi normájú, és amelyet a $v_1, \ldots, v_p$ bázisban az

$$x = \sum_{i=1}^{p} x_i v_i$$

lineáris kombináció állít elő. Ekkor

$$\begin{aligned} \|Ax\|^2 &= \langle Ax, Ax\rangle = \langle x, A^*Ax\rangle = \langle \sum_{i=1}^{p} x_i v_i, \sum_{i=1}^{p} \lambda_i x_i v_i\rangle \\ &= \sum_{i=1}^{p} \lambda_i x_i^2 \leq \sum_{i=1}^{p} \lambda_k x_i^2 = \lambda_k , \end{aligned}$$

hiszen $\|x\| = 1$. Ez azt jelenti, hogy $\|Ax\| \leq \sqrt{\lambda_k}$ minden egységnyi normájú x vektor esetén, azaz $\|A\| \leq \sqrt{\lambda_k}$. Másrészt ha a fenti levezetésben az x vektornak éppen a v_k bázisvektort választjuk, akkor azt kapjuk, hogy

$$\|Av_k\|^2 = \langle v_k, A^*Av_k\rangle = \langle v_k, \lambda_k v_k\rangle = \lambda_k ,$$

azaz $\|A\| \geq \sqrt{\lambda_k}$, ami az állításunkat bizonyítja. □

6.1.6 Következmény. *Tegyük fel, hogy A szimmetrikus mátrix. Jelölje $\lambda_{\min}$, illetve $\lambda_{\max}$ az A legkisebb, illetve legnagyobb sajátértékét. Ekkor*

$$\lambda_{\min}\|v\|^2 \leq \langle v, Av\rangle \leq \lambda_{\max}\|v\|^2$$

minden $v \in X$ esetén. Nevezetesen $\|A\|$ megegyezik a $|\lambda_{\max}|$ és $|\lambda_{\min}|$ közül a nagyobbikkal.

Bizonyítás. Tekintsük az X egy olyan ortonormált bázisát, amely az A sajátvektoraiból áll. Ebben a bázisban a fenti kvadratikus alak négyzetösszegként áll elő, azaz ha a v vektor koordinátái ebben a bázisban $v_1, \dots, v_p$, akkor

$$\langle v, Av\rangle = \sum_{i=1}^{p} \lambda_i v_i^2 \ ,$$

ahol a λ_i együtthatók az A megfelelő sajátértékei. Innen azonnal adódik a fenti egyenlőtlenség, hiszen $\sum_{i=1}^{p} v_i^2 = \|v\|^2$.

Az $\|A\|$ előállítása a 6.1.5 Állítás közvetlen következménye, hiszen ekkor $A^*A = A^2$, és A^2 sajátértékei éppen az A sajátértékeinek négyzetei. □

A későbbiekben egy feltételes szélsőértéken alapuló módszerrel is találkozunk majd a norma meghatározására.

6.1.7 Példa. A norma segítségével megfogalmazhatjuk a geometriai sorok összegképletének mátrixokra érvényes általánosítását is. Megmutatjuk, hogy ha $\|A\| < 1$, akkor $I - A$ invertálható, ahol I az egységmátrix, továbbá

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k \ . \tag{6.4}$$

Itt a végtelen sor konvergenciája normában értendő, azaz azt mondjuk, hogy a $\sum_{k=0}^{\infty} A^k$ sor konvergens, és az összege az S mátrix, ha az $S_n = \sum_{k=0}^{n} A^k$ részletösszegekre igaz, hogy

$$\lim_{n\to\infty} \|S_n - S\| = 0 \ .$$

Az abszolút konvergens sorokról szóló tételhez teljesen hasonlóan megmutatható, hogy a $\sum_{k=0}^{\infty} A^k$ sor konvergenciájának elégséges (de nem szükséges) feltétele a $\sum_{k=0}^{\infty} \|A^k\|$ numerikus sor konvergenciája. Ez utóbbi azonban esetünkben nyilvánvaló, hiszen nemnegatív tagú sorról van szó, és a részletöszszegek az $\|A^k\| \leq \|A\|^k$ egyenlőtlenség alapján (lásd a 6.1.4 Állítást) felülről becsülhetők a $\sum_{k=0}^{\infty} \|A\|^k$ konvergens geometriai sor részletösszegeivel. Tehát az összehasonlító kritérium szerint a (6.4) alatti végtelen sor konvergens.

Megmutatjuk, hogy a fenti sor összege éppen az $I - A$ mátrix inverze. Az eddigi jelöléseinket használva

$$S_n(I - A) = I - A^{n+1} \to I \ ,$$

hiszen $A^{n+1} \to 0$ a feltételünk szerint. Másrészt nyilvánvalóan $S_n(I-A) \to S(I-A)$, hiszen

$$\|S_n(I - A) - S(I - A)\| \leq \|I - A\| \cdot \|S_n - S\| \to 0 \ .$$

Ez azt jelenti, hogy $S(I-A) = I$, azaz $S = (I-A)^{-1}$, amit igazolnunk kellett.

Megjegyezzük, hogy a (6.4) formula fontos szerepet játszik az elméleti közgazdaságtanban. Az irodalomban a (6.4) alatti végtelen sort Neumann-sornak is nevezik . A közgazdasági input-output modellekben, ha A jelöli a fajlagos ráfordítási mátrixot, akkor az $(I-A)^{-1}$ mátrixot az A Leontief-inverzének nevezik. Fontos kérdés ezekben a modellekben, hogy melyek azok a mátrixok, amelyeknek létezik csupa nemnegatív elemű Leontief-inverze. Az ilyen mátrixokat produktívnak nevezik. Amint azt a (6.4) formulából azonnal láthatjuk, a nemnegatív elemű A fajlagos ráfordítási mátrix produktív, ha teljesül rá az $\|A\| < 1$ feltétel.

6.1.8 Példa. Az elméleti közgazdaságtan irodalmában gyakran előfordul a domináns sajátérték fogalma, amely egy mátrix abszolút értékben legnagyobb sajátértékét jelenti. Későbbi tanulmányainkban látni fogjuk, hogy egy nemnegatív elemű mátrix produktivitásának szüséges és elégséges feltétele az, hogy a domináns sajátértéke kisebb, mint 1 (ez a nevezetes Perron-Frobenius-tétel). Nem árt felhívni a figyelmet arra, hogy ez a fogalom általában nem egyezik meg a mátrix normájával. Ha λ jelöli az A mátrix domináns sajátértékét, akkor mindenesetre

$$|\lambda| \leq \|A\| ,$$

ám egyenlőség pontosan akkor érvényes, ha A alkalmas bázisban diagonális alakra hozható (lásd a 7. gyakorlatot). Tekintsük például a nem diagonalizálható

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

mátrixot. Ekkor az A mátrixnak $\lambda = 1$ kétszeres sajátértéke, de a 6.1.5 Állítás alapján könnyen ellenőrízhető, hogy $\|A\| = (1+\sqrt{5})/2$.

6.2 Differenciálhatóság

Ebben a szakaszban bevezetjük a többváltozós függvények deriváltjának fogalmát. Vegyük észre, hogy a fogalom természetes általánosítása az első éves analízisben megismert *érintő approximáció* fogalmának, és tulajdonképpen semmi újat nem tartalmaz. Pusztán a lineáris leképezés fogalmát használjuk az egydimenziós esetnél általánosabb értelemben.

6.2.1 Definíció. *Legyenek X és Y euklideszi terek, és tekintsük az $f : X \to Y$ leképezést, amely értelmezve van az $x \in X$ pont egy környezetében. Azt mondjuk, hogy f* differenciálható *az x pontban, ha található olyan $A \in L(X,Y)$ lineáris leképezés, hogy bármely $v \in X$, $x+v \in D_f$ esetén*

$$f(x+v) = f(x) + Av + r(v) ,$$

ahol $\lim_{v\to 0} \|r(v)\|/\|v\| = 0$. *Ebben az esetben az A leképezést az f deriváltjának nevezzük az x pontban. Jelölése $A = f'(x)$.*

Megjegyezzük, hogy az érintő approximáció fogalmához hasonlóan a fenti definíció azt fogalmazza meg, hogy az x pont egy környezetében az f függvény jól, azaz kis ordó nagyságrendben közelíthető az A lineáris leképezéssel. Világos ugyanis a definícióból, hogy az $r : X \to Y$ függvény kis ordó nagyságrendű az x környezetében.

Nem látszik a definícióból, hogy a derivált egyértelműen meghatározott, azaz csak egyetlen olyan A lineáris leképezés létezhet, amely kielégíti a fenti definíciót. Erre ad választ az alábbi állítás.

6.2.2 Állítás. *A derivált egyértelműen meghatározott.*

Bizonyítás. Tegyük fel, hogy az A és B lineáris leképezések egyaránt eleget tesznek a definíció követelményeinek, azaz, ha $x + v \in D_f$, úgy

$$\begin{aligned} f(x+v) &= f(x) + Av + r(v) \\ f(x+v) &= f(x) + Bv + q(v)\ , \end{aligned}$$

ahol r és q kis ordó függvények. Ekkor a $C = A - B$ jelöléssel a $Cv = r(v) - q(v) = o(v)$ egyenlőséghez jutunk, amely ugyancsak kis ordó függvény. Tehát tetszőleges $v \neq 0$ vektor mellet

$$\frac{\|Cv\|}{\|v\|} = \frac{\|C(\frac{1}{n}v)\|}{\|\frac{1}{n}v\|} = \frac{\|o(\frac{1}{n}v)\|}{\|\frac{1}{n}v\|} \to 0\ ,$$

ha $n \to \infty$. Ez azt jelenti, hogy $Cv = 0$, azaz $C = A - B = 0$. □

Nyilvánvaló, hogy ha f differenciálható az x pontban, akkor ott folytonos is. (Lásd a 8. gyakorlatot.) Azt is belátjuk, hogy egy lineáris leképezés mindenütt differenciálható, és a deriváltja saját maga.

6.2.3 Állítás. *Ha f lineáris, akkor minden $x \in X$ pontban differenciálható, és $f'(x) = f$.*

Bizonyítás. Valóban, alkalmazzuk a definíciót az $A = f$, $r = 0$ szereposztás mellett. □

6.2.4 Példa. Tekintsük az $f : X \to \mathbb{R}$, $f(x) = \langle x, Bx\rangle$ kvadratikus alakot, ahol $B \in L(X)$ szimmetrikus transzformáció. Megmutatjuk, hogy f minden $x \in X$ pontban differenciálható, éspedig $f'(x) = 2Bx$.

Valóban, bármely $v \in X$ vektor mellett

$$\begin{aligned} f(x+v) - f(x) &= \langle x+v, B(x+v)\rangle - \langle x, Bx\rangle \\ &= \langle v, Bx\rangle + \langle x, Bv\rangle + \langle v, Bv\rangle \\ &= \langle v, 2Bx\rangle + \langle v, Bv\rangle\ , \end{aligned}$$

hiszen B szimmetrikus. Állításunk igazolásához tehát elég megmutatni, hogy $\langle v, Bv\rangle$ kis ordó nagyságrendű. Ez azonban egyszerűen látható a

$$|\langle v, Bv\rangle| \leq \|B\| \cdot \|v\|^2$$

egyenlőtlenségből.

Megjegyezzük, hogy ebben a példában $2Bx$ azt az $L(X, \mathbb{R}) = X^*$ térbeli lineáris függvényt jelenti, amelynek mátrixa az a sorvektor, amelynek elemei éppen a $2Bx$ koordinátái. Nevezetesen

$$2Bx(v) = \langle v, 2Bx\rangle$$

bármely $v \in X$ esetén.

Az alábbiakban összefoglaljuk a derivált legfontosabb tulajdonságait. A következő állítás egyszerűen adódik a definícióból.

6.2.5 Állítás. *Tegyük fel, hogy az f és g függvények egyaránt differenciálhatók az $x \in X$ pontban, és legyen $\lambda \in \mathbb{R}$ tetszőleges. Akkor $f + g$, illetve λf is differenciálhatók az x pontban, és*

$$\begin{aligned}(f+g)'(x) &= f'(x) + g'(x)\\ (\lambda f)'(x) &= \lambda f'(x)\end{aligned}$$

Az alábbi tétel az összetett függvény deriválási szabályát általánosítja euklideszi terekre. Vegyük észre azonban, hogy e tétel bizonyítása szinte szó szerint megegyezik az analízisben tanulttal.

Legyenek tehát X, Y és Z euklideszi terek, és tekintsük az $f : X \to Y$, valamint a $g : Y \to Z$ függvényeket. Tegyük fel, hogy x belső pontja az f értelmezési tartományának, és $f(x)$ is belső pontja a g értelmezési tartományának.

6.2.6 Tétel. *Ha f differenciálható az x pontban, továbbá g differenciálható az $f(x)$ pontban, akkor $g \circ f$ is differenciálható az x pontban, éspedig*

$$(g \circ f)'(x) = g'(f(x))f'(x) \; .$$

Bizonyítás. A feltételeink azt jelentik, hogy

$$f(x+v) = f(x) + f'(x)v + r(v) \; ,$$

illetve

$$g(f(x)+u) = g(f(x)) + g'(f(x))u + q(u) \; ,$$

ahol r és q egyaránt kis ordó nagyságrendűek. Ha most $v \in X$ tetszőleges, akkor az $u = f(x+v) - f(x)$ jelöléssel

$$\begin{aligned}g(f(x+v)) - g(f(x)) &= g'(f(x))u + q(u)\\ &= g'(f(x))(f(x+v) - f(x)) + q(u)\\ &= g'(f(x))(f'(x)v + r(v)) + q(u)\\ &= g'(f(x))f'(x)v + g'(f(x))r(v) + q(u) \; .\end{aligned}$$

Azt kell igazolni, hogy $g'(f(x))r(v) + q(u)$ kis ordó nagyságrendű v szerint. Ezt tagonként mutatjuk meg. Az első tagra ez a megállapítás nyilvánvaló, hiszen

$$\lim_{v\to 0} \frac{\|g'(f(x))r(v)\|}{\|v\|} \le \|g'(f(x))\| \lim_{v\to 0} \frac{\|r(v)\|}{\|v\|} = 0 \; .$$

A második tag kis ordó nagyságrendű u szerint. Ez azonban v szerint is igaz, ugyanis

$$\frac{\|q(u)\|}{\|v\|} = \begin{cases} 0 \; , & \text{ha } f(x+v) - f(x) = 0 \\ \frac{\|q(u)\|}{\|u\|}\frac{\|f(x+v)-f(x)\|}{\|v\|} \; , & \text{ha } f(x+v) - f(x) \neq 0 \end{cases} \; .$$

Mivel az f folytonossága miatt $v \to 0$ esetén $u \to 0$ is fennáll, azért

$$\lim_{v\to 0} \frac{\|q(u)\|}{\|v\|} = 0 ,$$

hiszen az

$$\frac{\|f(x+v) - f(x)\|}{\|v\|} = \frac{\|f'(x)v + r(v)\|}{\|v\|} \leq \|f'(x)\| + \frac{\|r(v)\|}{\|v\|}$$

tört korlátos. □

Ha esetünkben $\dim X = p$, $\dim Y = q$ és $\dim Z = r$, akkor $g'(f(x))$ $r \times q$, illetve $f'(x)$ $q \times p$ méretű mátrixok, és ennek megfelelően a $(g \circ f)'(x)$ szorzatmátrix $r \times p$ méretű.

6.2.7 Tétel. *Legyen $f : X \to \mathbb{R}$ differenciálható az x pontban, és tegyük fel, hogy az x pontban az f függvénynek lokális szélsőértéke van. Akkor $f'(x) = 0$.*

Bizonyítás. A feltételünk mellett tetszőleges $v \in X$ esetén a $g : \mathbb{R} \to \mathbb{R}$,

$$g(t) = f(x + tv)$$

függvénynek a 0 pontban lokális szélsőértéke van. Másrészt a 6.2.6 Tétel szerint g differenciálható a 0 pontban, és

$$0 = g'(0) = f'(x)v .$$

Ez éppen azt jelenti, hogy $f'(x) = 0$. □

A függvény értelmezési tartományának azon pontjait, ahol a függvény differenciálható, és a derivált zérus, *kritikus pontoknak* nevezzük.

6.2.8 Példa. Legyen $B \in L(X)$ szimmetrikus transzformáció, és tekintsük a $Q(x) = \langle x, Bx\rangle$ kvadratikus alakot. Keressük meg a Q szélsőértékeit. A 6.2.4 Példa szerint Q differenciálható, és $Q'(x) = 2Bx$. A 6.2.7 Tétel alapján a kritikus pontok a

$$Q'(x) = 2Bx = 0$$

homogén lineáris egyenletrendszer megoldásai, azaz a $\ker B$ altér elemei. Világos, hogy ezek a kritikus pontok minimumhelyek, ha B pozitív szemidefinit, maximumhelyek, ha B negatív szemidefinit, illetve egyikük sem szélsőértékhely, ha B indefinit.

6.3 Parciális deriváltak

Természetes kérdés a derivált fogalmának bevezetése után, hogy vajon hogyan határozható meg a derivált mátrixa. Ezt a kérdést vizsgáljuk meg ebben a szakaszban.

Legyenek tehát a továbbiakban X és Y olyan euklideszi terek, amelyekre $\dim X = p$, és $\dim Y = q$. Tekintsünk egy olyan $f : X \to Y$ függvényt, amely differenciálható az $x \in X$ pontban. Ekkor a definíció szerint $f'(x) \in L(X, Y)$, azaz a mátrixa $q \times p$ méretű. Mivel

$$f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_q \end{bmatrix},$$

ahol az f_i függvények az f koordinátafüggvényei, azért az $f'(x)$ mátrix sorait az egyes koordinátafüggvények deriváltjai alkotják, azaz

$$f'(x) = \begin{bmatrix} f_1'(x) \\ f_2'(x) \\ \vdots \\ f_q'(x) \end{bmatrix}.$$

Elegendő tehát megvizsgálni, hogy hogyan állítható elő egyetlen koordinátafüggvény deriváltjának a mátrixa. Ezért feltehető, hogy $Y = \mathbb{R}$.

Megjegyezzük, hogy az f differenciálhatóságából következik a koordinátafüggvények differenciálhatósága, és fordítva, ha az f minden koordinátafüggvénye differenciálható, akkor f is differenciálható (lásd a 9. gyakorlatot).

Tekintsünk tehát egy $f : X \to \mathbb{R}$ függvényt, és jelölje $e_1, \ldots, e_p$ az X rögzített ortonormált bázisát.

6.3.1 Definíció. *Legyen $x \in X$ az f értelmezési tartományának belső pontja. Azt mondjuk, hogy f parciálisan differenciálható az i-ik változó szerint az x pontban, ha létezik a*

$$\lim_{t \to 0} \frac{1}{t} \left(f(x + te_i) - f(x) \right) = D_i f(x)$$

határérték, és ez véges. A $D_i f(x)$ határértéket az f parciális deriváltjának nevezzük az x pontban.

Ha bevezetjük a $g(t) = f(x+te_i)$ függvényt a számegyenesen, akkor az f parciális differenciálhatósága az i-ik változó szerint az x pontban azt jelenti, hogy g differenciálható a 0 pontban, és $g'(0) = D_i f(x)$. Ennek az a szemléletes tartalma, hogy az f függvényt csak az i-ik változójában vizsgáljuk, a többi változót rögzített konstansnak tekintjük az x pontban.

6.3.2 Példa. A definíció figyelmes átolvasásával láthatjuk, hogy először a behelyettesítést végezzük el, csak utána a formális deriválást. Tekintsük például az $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x, y) = e^{x-y+2} \sqrt{3 + x^2 + y^2} (2x - 3y - 6)^5 \sin^2(\pi + x) \cos^2(\pi + y)$$

függvényt, és határozzuk meg az y szerinti parciális deriváltját az origóban. Minden számolás nélkül azonnal látható, hogy $D_2 f(0, 0) = 0$, ugyanis az $x = 0$ tengely mentén az f függvény azonosan nulla.

6.3.3 Állítás. *Ha f differenciálható az x pontban, akkor f minden változója szerint parciálisan differenciálható az x pontban, éspedig*

$$D_i f(x) = f'(x) e_i\,.$$

Bizonyítás. Valóban, a differenciálhatóság miatt

$$\frac{1}{t}(f(x+te_i)-f(x)) = \frac{1}{t}(f'(x)(te_i)+r(te_i)) = f'(x)e_i + \frac{r(te_i)}{t} ,$$

amiből $t \to 0$ mellett azonnal adódik az állítás. □

6.3.4 Példa. Megjegyzendő, hogy a fenti állítás nem fordítható meg. Nevezetesen nem nehéz példát mutatni olyan függvényre, amely valamely pontban parciálisan differenciálható az összes változója szerint, de a függvény még csak nem is folytonos abban a pontban. Tekintsük például a síkon az

$$f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2} , & \text{ha } x^2+y^2 \neq 0 \\ 0 , & \text{különben} \end{cases}$$

függvényt. Könnyen látható, hogy $D_1 f(0,0) = D_2 f(0,0) = 0$, azonban f nem folytonos az origóban. Valóban, f a koordinátatengelyek mentén zérus, míg a 45°-os egyenes mentén 1, így f az origó bármely környezetében egyaránt felveszi a 0 és az 1 értékeket is.

A parciális deriváltak ismerete már lehetővé teszi a derivált mátrixának felépítését. Amint láthatjuk, ha $f : X \to \mathbb{R}$ az x pontban differenciálható függvény, akkor $f'(x)$ olyan $1 \times p$ méretű mátrix, amelynek i-ik eleme éppen $D_i f(x)$. Ezen észrevétel alapján a keresett mátrix már könnyen megadható.

6.3.5 Következmény. *Tegyük fel, hogy az $f : X \to Y$ függvény differenciálható az x pontban. Akkor az eddigi jelöléseinket megtartva*

$$f'(x) = \begin{bmatrix} D_1 f_1(x) & D_2 f_1(x) & \dots & D_p f_1(x) \\ D_1 f_2(x) & D_2 f_2(x) & \dots & D_p f_2(x) \\ \vdots & & & \vdots \\ D_1 f_q(x) & D_2 f_q(x) & \dots & D_p f_q(x) \end{bmatrix}$$

Megjegyezzük, hogy a derivált fentebb megadott mátrixát néha az f Jacobi-mátrixának nevezik az x pontban. Amikor f valós számértékű függvény, tehát a Jacobi-mátrixa csak egyetlen sort tartalmaz, akkor a Jacobi-mátrix helyett elterjedt a gradiens vektor elnevezés is. Mi azonban a továbbiakban is kizárólag a derivált elnevezést használjuk.

6.3.6 Példa. Tekintsük például azt az $f : \mathbb{R}^2 \to \mathbb{R}^2$ függvényt, amely a sík pontjainak poláris koordinátáit derékszögű koordinátákra váltja, azaz

$$f(r,\theta) = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} ,$$

és legyen $g : \mathbb{R}^2 \to \mathbb{R}^3$ a következő:

$$g(x,y) = \begin{bmatrix} x^2 - xy \\ y^2 - xy \\ 2xy \end{bmatrix} .$$

Határozzuk meg a $(g \circ f)'(1, \pi/3)$ mátrixot.

A 6.3 Következmény szerint

$$f'(r,\theta) = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix},$$

továbbá

$$g'(x,y) = \begin{bmatrix} 2x-y & -x \\ -y & 2y-x \\ 2y & 2x \end{bmatrix}.$$

Igy a 6.2.6 Tétel alapján

$$\begin{aligned} (g \circ f)'(1,\pi/3) &= \begin{bmatrix} 1-\sqrt{3}/2 & -1/2 \\ -\sqrt{3}/2 & \sqrt{3}-1/2 \\ \sqrt{3} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 1/2-\sqrt{3}/2 & 1/2-\sqrt{3}/2 \\ 3/2-\sqrt{3}/2 & 1/2+\sqrt{3}/2 \\ \sqrt{3} & 1 \end{bmatrix}, \end{aligned}$$

amely természetesen 3×2 méretű.

6.3.7 Példa. Az összetett függvény deriválási szabályának gyakorta használt speciális esete az, amikor $f : \mathbb{R} \to \mathbb{R}^p$ és $g : \mathcal{R}^p \to \mathbb{R}$ differenciálható függvények. Ekkor

$$(g \circ f)'(t) = \sum_{i=1}^{p} D_i g(f(t)) f_i'(t)\,,$$

ahol $t \in \mathbb{R}$, és az f_i függvények az f koordinátafüggvényei.

6.3.8 Példa. A 6.2.7 Tétel szerint a szélsőértéknek nyilván szükséges feltétele a parciális deriváltak eltűnése. Keressük meg például az $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = 5x^2 + xy^2 - y^4$$

formulával definiált függvény szélsőértékeit. A kritikus pontokat a

$$\begin{aligned} D_1 f(x,y) &= 10x + y^2 = 0 \\ D_2 f(x,y) &= 2xy - 4y^3 = 0 \end{aligned}$$

egyenletrendszer megoldásával nyerjük. Ennek egyetlen megoldása az origó, amely azonban nyilván nem lokális szélsőérték, hiszen az f függvény az origó bármely környezetében egyaránt felvesz pozitív és negatív értékeket is. Ezért az f függvénynek nincs szélsőértéke.

6.4 Folytonos differenciálhatóság

Az előző szakaszban már láttunk példát arra, hogy a parciális deriváltak létezése nem feltétlenül jelenti a függvény differenciálhatóságát. Most azt fogjuk megvizsgálni, hogy milyen pótlólagos feltételek mellett igazolható a differenciálhatóság.

Mindenekelőtt megjegyezzük, hogy ha $f : X \to \mathbb{R}$ differenciálható valamely x pontban, akkor $f'(x)$ a definíció szerint az X^* duális tér egy eleme. Azonban X^* és X természetes módon izomorfak, ezt az izomorfizmust az X bázisa, illetve az X^* duális bázisa közötti bijekció adja meg. Ezért az $f' : X \to X^*$ leképezés úgy is tekinthető, mint egy $f' : X \to X$ leképezés. Formálisan nézve itt arról van szó, hogy az $f'(x)$ sorvektorokat oszlopvektorok gyanánt kezeljük.

6.4.1 Definíció. *Legyen M az X euklideszi tér valamely nyílt részhalmaza, és tegyük fel, hogy az $f : X \to \mathbb{R}$ függvény differenciálható az M halmaz minden pontjában. Azt mondjuk, hogy f folytonosan differenciálható az $x \in M$ pontban, ha $f' : X \to X$ folytonos az x pontban.*

6.4.2 Tétel. *Az f függvény akkor és csak akkor folytonosan differenciálható az $x \in M$ pontban, ha itt a parciális deriváltjai léteznek és folytonosak.*

Bizonyítás. Először a szükségességet igazoljuk. Tekintsük az $x \in M$ pontot, és legyen $\epsilon > 0$. Ekkor az f' folytonossága miatt létezik olyan $\delta > 0$, hogy $x, y \in M$, $\|x - y\| < \delta$ esetén $\|f'(x) - f'(y)\| < \epsilon$. Ekkor a 6.3.3 Állítás folytán a parciális deriváltak léteznek, és

$$\begin{aligned} |D_i f(x) - D_i f(y)| &= \|(f'(x) - f'(y))e_i\| \\ &\leq \|f'(x) - f'(y)\| < \epsilon \end{aligned}$$

bármely $i = 1, \ldots, p$ mellett. Ez éppen a parciális deriváltak folytonosságát jelenti.

Térjünk rá az elegendőség bizonyítására. Legyen adott $x \in M$ és $\epsilon > 0$. Ekkor a parciális deriváltak folytonossága alapján van olyan $\delta > 0$, hogy minden $x, y \in M$, $\|x - y\| < \delta$ esetén

$$|D_i f(x) - D_i f(y)| < \epsilon/p$$

bármely $i = 1, \ldots, p$ mellett. (Ez nyilván megtehető úgy, hogy minden i esetén választunk egy ilyen δ számot, majd az így kapott p darab δ közül kiválasztjuk a legkisebbet.)

Válasszunk ezután egy olyan $v \in X$ vektort, amelyre $\|v\| < \delta$. Ha v koordinátái rendre a $v_1, \ldots, v_p$ valós számok, úgy vezessük be a

$$v^i = \begin{bmatrix} v_1 \\ \vdots \\ v_i \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

és $v^0 = 0$ jelöléseket ($i = 1, \ldots, p$). Ekkor a Lagrange-féle középérték-tétel szerint vannak olyan $0 < t_i < 1$ számok, hogy

$$\begin{aligned} f(x + v) - f(x) &= \sum_{i=1}^{p} \Big(f(x + v^i) - f(x + v^{i-1}) \Big) = \\ &= \sum_{i=1}^{p} D_i f(x + v^{i-1} + t_i v_i e_i) v_i = \end{aligned}$$

$$\sum_{i=1}^{p} D_i f(x) v_i \quad + \quad \sum_{i=1}^{p} \Big(D_i f(x + v^{i-1} + t_i v_i e_i) - D_i f(x) \Big) v_i .$$

Itt a második szumma kis ordó nagyságrendű $v \to 0$ esetén, hiszen

$$\frac{1}{\|v\|} \left| \sum_{i=1}^{p} \Big(D_i f(x + v^{i-1} + t_i v_i e_i) - D_i f(x) \Big) v_i \right| \le \sum_{i=1}^{p} \frac{\epsilon}{p} \frac{|v_i|}{\|v\|} < \epsilon \, .$$

Ez éppen azt jelenti, hogy f differenciálható az $x \in X$ pontban. Mivel $f'(x) = [D_1 f(x), \ldots, D_p f(x)]$, azért az összetett függvény folytonossága szerint f' folytonos is az x pontban. □

6.5 Másodrendű deriváltak

Már láttuk, hogy az egyváltozós esethez hasonlóan a derivált zérus volta a szélsőértéknek csak szükséges feltétele. Ebben a szakaszban bevezetjük a második derivált fogalmát, amelyre szükségünk lesz az elégséges feltételek megfogalmazásához.

Tekintsünk egy X euklideszi teret és egy $f : X \to \mathbb{R}$ differenciálható függvényt. Amint azt már említettük, ilyenkor a derivált függvény olyan $f' : X \to X$ függvénynek is tekinthető, amelynek koordinátafüggvényei a $D_i f$ parciális deriváltak.

6.5.1 Definíció. *Azt mondjuk, hogy f kétszer differenciálható az $x \in X$ pontban, ha $f' : X \to X$ differenciálható az x pontban.*

Világos, hogy ha f kétszer differenciálható az x pontban, akkor $f''(x) \in L(X)$ az X euklideszi tér egy lineáris transzformációja. Ez azt jelenti, hogy a mátrixa egy $p \times p$ méretű négyzetes mátrix. Mivel

$$f' = \begin{bmatrix} D_1 f \\ \vdots \\ D_p f \end{bmatrix} ,$$

azért $f''(x)$ mátrixa felírható a parciális deriváltfüggvények parciális deriváltjaival, azaz a másodrendű parciális deriváltak segítségével.

6.5.2 Következmény. *Ha $f : X \to \mathbb{R}$ kétszer differenciálható az x pontban, akkor itt léteznek a másodrendű parciális deriváltjai, és*

$$f''(x) = \begin{bmatrix} D_{11} f(x) & D_{12} f(x) & \ldots & D_{1p} f(x) \\ D_{21} f(x) & D_{22} f(x) & \ldots & D_{2p} f(x) \\ \vdots & & & \vdots \\ D_{p1} f(x) & D_{p2} f(x) & \ldots & D_{pp} f(x) \end{bmatrix} .$$

ahol $D_{ij} f(x) = D_j(D_i f)(x)$.

Érdemes megjegyezni, hogy a fenti mátrixot az irodalomban néha az f függvény Hesse-mátrixának nevezik az x pontban. Mi azonban továbbra is a második derivált elnevezést használjuk.

6.5.3 Példa. Legyen $f : X \to \mathbb{R}$ kétszer differenciálható az $x \in X$ pont egy környezetében, és legyen $v \in X$ adott. Tekintsük a $g(t) = f(x + tv)$ függvényt a számegyenesen. Ekkor az összetett függvény deriválási szabálya alapján g is kétszer differenciálható a 0 pont egy környezetében, és itt

$$g''(t) = \langle v, f''(x + tv)v \rangle ,$$

ahol $t \in \mathbb{R}$.

6.5.4 Példa. Legyen például $f : \mathbb{R}^3 \to \mathbb{R}$ az

$$f(x, y, z) = 2x^2 y + xyz - y^2 z^2$$

formulával értelmezett függvény. A 6.5 Következmény szerint a második deriváltat az

$$f''(x, y, z) = \begin{bmatrix} 4y & 4x + z & y \\ 4x + z & -2z^2 & x - 4yz \\ y & x - 4yz & -2y^2 \end{bmatrix}$$

mátrix adja meg.

A fenti példában az $f''(x)$ mátrix szimmetrikus. Megmutatjuk, hogy ez általában is érvényes.

6.5.5 Tétel. **(Young tétele)** *Tegyük fel, hogy $f : X \to \mathbb{R}$ kétszer folytonosan differenciálható az $x \in M$ pontban. Akkor $f''(x)$ szimmetrikus mátrix.*

Bizonyítás. Nyilván elég a bizonyítást kétváltozós függvényekre elvégezni. Tegyük fel, hogy $f : \mathbb{R}^2 \to \mathbb{R}$ kétszer folytonosan differenciálható az (x, y) pontban. Legyen $v \in \mathcal{R}$ rögzített, és tekintsük az

$$F(t) = f(t, y + v) - f(t, y) , \qquad G(t) = f(x + v, t) - f(x, t)$$

függvényeket. A feltevésünk szerint ezek differenciálhatók az x, illetve az y pont egy környezetében, és

$$F(x + v) - F(x) = G(y + v) - G(y) . \tag{6.5}$$

A Lagrange-féle középértéktétel szerint található olyan $0 < t < 1$ szám, amelyre

$$F(x + v) - F(x) = F'(x + tv)v ,$$

azaz az F definíciójára tekintettel

$$\begin{aligned} F(x + v) - F(x) &= (D_1 f(x + tv, y + v) - D_1 f(x + tv, y))\, v \\ &= (D_{12} f(x + tv, y)v + o(v))\, v . \end{aligned}$$

Innen a második derivált folytonossága alapján

$$\lim_{v \to 0} \frac{F(x + v) - F(x)}{v^2} = D_{12} f(x, y) .$$

Teljesen hasonló gondolatmenettel az adódik, hogy

$$\lim_{v \to 0} \frac{G(y + v) - G(y)}{v^2} = D_{21} f(x, y) .$$

Ezért a (6.5) egyenlőségből azonnal következik, hogy

$$D_{12}f(x,y) = D_{21}f(x,y) ,$$

azaz a második derivált szimmetrikus mátrix. □

A következő tételünk lényegében a Taylor-formula kiterjesztése többváltozós függvényekre.

6.5.6 Tétel. *Tegyük fel, hogy f kétszer folytonosan differenciálható az $x \in X$ pont egy környezetében. Akkor*

$$f(x+v) = f(x) + f'(x)v + \frac{1}{2}\langle v, f''(x)v\rangle + o(\|v\|^2) ,$$

ahol

$$\lim_{v\to 0} \frac{o(\|v\|^2)}{\|v\|^2} = 0 .$$

Bizonyítás. Legyen adott $\epsilon > 0$. A második derivált folytonossága miatt az x pontnak van olyan környezete, amelyben

$$\|f''(x+v) - f''(x)\| < \epsilon .$$

Vezessük be a számegyenesen a

$$g(t) = f(x+tv)$$

függvényt. Mivel ekkor g egy elsőfokú függvény és az f kompozíciójaként áll elő, az összetett függvény differenciálhatósága alapján világos, hogy g kétszer folytonosan differenciálható a 0 egy környezetében, és

$$g'(0) = f'(x)v \qquad\qquad g''(0) = \langle v, f''(x)v\rangle .$$

Alkalmazzuk a g függvényre a Taylor-formulát, akkor található olyan $t \in [0,1]$ pont, amelyre

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(t) .$$

Mivel $g''(t) = \langle v, f''(x+tv)v\rangle$, innen azt kapjuk, hogy

$$f(x+v) = f(x) + f'(x)v + \frac{1}{2}\langle v, f''(x)v\rangle + \frac{1}{2}\langle v, (f''(x+tv) - f''(x))v\rangle .$$

Itt az $r(v) = 1/2\langle v, (f''(x+tv) - f''(x))v\rangle$ jelöléssel világos, hogy

$$|r(v)| \le \frac{1}{2}\|f''(x+tv) - f''(x)\|\|v\|^2 < \epsilon\|v\|^2$$

Ez éppen azt jelenti, hogy $r(v) = o(\|v\|^2)$, amit igazolnunk kellett. □

6.6 A szélsőérték másodrendű feltételei

Egyváltozós függvények esetében a derivált valamely zérushelye biztosan szélsőértékhely, ha ott a második derivált nem nulla, sőt az előjel azt is eldönti, hogy maximumról, vagy minimumról van szó. Ebben a szakaszban látni fogjuk, hogy többváltozós függvényekre analóg feltételek érvényesek, csupán a második derivált előjele helyett a megfelelő szimmetrikus mátrix definitségével van dolgunk.

Tekintsünk tehát egy X euklideszi teret, valamint egy $f : X \to \mathbb{R}$ kétszer folytonosan differenciálható függvényt. Első tételünk a szélsőérték szükséges feltételét fogalmazza meg.

6.6.1 Tétel. *Ha x az f lokális minimumhelye, akkor $f''(x)$ pozitív szemidefinit mátrix.*

Bizonyítás. Legyen $v \in X$ tetszőleges vektor, és tekintsük a $g(t) = f(x + tv)$ egyváltozós függvényt. A feltételünk szerint a 0 pont a g lokális minimumhelye. Másrészt a 6.2.6 Tétel szerint g kétszer differenciálható, ezért

$$0 \le g''(0) = \langle v, f''(x)v \rangle \ ,$$

és éppen ezt kellett igazolnunk. □

Természetesen analóg tétel érvényes maximum esetére is, akkor a második derivált negtív szemidefinit. Ezután rátérünk az elégséges feltétel bizonyítására.

6.6.2 Tétel. *Tegyük fel, hogy $f : X \to \mathbb{R}$ olyan kétszer folytonosan differenciálható fügvény, amelyre $f'(x) = 0$, valamint $f''(x)$ pozitív definit. Akkor x az f lokális minimumhelye.*

Bizonyítás. A Taylor-formula alapján (lásd a 6.5.6 Tételt) az x valamely környezetében

$$f(x + v) - f(x) = \frac{1}{2}\langle v, f''(x)v \rangle + o(\|v\|^2) \tag{6.6}$$

Jelölje λ az $f''(x)$ mátrix legkisebb sajátértékét, akkor a feltételünk szerint λ pozitív, és a 6.1 Következmény alapján

$$\langle v, f''(x)v \rangle \ge \lambda \|v\|^2$$

minden $v \in X$ vektor mellett. Legyen $\delta > 0$ olyan, hogy bármely $\|v\| < \delta$ esetén

$$\left| \frac{o(\|v\|^2)}{\|v\|^2} \right| < \frac{\lambda}{3} \ .$$

Ezeket a (6.6) egyenlőségbe visszahelyettesítve azt kapjuk, hogy

$$f(x + v) - f(x) \ge \frac{\lambda}{6}\|v\|^2 > 0 \ ,$$

hacsak $\|v\| < \delta$, $v \ne 0$. Ez éppen azt jelenti, hogy x az f lokális (szigorú) minimumhelye. □

Magától értetődően az $f''(x)$ negatív definitsége lokális maximumot jelent.

6.6.3 Példa. Felvetődhet a kérdés, hogy miért nem használtuk a 6.6.1 Tétel bizonyításának módszerét a 6.6.2 Tételre is. Abból ugyanis az adódna, hogy bármely

$v \in X$ mellett a $g(t) = f(x + tv)$ függvénynek a 0 pontban lokális minimumhelye van. Ebből azonban nem következik, hogy az f függvénynek az x pontban lokális mimimumhelye lenne, amint azt az alábbi példa is mutatja.

Tekintsük a síkon az

$$f(x, y) = (y - x^2)(y - 2x^2)$$

függvényt. Könnyen ellenőrízhető, hogy az origó kritikus pont. Másrészt e függvény az $y = x^2$, és $y = 2x^2$ parabolák között negatív értékeket, azokon kívül pedig pozitív értékeket vesz fel. Ezért bármely, az origón átmenő egyenesre leszűkítve a 0 pontban az f függvénynek lokális minimuma van, de az origó az f függvénynek nem lokális minimumhelye, hiszen f az origó bármely környezetében egyaránt felvesz pozitív és negatív értékeket is. Mellesleg

$$f''(0, 0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

nyilvánvalóan pozitív szemidefinit.

6.6.4 Példa. Vizsgáljuk most meg a háromváltozós

$$f(x, y, z) = xy^2z^3(7 - x - 2y - 3z)$$

függvényt. Ekkor a kritikus pontokra a

$$\begin{aligned} D_1 f(x, y, z) &= y^2z^3(7 - 2x - 2y - 3z) = 0 \\ D_2 f(x, y, z) &= 2xyz^3(7 - x - 3y - 3z) = 0 \\ D_3 f(x, y, z) &= 3xy^2z^2(7 - x - 2y - 4z) = 0 \end{aligned}$$

egyenletrendszer adódik, amelynek egyik megoldása az $(1, 1, 1)$ pont. Ezen a helyen a második derivált

$$f''(1, 1, 1) = \begin{bmatrix} -2 & -2 & -3 \\ -2 & -6 & -6 \\ -3 & -6 & -12 \end{bmatrix}.$$

Az $A - \lambda I$ mátrix elemi bázistranszformációjának elvégzése után az utolsó sorban a

$$\lambda^3 + 20\lambda^2 + 59\lambda + 42$$

polinomhoz jutunk. Ennek természetesen csak valós gyökei vannak, amelyek szükségképpen mind negatívok, hiszen a polinomnak minden együtthatója pozitív. Ez azt jelenti, hogy a második derivált negatív definit, azaz az $(1, 1, 1)$ pontban az f függvénynek lokális maximuma van.

6.6.5 Példa. Kétváltozós függvények esetében a második derivált definitsége egyszerűen ellenőrízhető. Tekintsük ugyanis az

$$f''(x, y) = \begin{bmatrix} D_{11} f(x, y) & D_{12} f(x, y) \\ D_{21} f(x, y) & D_{22} f(x, y) \end{bmatrix}$$

mátrixot az (x, y) kritikus pontban, és vezessük be a

$$D(x, y) = D_{11} f(x, y) D_{22} f(x, y) - (D_{12} f(x, y))^2$$

kifejezést. Ekkor a következő esetek lehetségesek.

- $D(x, y) < 0$ esetén a karakterisztikus polinom gyökei ellenkező előjelűek, így a második derivált indefinit. Ilyenkor az (x, y) pontban nincs szélsőérték.
- $D(x, y) > 0$ esetén a karakterisztikus polinom gyökei azonos előjelűek, így a második derivált definit mátrix. Méghozzá pozitív definit, ha $D_{11}f(x, y) > 0$ (azaz (x, y) lokális minimumhely), illetve negatív definit, ha $D_{11}f(x, y) < 0$ (azaz (x, y) lokális maximumhely).
- $D(x, y) = 0$ esetén minden lehetséges. Az $f(x, y) = x^3 + y^3$ esetében az origó nem szélsőértékhely, míg az $f(x, y) = x^4 + y^4$, illetve ennek negatívjának esetében az origó minimumhely, illetve maximumhely. Könnyen ellenőrízhető azonban, hogy mindhárom esetben $D(0, 0) = 0$.

6.6.6 Példa. Tegyük fel, hogy valamely kísérlet kimenetelére p számú megfigyelést végeztünk, és az $x_1, \ldots, x_p$ különböző helyeken az $y_1, \ldots, y_p$ értékek adódtak. Az az elképzelésünk, hogy ezekre a tapasztalati adatokra lineáris modell illeszthető, azaz egy olyan $y = mx + b$ egyenletű egyenest keresünk, amelyre

$$mx_1 + b = y_1 \qquad \ldots \qquad mx_p + b = y_p$$

Természetesen az adatok nem követik a mi hipotézisünket, ezért általában ilyen egyenes nem létezik. Ha ezt "mérési hibának" tudjuk be, és megelégszünk egy jó közelítéssel, akkor egy olyan egyenest keresünk, amely az adatainkat "jól" közelíti. Jó közelítésen azt értjük, hogy az

$$f(m, b) = \sum_{i=1}^{p} (mx_i + b - y_i)^2$$

négyzetösszeg minimális. Ezt a közelítő eljárást *legkisebb négyzetek* módszerének nevezzük.

A minimumhelyre a parciális deriváltakból a

$$\begin{aligned} D_1 f(m, b) &= \sum_{i=1}^{p} 2x_i(mx_i + b - y_i) = 0 \\ D_2 f(m, b) &= \sum_{i=1}^{p} 2(mx_i + b - y_i) = 0 \end{aligned}$$

egyenletrendszer adódik. Innen a

$$\begin{aligned} \sum_{i=1}^{p} x_i y_i &= m \sum_{i=1}^{p} x_i^2 + b \sum_{i=1}^{p} x_i \\ \sum_{i=1}^{p} y_i &= m \sum_{i=1}^{p} x_i + bp \end{aligned} \tag{6.7}$$

egyenletrendszert kapjuk, amiből az m és b ismeretlenek már könnyen meghatározhatók. Világos, hogy így minimumhoz jutunk, hiszen f teljes négyzetek összegeként áll elő.

Ezt a minimumhelyet deriválás nélkül, pusztán algebrai eszközökkel is megkaphatjuk. Ha bevezetjük az

$$A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_p & 1 \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}, \qquad z = \begin{bmatrix} m \\ b \end{bmatrix}$$

jelöléseket, akkor az $\mathbb{R}^p$ térben

$$f(z) = \|Az - y\|^2$$

alakban írható. Ez nyilván pontosan akkor minimális, ha az $Az-y$ vektor ortogonális az $\operatorname{im} A$ altérre. Ez azt jelenti, hogy az $\mathbb{R}^2$ mindkét e_i bázisvektorára $\langle y - Az, Ae_i\rangle = 0$. Innen egyszerű átalakítással az

$$\langle A^*y, e_i\rangle = \langle A^*Az, e_i\rangle$$

egyenlet adódik $i = 1, 2$ mellett, amiből

$$A^*y = A^*Az\ .$$

Itt A^*A nyilván invertálható, hiszen 2 rangú 2×2-es mátrix. Következésképpen

$$\begin{bmatrix} m \\ b \end{bmatrix} = z = (A^*A)^{-1}\, A^*y\ .$$

A kijelölt műveletek elvégzésével könnyen ellenőrízhető, hogy így is a (6.7) alatti egyenletrendszer megoldásához jutottunk.

6.7 Az implicitfüggvény-tétel

Számos feladatban felmerülő probléma, hogy valamely implicit módon megadott

$$f(x, y) = 0$$

egyenletből az y változó mikor fejezhető ki mint az x függvénye. Másként megfogalmazva, mikor található olyan g adott tulajdonságú függvény, amelyre a fenti egyenlet azonosság lesz, azaz

$$f(x, g(x)) = 0$$

teljesül.

Például a mikroökonómiában kézenfekvőnek tűnik (bár nem nyilvánvaló) az a feltevés, hogy a hasznossági illetve a termelési függvény szintvonalai (közömbösségi görbéi) két termék közötti függvénykapcsolatot fejeznek ki. Ezzel a kérdéskörrel foglalkozunk a következő szakaszban.

6.7.1 Tétel. (Implicitfüggvény-tétel) *Legyenek X, Y és Z euklideszi terek,* $\dim \mathrm{X} = \mathrm{p}$, $\dim \mathrm{Y} = \dim \mathrm{Z} = \mathrm{q}$, *legyen $(x_0, y_0) \in X \times Y$ adott pont, legyen $f : X \times Y \to Z$ adott függvény, $f(x_0, y_0) = 0$, legyen f folytonosan differenciálható az (x_0, y_0) pont egy környezetében, és tegyük fel, hogy $D_2f(x_0, y_0) \in L(Y)$ $m\times m$-méretű invertálható mátrix.*

Akkor létezik az x_0 pontnak olyan U, az y_0 pontnak olyan V környezete, és létezik pontosan egy olyan $g : U \to V$ folytonosan differenciálható függvény, amelyre

- $U \times V \subset D(f)$, $D(g) = U$, $R(g) = V$,
- $g(x_0) = y_0$,
- $\forall x \in U$ *esetén* $f(x, g(x)) = 0$,
- $\forall x \in U$ *esetén* $g'(x) = -D_2 f(x, g(x))^{-1} D_1 f(x, g(x))$.

Átfogalmazás: *Az*

$$f^{-1}(0) \cap (U \times V) = \{(x, y) \in X \times Y : f(x, y) = 0\} \cap (U \times V) \subset X \times Y$$

halmaz (reláció) az U *halmazon értelmezett differenciálható függvény, azaz* $f^{-1}(0) \cap U \times V = g : U \to Y$ *folytonosan differenciálható függvény.*

6.7.2 Megjegyzés.

- A fenti tétel igaz a 0 helyett $\forall\ z \in Z$ esetén: az $f^{-1}(z) \cap (U \times V) \subset X \times Y$ halmaz (reláció) az U halmazon értelmezett differenciálható függvény.
- A tétel nem állítja, hogy az $f^{-1}(z)$ az egész X-en függvény, csupán azt, hogy az x_0 egy környezetében az.
- A tételt ilyen általánosan most nem bizonyítjuk. A közgazdaságtanban azonban sokszor elég a fenti tétel kétdimenziós speciális esete is, amely viszonylag könnyen belátható.

6.7.3 Tétel. **(Implicitfüggvény-tétel, speciális eset)** *Legyenek* $I, J \subseteq \mathbb{R}$ *nyílt intervallumok,* $(x_0, y_0) \in I \times J$ *adott pont, legyen* $f : I \times J \to \mathbb{R}$ *adott függvény, amelyre* $f((x_0, y_0)) = 0$*, tegyük fel, hogy az* $f : I \times J \to \mathbb{R}$ *függvény folytonosan diffható az* $(x_0, y_0) \in I \times J$ *egy környezetében, és a* $D_2 f((x_0, y_0)) \neq 0$.

Akkor létezik az x_0 *pontnak olyan* $U = [x_0 - \delta, x_0 + \delta]$ *és az* y_0 *pontnak olyan* $V = [y_0 - \varepsilon, y_0 + \varepsilon]$ *környezete, és létezik pontosan egy* $g : [x_0 - \delta, x_0 + \delta] \to [y_0 - \varepsilon, y_0 + \varepsilon]$ *folytonosan differenciálható függvény, hogy*

- $[x_0 - \delta, x_0 + \delta] \times [y_0 - \varepsilon, y_0 + \varepsilon] \subset I \times J$, *és* $R(g) \subset [y_0 - \varepsilon, y_0 + \varepsilon]$,
- $g(x_0) = y_0$,
- $\forall x \in [x_0 - \delta, x_0 + \delta]$ *esetén* $f(x, g(x)) = 0$,
- $\forall x \in (x_0 - \delta, x_0 + \delta)$ *esetén* $g'(x) = -\frac{D_1 f(x, g(x))}{D_2 f(x, g(x))}$.

Bizonyítás. Nyilván feltehető, hogy $D_2 f(x_0, y_0) > 0$. Mivel az f folytonosan differenciálható az $(x_0, y_0) \in I \times J$ pont egy környezetében, azért $\exists\ \gamma, \varepsilon > 0$ számok, hogy $\forall\ (x, y) \in [x_0 - \gamma, x_0 + \gamma] \times [y_0 - \varepsilon, y_0 + \varepsilon]$ esetén $D_2 f(x, y) > 0$, így $\forall\ x \in [x_0 - \gamma, x_0 + \gamma]$ esetén az $[y_0 - \varepsilon, y_0 + \varepsilon]$ intervallumon $D_2 f(x, \cdot) > 0$, így az $f(x, \cdot) : [y_0 - \varepsilon, y_0 + \varepsilon] \to \mathbb{R}$ függvény szigorúan monoton növekedő.

Mivel $f(x_0, y_0) = 0$, azért $f(x_0, y_0 - \varepsilon) < 0$ és $f(x_0, y_0 + \varepsilon) > 0$. Mivel az f folytonos, azért $\exists \delta \in (0, \gamma)$, hogy $\forall x \in [x_0 - \delta, x_0 + \delta]$ esetén $f(x, y_0 - \varepsilon) < 0$ és $f(x, y_0 + \varepsilon) > 0$.

Mivel az $f(x,\cdot) : [y_0-\varepsilon, y_0+\varepsilon] \to \mathbb{R}$ függvény folytonos, azért a Bolzano-tétel szerint létezik, mivel szigorúan monoton, azért pontosan egy $y_x \in [y_0-\varepsilon, y_0+\varepsilon]$ létezik, amelyre $f(x,y_x) = 0$.

Legyen $g : [x_0-\delta, x_0+\delta] \to [y_0-\varepsilon, y_0+\varepsilon]$ az a függvény, amelyre $\forall x \in [x_0-\delta, x_0+\delta]$ esetén $g(x) := y_x$.

A g definíciójából következik, hogy $\forall x \in [x_0-\delta, x_0+\delta]$ esetén $f(x,g(x)) = 0$, mivel $\forall x \in [x_0-\delta, x_0+\delta]$ esetén pontosan egy fenti tulajdonságú y_x található, azért pontosan egy ilyen g függvény létezik.

Megmutatjuk, hogy a g függvény differenciálható $\forall\ x \in (x_0-\delta, x_0+\delta)$ pontban. Legyen $z \in (x_0-\delta, x_0+\delta)$ tetszőleges pont, ekkor a Lagrange-középértéktétel szerint létezik olyan u az x és a z között, valamint v a $g(x)$ és a $g(x)$ között, melyekre

$$\begin{aligned}
0 &= f(z,g(z)) - f(x,g(x)) = \\
&= f(z,g(z)) - f(x,g(z)) + f(x,g(z)) - f(x,g(x)) = \\
&= D_1f(u,g(z)) \cdot (z-x) + D_2f(x,v) \cdot (g(z)-g(x)).
\end{aligned}$$

Így $D_2f(x,v) \neq 0$ miatt

$$\frac{g(z)-g(x)}{z-x} = -\frac{D_1f(u,g(z))}{D_2f(x,v)}.$$

Ha most tudnánk, hogy a fenti egyenlőség jobboldala korlátos, akkor abból már adódna, hogy a g függvény folytonos. Ez sajnos az eddigiekből nem következik, de könnyen látható, hogy ha a bizonyítás elején körültekintőbben választjuk meg a δ-t és az ε-t, akkor a fenti egyenlőség jobboldala korlátos lesz. Nevezetesen a D_1f és a D_2f (x_0,y_0)-beli folytonossága és $D_2f((x_0,y_0)) > 0$ miatt a δ és $\varepsilon > 0$ számok vállaszthatók olyan kicsire, hogy $\forall\ (x,y) \in [x_0-\delta, x_0+\delta] \times [y_0-\varepsilon, y_0+\varepsilon]$ esetén

$$D_2f((x,y)) \geq \frac{D_2f((x_0,y_0))}{2} \quad \text{és} \quad |D_1f(x,y)| \leq |D_1f(x_0,y_0)| + 1.$$

Ekkor $\forall\ (x,y),(u,v) \in [x_0-\delta, x_0+\delta] \times [y_0-\varepsilon, y_0+\varepsilon]$ esetén

$$\left| -\frac{D_1f(x,y)}{D_2f(u,v)} \right| \leq 2\frac{|D_1(x_0,y_0)|+1}{D_2f(x_0,y_0)} =: K,$$

amit éppen akartunk. Ezek szerint ha a δ-t és az ε-t a fentiek szerint választjuk meg, akkor a g függvény folytonos.

Továbbá mivel az f' és a g folytonos függvények, azért $\forall\ \alpha > 0$ esetén létezik az x-nek olyan W környezete, hogy $\forall\ z \in W$ esetén

$$\left| \frac{D_1f(x,g(x))}{D_2f(x,g(x))} - \frac{D_1f(u,g(z))}{D_2f(x,v)} \right| < \alpha,$$

Mivel pedig $\frac{g(z)-g(x)}{z-x} = -\frac{D_1f(u,g(z))}{D_2f(x,v)}$, azért

$$\left| \frac{g(z)-g(x)}{z-x} + \frac{D_1f(x,g(x))}{D_2f(x,g(x))} \right| < \alpha,$$

ezért a g függvény differenciálható az x pontban és

$$g'(x) = \frac{D_1f(x,g(x))}{D_2f(x,g(x))}.$$

Innen pedig a g, a $D_1 f$ és a $D_2 f$ folytonossága alapján adódik, hogy g' is folytonos. Mivel $f(x_0, y_0) = 0$, így a $g(x_0)$-ra vonatkozó egyértelműségi feltétel miatt $g(x_0) = y_0$. Ezzel a tételt bebizonyítottuk. □

6.7.4 Példa. Tekintsük a síkon az

$$f(x, y) = e^{x+y} + x + y - 1 = 0$$

egyenletet. Világos, hogy $f(0, 0) = 0$, és $D_2 f(0, 0) = 2$, így teljesülnek az implicitfüggvény-tétel feltételei. Tehát található egyetlen olyan folytonosan differenciálható $y = g(x)$ függvény a 0 pont környezetében, amelyre a fenti egyenlet azonosság. Erre a függvényre a tételünkből

$$g'(x) = -\frac{1}{e^{x+g(x)} + 1}(e^{x+g(x)} + 1) = -1$$

adódik, azaz $g(x) = -x$, amelyet y helyére írva valóban azonossághoz jutunk.

6.7.5 Példa. Az y változó kifejezhetőségét nem algebrai értelemben kell értenünk, tehát előfordulhat, hogy a g függvény létezését igazolni tudjuk, de azt algebrai átalakításokkal a fenti egyenletből nem tudjuk előállítani. Tekintsük például az

$$e^{x+y} - 2\cos y + 1 = 0$$

egyenletet. Meg lehet mutatni, hogy ez az egyenlet egy folytonosan differenciálható függvényt definiál, azaz létezik pontosan egy olyan g folytonosan differenciálható függvény, amelyre $g(0) = 0$, és

$$e^{x+g(x)} - 2\cos g(x) + 1 = 0$$

minden x esetén, de persze az y változó a fenti egyenletből algebrai átalakításokkal nem fejezhető ki.

6.7.6 Példa. (Egy mikroökonómiai példa: A helyettesítési határarány):

A mikroökonómiában a hasznossági függvény a jószágtéren értelmezett olyan függvény, amely a fogyasztó preferenciáit fejezi ki. Feltéve, hogy két jószágunk van, legyen $u : \mathbb{R}^2_+ \to \mathbb{R}$ egy hasznossági függvény, ekkor egy adott $\alpha \in \mathbb{R}$ hasznossági szinthez tartozó közömbösségi görbe az

$$u^{-1}(\alpha) = \{(x_1, x_2) : u(x_1, x_2) = \alpha\} \subset \mathbb{R}^2_+$$

szinthalmaz. Ez a halmaz (reláció) nem biztos, hogy függvény, de ha az u-ra teljesülnek a fenti tétel feltételei, akkor egy U környezetben az, azaz $u^{-1}(\alpha) = g : U \to \mathbb{R}$ függvény, ami differenciálható is, és

$$g'(x_1) = -\frac{D_1 u(x_1, g(x_1))}{D_2 u(x_1, g(x_1))}$$

(a mikroökonómiában a g függvényt x_2-vel szokták jelölni, ekkor az előbbi összefüggés a következő alakú: $x_2'(x_1)(= \frac{dx_2}{dx_1}) = -\frac{D_1 u(x_1, x_2(x_1))}{D_2 u(x_1, x_2(x_1))}$), azaz helyettesítési határráta megyegyezik a határhasznok hányadosának az ellentettjével.

Ebből adódik az is, hogy ha $u : \mathbb{R}^2_+ \to \mathbb{R}$ függvény monoton növekedő, (a deriváltja nemnegatív) akkor a $g : U \to \mathbb{R}$ függvény monoton csökkenő. Könnyen látható továbbá, hogy ha az $u : \mathbb{R}^2_+ \to \mathbb{R}$ függvény konkáv, akkor a $g : U \to \mathbb{R}$ függvény konvex, így a g' derivált függvény nő, mivel g' negatív, azért abszolútértékben csökken, azaz a helyettesítési határarány abszolútértékben csökken.

Ugyanez mondható el a termelési függvények esetében:

Egy $f : \mathbb{R}^2_+ \to \mathbb{R}$ termelési függvényre feltéve a fenti tétel feltételeit, azt kapjuk, hogy egy környezetben az $f^{-1}(\alpha)$ halmaz függvény, azaz $f^{-1}(\alpha) = g : U \to \mathbb{R}$ függvény, ami differenciálható is, és $g'(x_1) = -\frac{D_1 f(x_1,g(x_1))}{D_2 f(x_1,g(x_1))}$, (a mikroökonómiában megszokott jelölésekkel: $x_2'(x_1)(=\frac{dx_2}{dx_1}) = -\frac{D_1 f(x_1,x_2(x_1))}{D_2 f(x_1,x_2(x_1))}$,) azaz a technikai helyettesítési határráta megyegyezik a határtermékek hányadosának az ellentettjével.

6.7.7 Példa. (Egy makroökonómiai példa: Az IS és az LM görbék):

1. Az IS (investment–saving) görbe:

A beruházás a kamatlábtól függ: $I(i)$, a megtakarítás a kibocsátástól függ: $S(Y)$. Legyen $F : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ az a függvény, amelyre $\forall\ Y, i \in \mathbb{R}_+$ esetén $F(Y,i) := S(Y) - I(i)$, ekkor az $S(Y) = I(i)$ egyenlőségnek eleget tévő kamatláb–jövedelem párok halmaza az

$$F^{-1}(0) = \{(Y,i) \in \mathbb{R}^2_+ : F(Y,i) = 0\} \subset \mathbb{R}^2_+$$

halmaz (reláció), amely, ha az F függvényre igazak a fenti tétel feltételei, akkor egy környezetben differenciálható függvény:

$$F^{-1}(0) = i : U \to \mathbb{R}.$$

Kérdés, hogy milyen az i függvényalakja? Mivel a fenti tétel szerint az $i'(Y)(=\frac{di}{dY}) = -\frac{D_1 F(Y,i(Y))}{D_2(Y,i(Y))} = -\frac{S'(Y)}{I'(i)}$, ezért feltéve, hogy $S'(Y) > 0$ és $I'(i) < 0$, (azaz a megtakarítás a kibocsátás esetén nő, a beruházás a kamatláb növekedése esetén csökken,) adódik, hogy $i'(Y) < 0$, azaz az i csökkenő függvény. (Ha a kormányzat a kibocsátást növeli, akkor a kamatláb csökken.)

2. Az LM (liquidity–money) görbe:

A pénzkereslet a kibocsátástól és a kamatlábtól függ: $M_D(Y,i)$, a pénzkínálat állandó: M/P. Legyen $F : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ az a függvény, amelyre $\forall\ Y, i \in \mathbb{R}_+$ esetén $F(Y,i) := M_D(Y,i) - M/P$, ekkor az $M_D(Y,i) = M/P$ egyenlőségnek eleget tévő kamatláb-jövedelem párok halmaza az

$$F^{-1}(0) = \{(Y,i) \in \mathbb{R}^2_+ : F(Y,i) = 0\} \subset \mathbb{R}^2_+$$

halmaz (reláció), amely, ha az F függvényre igazak a fenti tétel feltételei, akkor egy környezetben függvény:

$$F^{-1}(0) = Y : U \to \mathbb{R},$$

amely differenciálhatóés $Y'(i)(=\frac{dY}{di}) = -\frac{D_2 F(Y(i),i)}{D_1 F(Y(i),i)} = -\frac{D_2 M(Y(i),i)}{D_1 M(Y(i),i)}$, ezért feltéve, hogy $D_1 M(Y,i) > 0$ és $D_2 M(Y,i) < 0$, (azaz a pénzkereslet a kibocsátás növekedése esetén nő, a kamat növekedése esetén csökken,) adódik, hogy $Y'(i) > 0$, azaz az Y növekvő függvény. (Ha a központi bank a kamatlábat növeli, akkor a kibocsátás nő.)

6.8 Feltételes szélsőérték

Számos szélsőérték probléma vezet olyan feladathoz, amelyben az f függvény szélsőértékét egy adott K halmazon kell meghatározni. Ilyen esetekben az $f'(x) = 0$ feltétel már nem feltétlenül szükséges feltétele a szélsőértéknek, hiszen elképzelhető, hogy az f függvény a szélsőértékét a K halmaz határán veszi fel.

Tekintsük például az $f(x, y) = x + y$ függvényt, és keressük az f minimumát az $|x| \leq 1$, $|y| \leq 1$ feltételek mellett. Ha bevezetjük a

$$K = \Big\{(x, y) \in \mathbb{R}^2 : |x| \leq 1,\ |y| \leq 1\Big\}$$

halmazt (amely egy origó középpontú, két egységnyi oldalú négyzet), akkor a feladatunk az f minimumhelyének megkeresése a K halmazon. Könnyen látható, hogy f a minimumát ezen a halmazon az $(x, y) = (-1, -1)$ pontban veszi fel, de f deriváltja természetesen sehol sem nulla. Az f függvénynek persze az egész $\mathbb{R}^2$ téren nincs minimuma.

Legyenek tehát X és Y valós euklideszi terek, $\dim X = p$, $\dim Y = q$, és tegyük fel, hogy $q \leq p$. Tekintsük az $f : X \to \mathbb{R}$ és $F : X \to Y$ folytonosan differenciálható függvényeket. Legyen $a \in Y$ tetszőleges adott pont. Keressük az f függvény lokális minimumhelyét az $F(x) = a$ feltétel mellett, jelölésben

$$\begin{aligned} &f(x) \to \min \\ &F(x) = a\,. \end{aligned} \tag{6.8}$$

Ha bevezetjük a

$$K = \{x \in X : F(x) = a\} = F^{-1}(a)$$

jelölést, akkor a fenti feladat az

$$\begin{aligned} &f(x) \to \min \\ &x \in K \end{aligned}$$

alakban is felírható.

6.8.1 Definíció. Azt mondjuk, hogy az $x_0 \in X$ pont a (6.8) feladat *megoldása*, ha egyrészt $F(x_0) = a$, másrészt az x pontnak van olyan U környezete, hogy

$$f(x_0) \leq f(x)$$

bármely $x \in U \cap K$ esetén.

6.8.2 Definíció. A (6.8) feladat *Lagrange-függvényén* az $\mathcal{L} : X \times Y \to \mathbb{R}$,

$$\mathcal{L}(x, y) = f(x) + \langle y, F(x)\rangle$$

függvényt értjük.

Nyilvánvaló, hogy a Lagrange-függvény mindkét változója szerint folytonosan differenciálható. Az egyszerűbb jelölésmód érdekében a következőkben az L első, illetve második változó szerinti parciális deriváltján az x, illetve az y szerinti deriváltakat értjük.

6.8.3 Tétel. *(***Lagrange-féle multiplikátor-tétel***) Tegyük fel, hogy x_0 a (6.8) feladat megoldása, és* $\operatorname{im} F'(x_0) = Y$*, azaz az $F'(x_0)$ mátrix sorai lineárisan függetlenek. Ekkor található olyan $y \in Y$ vektor, hogy*

$$D_1\mathcal{L}(x_0, y) = f'(x_0) + F'(x_0)^* y = 0 \, . \tag{6.9}$$

Legyenek az Y egy ortonormált bázisára nézve az $y \in Y$ vektornak a koordinátái $\lambda_1, ...\lambda_q$, ekkor a (6.9) egyenletet az

$$f'(x_0) + \sum_{i=1}^{q} \lambda_i f_i'(x_0) = 0 \tag{6.10}$$

alakban is felírhatjuk, ahol az f_i függvények az F koordinátafüggvényei. Ezek szerint a fenti tétel úgy is fogalmazható, hogy az optimális pontban a feltételi függvények deriváljainak van olyan lineáris kombinációja amely a célfüggvény deriváltját állítja elő. A $\lambda_1, ...\lambda_q$ együtthatókat *Lagrange-féle multiplikátoroknak* nevezzük.

A fenti tételben természetesen a $D_2\mathcal{L}(x_0, y) = a$ feltétel is teljesül, hiszen $D_2\mathcal{L}(x_0, y) = F(x) = a$. Ha ezt az egyenletet a (6.10) egyenlethez csatoljuk, akkor az x és y koordinátáiból álló $p + q$ darab ismeretlenre $p + q$ darab egyenlet adódik, azaz

$$\begin{aligned} \begin{bmatrix} D_1 f(x_0) \\ \vdots \\ D_p f(x_0) \end{bmatrix} + \sum_{i=1}^{q} \lambda_i \begin{bmatrix} D_1 f_i(x_0) \\ \vdots \\ D_p f_i(x_0) \end{bmatrix} &= \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \\ \begin{bmatrix} f_1(x_0) \\ \vdots \\ f_q(x_0) \end{bmatrix} &= \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix} . \end{aligned} \tag{6.11}$$

A Lagrange-féle multiplikátor-tételt is csak a kétdimenziós speciális esetben bizonyítjuk be. Ekkor a tételnek igen szemléletes a tartalma.

Legyenek $I,\ J \subset \mathbb{R}$ nyílt intervallumok, legyenek $f_0 : I \times J \to \mathbb{R}$ és $f_1 : I \times J \to \mathbb{R}$ adott függvénynyek, tekintsük a fenti (6.8) feladatnak a következő speciális esetét:

$$\begin{aligned} f_0(x, y) &\to \min \\ f_1(x, y) &= \alpha \, . \end{aligned} \tag{6.12}$$

6.8.4 Tétel. (Lagrange-féle multiplikátor-tétel, speciális eset) *Legyen $(x_0, y_0) \in I \times J$ a (6.12) feladat megoldása, legyen az $f_0 : I \times J \to \mathbb{R}$ függvény differenciálható az (x_0, y_0) pontban, legyen az $f_1 : I \times J \to \mathbb{R}$ függvény olyan, amelyre teljesülnek az implicitfüggvény-tétel feltételei, azaz folytonosan differenciálható az (x_0, y_0) pont egy környezetében és a $D_2 f_1(x_0, y_0) \neq 0$.*

Akkor $\exists\ \lambda \in \mathbb{R}$ Lagrange-szorzó, hogy

$$D_{1,2}\mathcal{L}((x_0, y_0), \lambda) = (0, 0), \text{ azaz}$$

$$f_0'(x_0, y_0) + \lambda f_1'(x_0, y_0) = (0, 0), \text{ azaz}$$

$$D_1f_0(x_0,y_0)+\lambda D_1f_1(x_0,y_0)=0 \quad \textit{és} \quad D_2f_0(x_0,y_0)+\lambda D_2f_1(x_0,y_0)=0,$$

továbbá

$$\lambda=-\frac{D_2f_0(x_0,y_0)}{D_2f_1(x_0,y_0)},$$

valamint feltéve, hogy $D_2f_0(x_0,y_0)\neq 0$, *teljesül, hogy*

$$\frac{D_1f_0(x_0,y_0)}{D_2f_0(x_0,y_0)}=\frac{D_1f_1(x_0,y_0)}{D_2f_1(x_0,y_0)}$$

Bizonyítás. Mivel az f_1 függvényre, így az $f_1-\alpha$ függvényre is fennállnak az implicitfüggvény-tétel feltételei, azért $\exists\,[x_0-\delta,x_0+\delta]$ környezet, hogy az $f_1^{-1}(\alpha)\subset[x_0-\delta,x_0+\delta]\times J$ halmaz (reláció) függvény, azaz $\exists!\ g:[x_0-\delta,x_0+\delta]\to J$ függvény, hogy $g(x_0)=y_0$, $\forall x\in[x_0-\delta,x_0+\delta]$ esetén $f_1(x,g(x))=\alpha$, valamint $g'(x_0)=-\frac{D_1f_1(x_0,y_0)}{D_2f_1(x_0,y_0)}$.

Legyen $h:[x_0-\delta,x_0+\delta]\to\mathbb{R}$ az a függvény, amelyre $\forall\ x\in[x_0-\delta,x_0+\delta]$ esetén $h(x):=f_0(x,g(x))$. Mivel az f_0 differenciálható az (x_0,y_0) pontban és $g(x_0)=y_0$, azért a h is differenciálható az x_0 pontban, és

$$\begin{aligned}
h'(x_0) &= f_0'(x_0,g(x_0))\cdot\begin{bmatrix}1\\ g'(x_0)\end{bmatrix}\\
&= [D_1f_0(x_0,y_0),D_2f_0(x_0,y_0)]\cdot\begin{bmatrix}1\\ g'(x_0)\end{bmatrix}\\
&= D_1f_0(x_0,y_0)+D_2f_0(x_0,y_0)\cdot g'(x_0)\\
&= D_1f_0(x_0,y_0)-D_2f_0(x_0,y_0)\cdot\frac{D_1f_1(x_0,y_0)}{D_2f_1(x_0,y_0)}\\
&= D_1f_0(x_0,y_0)-D_1f_1(x_0,y_0)\cdot\frac{D_2f_0(x_0,y_0)}{D_2f_1(x_0,y_0)}\ .
\end{aligned}$$

Továbbá, mivel egyrészt az (x_0,y_0) a fenti feladat megoldása, másrészt $g(x_0)=y_0$ és $\forall x\in[x_0-\delta,x_0+\delta]$ esetén $f_1(x,g(x))=\alpha$, azért az x_0 a h függvény minimumhelye. Ezért $h'(x_0)=0$, azaz

$$D_1f_0(x_0,y_0)-D_1f_1(x_0,y_0)\cdot\frac{D_2f_0(x_0,y_0)}{D_2f_1(x_0,y_0)}=0. \tag{6.13}$$

Ezek szerint a $\lambda:=-\frac{D_2f_0(x_0,y_0)}{D_2f_1(x_0,y_0)}$ választással egyrészt a λ definíciójából nyilván

$$D_2f_0(x_0,y_0)+\lambda D_2f_1(x_0,y_0)=0,$$

másrészt a 6.13 szerint

$$D_1f_0(x_0,y_0)+\lambda D_1f_1(x_0,y_0)=0\ .$$

Szintén a 6.13 szerint ha $D_2f_0(x_0,y_0)\neq 0$, akkor

$$\frac{D_1f_0(x_0,y_0)}{D_2f_0(x_0,y_0)}=\frac{D_1f_1(x_0,y_0)}{D_2f_1(x_0,y_0)}\ . \qquad \square$$

6.8.5 Megjegyzés. A fenti tételben tegyük fel, hogy nem csak az $f_1 : I \times J \to \mathbb{R}$ függvényre, hanem az $f_0 : I \times J \to \mathbb{R}$ függvényre is teljesülnek az implicitfüggvény-tétel feltételei. Ekkor a tétel jelentése igen szemléletes és úgy fogalmazható, hogy ha az (x_0, y_0) a fenti feladat megoldása, akkor a két függvény szintvonalai érintik egymást az (x_0, y_0) pontban.

Ugyanis ekkor egyrészt, mint a bizonyításban már láttuk, mivel az f_1 függvényre, így az $f_1 - \alpha$ függvényre is fennállnak az implicitfüggvény-tétel feltételei, azért $\exists\ [x_0 - \delta, x_0 + \delta]$ környezet, hogy az $f^{-1}(\alpha) \subset [x_0 - \delta, x_0 + \delta] \times J$ halmaz (reláció) differenciálható függvény, azaz $\exists!\ g_1 : [x_0 - \delta, x_0 + \delta] \to J$ függvény, hogy $g_1(x_0) = y_0$, $\forall x \in [x_0 - \delta, x_0 + \delta]$ esetén $f_1(x, g_1(x)) = \alpha$, valamint $g_1'(x_0) = -\frac{D_1 f_1(x_0,y_0)}{D_2 f_1(x_0,y_0)}$.

Másrészt, ha az f_0 függvényre is fennállnak az implicitfüggvény-tétel feltételei, akkor ugyanígy $\exists\ [x_0 - \delta_0, x_0 + \delta_0]$ környezet, hogy az $f^{-1}(f_0(x_0, x_0)) \subset [x_0 - \delta_0, x_0 + \delta_0] \times I$ halmaz (reláció) differenciálható függvény, azaz $\exists!\ g_0 : [x_0 - \delta_0, x_0 + \delta_0] \to I$ függvény, hogy $g_0(x_0) = y_0$, $\forall x \in [x_0 - \delta_0, x_0 + \delta_0]$ esetén $f_0(x, g(x)) = f_0(x_0, y_0)$, valamint $g_0'(x_0) = -\frac{D_1 f_0(x_0,y_0)}{D_2 f_0(x_0,y_0)}$.

A tétel szerint
$$\frac{D_1 f_0(x_0, y_0)}{D_2 f_0(x_0, y_0)} = \frac{D_1 f_1(x_0, y_0)}{D_2 f_1(x_0, y_0)} \ , \quad \text{ezért}$$
$$g_0'(x_0) = g_1'(x_0) \ ,$$
mivel $g_0(x_0) = y_0 = g_1(x_0)$, azért ez azt jelenti, hogy a két függvény szintvonalai érintik egymást az (x_0, y_0) pontban.

Az itt elmondottak jól illusztrálhatók a következű példán keresztül.

6.8.6 Példa. Oldjuk meg a következő feladatot:

$$\begin{aligned} & x \cdot y \to \max \\ & x^2 + y^2 = 1 \ . \end{aligned} \tag{6.14}$$

A fenti tétel jelöléseivel $f_0(x, y) = x \cdot y$ és $f_1(x, y) = x^2 + y^2$. A feladat Lagrange-függvénye $\mathcal{L}(x, y) = x \cdot y + \lambda(x^2 + y^2)$, továbbá $D_1 f_0(x, y) = y$ és $D_2 f_0(x, y) = x$, továbbá $D_1 f_1(x, y) = 2x$ és $D_2 f_1(x, y) = 2y$. Ezért ha (x_0, y_0)megoldása (6.14) -nak, akkor a Lagrange-elv szerint létezik $\lambda \in \mathbb{R}$, hogy

$$D_{1,2}\mathcal{L}((x_0, y_0), \lambda) = [y_0, x_0] + \lambda \cdot [2x_0, 2y_0] = (0, 0), \text{ azaz}$$
$$[y_0, x_0] = -\lambda \cdot [2x_0, 2y_0].$$

Az x_0 szám nyilván nem lehet 0, mivel ekkor y_0 is 0 lenne, ellentmondásban az $x_0^2 + y_0^2 = 1$ feltétellel. Így

$$x_0 = (-\lambda) \cdot 2y_0 = (-\lambda) \cdot 2 \cdot (-\lambda) \cdot 2x_0 = 4\lambda^2 \cdot x_0$$

alapján $4\lambda^2 = 1$. Innen $\lambda = \pm\frac{1}{2}$, tehát $x_0 = y_0$ vagy $x_0 = -y_0$. Behelyettesítve az egyenlőség-feltételbe, mindenképpen azt kapjuk, hogy $2x_0^2 = 1$, azaz a megoldások csak
$$\left(\sqrt{\frac{m}{2}}, \sqrt{\frac{m}{2}}\right), \left(-\sqrt{\frac{m}{2}}, -\sqrt{\frac{m}{2}}\right), \left(\sqrt{\frac{m}{2}}, -\sqrt{\frac{m}{2}}\right), \left(-\sqrt{\frac{m}{2}}, \sqrt{\frac{m}{2}}\right)$$
közül kerülhetnek ki. Az f függvény értéke az első két helyen $\frac{m}{2}$, a második két helyen $-\frac{m}{2}$. Mivel kompaktsági megfontolások miatt 6.14 -nak úgyis van

megoldása, ezért a fenti első két számpár valóban a megoldásokat szolgáltatja. A másik két számpár a minimum feladat megoldása. Látható továbbá, hogy például $g_0'\left(\sqrt{\frac{m}{2}}, \sqrt{\frac{m}{2}}\right) = -1 = g_1'\left(\sqrt{\frac{m}{2}}, \sqrt{\frac{m}{2}}\right)$

6.8.7 Példa. Mutassuk meg, hogy egy háromszög szögeire

$$\cos\alpha\cos\beta\cos\gamma \le \frac{1}{8}\,,$$

és egyenlőség csak szabályos háromszögekre érvényes.

Esetünkben legyen $f(\alpha,\beta,\gamma) = \cos\alpha\cos\beta\cos\gamma$, és $F(\alpha,\beta,\gamma) = \pi - \alpha - \beta - \gamma$. Ekkor a Lagrange-függvény az

$$\mathcal{L}(\alpha,\beta,\gamma,y) = \cos\alpha\cos\beta\cos\gamma + y(\pi - \alpha - \beta - \gamma)$$

alakot ölti. A megoldásra tehát a

$$\begin{aligned} -\sin\alpha\cos\beta\cos\gamma - y &= 0 \\ -\cos\alpha\sin\beta\cos\gamma - y &= 0 \\ -\cos\alpha\cos\beta\sin\gamma - y &= 0 \end{aligned}$$

szükséges feltétel adódik. Ha még figyelembe vesszük, hogy $F(\alpha,\beta,\gamma) = \pi - \alpha - \beta - \gamma = 0$, akkor az egyenletrendszer egyetlen megoldása $\alpha = \beta = \gamma = \pi/3$, azaz a feltételt csak a szabályos háromszögek elégítik ki.

Könnyen belátható, hogy ebben a példában a feladat feltétel nélküli szélsőérték-feladattá alakítható át. Valóban, az $F(\alpha,\beta,\gamma) = 0$ egyenletből $\gamma = \pi - \alpha - \beta$. Ezt az f függvénybe helyettesítve

$$f(\alpha,\beta) = -\cos\alpha\cos\beta\cos(\alpha+\beta)\,.$$

Ekkor a szélsőérték szükséges feltétele

$$\begin{aligned} D_1 f(\alpha,\beta) &= \sin\alpha\cos\beta\cos(\alpha+\beta) + \cos\alpha\cos\beta\sin(\alpha+\beta) = 0 \\ D_2 f(\alpha,\beta) &= \cos\alpha\sin\beta\cos(\alpha+\beta) + \cos\alpha\cos\beta\sin(\alpha+\beta) = 0\,. \end{aligned}$$

Ennek az egyenletrendszernek 0 és π között egyetlen megoldása van, méghozzá $\alpha = \beta = \pi/3$. Azonnal látható, hogy ez valóban maximumhely, hiszen itt

$$f''(\pi/3,\pi/3) = \begin{bmatrix} -1 & -1/2 \\ -1/2 & -1 \end{bmatrix}$$

ami nyilvánvalóan negatív definit.

6.8.8 Megjegyzés. A fenti példák nagyon egyszerűek voltak, inkább csak szemléltették a tételt. Ha a feltételi halmaz több egyenletből áll, akkor a (6.11), azaz az $\mathcal{L}'(x_0, y) = 0$ egyenletrendszer olyan bonyolult, amelyből a megoldás amúgy sem határozható meg. Ezért a Lagrange-féle multiplikátor-tétel az alkalmazások szempontjából inkább elvi jelentőségűnek tekinthető, azaz az igazi feladata az, hogy más tudományokban adott elméleteket alátámasszon. Erre mutatunk példát a következőkben a mikroökonómiából.

6.8.9 Példa. A mikroökonómiában központi szerepet játszanak a feltételes szélsőérték feladatok, így a Lagrange-féle multiplikátor-tétel, ugyanis a fogyasztókról felteszik, hogy a hasznosságukat maximalizálják, a termelőkről pedig felteszik, hogy a profitjukat maximalizálják.

I. A fogyasztói viselkedés

1. A Marshall-féle megközelítés: A fogyasztó adott jövedelem szint mellett a hasznosságát maximalizálja.

Legyen a fogyasztó hasznossági függvénye $u : \mathbb{R}^n_+ \to \mathbb{R}$, legyen $p \in \mathbb{R}^n_+$ az árak vektora, ekkor a költségfüggvénye a $\langle p, . \rangle : \mathbb{R}^n_+ \to \mathbb{R}$ lineáris funkcionál, (adott $x \in \mathbb{R}$ termék költsége $\langle p, x \rangle$,) legyen $m \in \mathbb{R}$ a fogyasztó jövedelme. A feladat:

$$\begin{aligned} &u(x) \to \max \\ &\langle p, x \rangle = m \, . \end{aligned} \tag{6.15}$$

A feladat Lagrange-függvénye az az $\mathcal{L} : \mathbb{R}^n_+ \times \mathbb{R} \to \mathbb{R}$ függvény, amelyre $\forall\ x \in \mathbb{R}^n_+$, $\forall\ \lambda \in \mathbb{R}$ esetén

$$\mathcal{L}(x, y) = u(x) + \lambda \cdot (\langle p, x \rangle - m) \, .$$

Ha az $x_0 \in \text{int}\mathbb{R}^n_+$ a 6.15 feladat megoldása, és a feladat függvényeire fennálnak a Lagrange-féle multiplikátor-tétel feltételei, akkor

$$u'(x_0) + \lambda \cdot p = 0,$$

azaz $\forall i = 1, ..., n$ esetén

$$\lambda = -\frac{D_i u(x_0)}{p_i} \, ,$$

amely szerint az optimális pontban minden termék parciális határhasznának és az árának az aránya megegyezik.

Továbbá ebből adódik az is, hogy $\forall i, j = 1, ..., n$ esetén

$$\frac{D_i u(x_0)}{D_j u(x_0)} = \frac{p_i}{p_j} \, ,$$

amely szerint az optimális pontban bármely két termék határhasznának az aránya megegyezik az árak arányával.

A továbbiakban kövessük a fenti 6.8.5 Megjegyzés menetét:

Legyen $i, j = 1, 2, ..., n$ tetszőleges, és tekintsük csupán az i és j-dik terméket, azaz a többi termék mennyiségét rögzítsük valamely adott szinten, ekkor $u : \mathbb{R}^2_+ \to \mathbb{R}$. Az egyszerűség kedvéért legyen $i = 1, j = 2$. A feladat ekkor az előző (6.15) feladatnak a következő speciális esete:

$$\begin{aligned} &u(x_1, x_2) \to \max \\ &p_1 x_1 + p_2 x_2 = m \, . \end{aligned} \tag{6.16}$$

Legyen $(x_1^0, x_2^0) \in \text{int}\mathbb{R}^2_+$ megoldás. Az előbb láttuk, hogy ebben a pontban

$$\frac{D_1 u(x_1^0, x_2^0)}{D_2 u(x_1^0, x_2^0)} = \frac{p_1}{p_2} \, .$$

Tegyük fel, hogy az $u : \mathbb{R}^2_+ \to \mathbb{R}$ függvényre fennállnak az implicitfüggvény-tétel feltételei, ekkor (mint a 6.7.6 Példában már láttuk,) az $u^{-1}(u(x_1^0, x_2^0)) \subset \mathbb{R}^2_+$ közömbösségi görbe az x_1^0 egy környezetében differenciálható függvény, azaz $\exists!\ g_0$ differenciálható függvény, hogy $g_0(x_1^0) = x_2^0$, az x_1^0 egy környezetében $u(x_1, g_0(x_1)) = u(x_1^0, x_2^0)$, valamint

$$g_0'(x_1^0) = -\frac{D_1u(x_1^0, x_2^0)}{D_2u(x_1^0, x_2^0)} .$$

Továbbá (most ebben a speciális esetben az implicitfüggvény-tétel alkalmazása nélkül is) látható, hogy $f_1 : \mathbb{R}^2_+ \to \mathbb{R},\ f_1(x_1, x_2) = p_1x_1 + p_2x_2$ költségvetési függvényre az $f_1^{-1}(m) = g_1 : \mathbb{R} \to \mathbb{R}$ (affin) függvény, amelyre $g_1(x_1) = -\frac{p_1}{p_2}x_1 + \frac{m}{p_2}$, így

$$g_1'(x_1) = -\frac{p_1}{p_2} .$$

Mivel a Lagrange-féle multiplikátor-tétel szerint $\frac{D_1u(x_1^0,x_2^0)}{D_2u(x_1^0,x_2^0)} = \frac{p_1}{p_2}$, azért

$$g_0'(x_1^0) = -\frac{D_1u(x_1^0, x_2^0)}{D_2u(x_1^0, x_2^0)} = -\frac{p_1}{p_2} = g_1'(x_1^0) .$$

A mikroökonómiában szokásos jelölésekkel leírva:

$$g_0'(x_1^0)(= x_2'(x_1^0) = \frac{dx_2}{dx_1}) = -\frac{p_1}{p_2} ,$$

azaz megkaptuk a mikroökonómia egy sarkalatos törvényét, amely szerint az optimális pontban a 2. terméknek az 1. termékre vonatkozó helyettesítési határrátája megegyezik árarányaik reciprokának az ellentettjével.

Mivel $g_0(x_1^0) = x_2^0 = g_1(x_1^0)$, azért ez azt jelenti, hogy az optimális pontban a hasznossági függvény közömbösségi görbéje érinti a költségvetési egyenest. Így szemléletesen az optimális pontot úgy "kapjuk meg", hogy a hasznossági függvény közömbösségi görbéit addig toljuk, amíg nem érinti a költségvetési egyenest.

2. A Hicks-féle megközelítés: A fogyasztó adott hasznossági szint mellett a költségét (kiadását) minimalizálja. Ez mondható a fenti megközelítés duálisának.

Legyen $\hat{u} \in \mathbb{R}$ adott hasznossági szint. A feladat:

$$\begin{aligned} &u(x) \to \max \\ &\langle p, x\rangle = \hat{u} . \end{aligned} \tag{6.17}$$

A feladat Lagrange-függvénye az az $\mathcal{L} : \mathbb{R}^n_+ \times \mathbb{R} \to \mathbb{R}$ függvény, amelyre $\forall\ x \in \mathbb{R}^n_+$, $\forall\ \mu \in \mathbb{R}$ esetén

$$\mathcal{L}(x, y) = \langle p, x\rangle + \mu \cdot (u(x) - \hat{u}) .$$

Ha az $x_0 \in \mathrm{int}\mathbb{R}^n_+$ a 6.17 feladat megoldása, és a feladat függvényeire fennálnak a Lagrange-féle multiplikátor-tétel feltételei, akkor

$$p + \mu \cdot u'(x_0) = 0,$$

azaz $\forall i = 1, ..., n$ esetén

$$\mu = -\frac{p_i}{D_iu(x_0)} ,$$

azaz a (6.15) feladat Lagrange szorzójának a reciproka. A levonható következtetédek így ugyanazok. Például végső következtetésként azt kapjuk, hogy az optimális pontban a költségvetési egyenes érinti a hasznossági függvény közömbösségi görbéjét. Azonban most ez azt jelenti, hogy az optimális pontot úgy "kapjuk meg", hogy a költségvetési egyenest addig toljuk, amíg nem érinti a hasznossági függvény közömbösségi görbéjét.

II. A termelői viselkedés A költségminimalizálási feladat

Legyen a termelő termelési függvénye $f : \mathbb{R}^n_+ \to \mathbb{R}$, legyen $p \in \mathbb{R}^n_+$ az árak vektora, legyen $y \in \mathbb{R}$ adott termelési szint. A feladat:

$$\begin{aligned} f(x) &\to \max \\ \langle p, x\rangle &= y\ . \end{aligned} \tag{6.18}$$

Ez formálisan ugyanaz a feladat, mint a (6.17). A feladat Lagrange-függvénye az az $\mathcal{L} : \mathbb{R}^n_+ \times \mathbb{R} \to \mathbb{R}$ függvény, amelyre $\forall\ x \in \mathbb{R}^n_+$, $\forall\ \mu \in \mathbb{R}$ esetén

$$\mathcal{L}(x, y) = \langle p, x\rangle + \mu \cdot (f(x) - y)\ .$$

Ha az $x_0 \in \text{int}\mathbb{R}^n$ a 6.18 feladat megoldása, és a feladat függvényeire fennálnak a Lagrange-féle multiplikátor-tétel feltételei, akkor

$$p + \mu \cdot f'(x_0) = 0,$$

azaz $\forall i = 1, ..., n$ esetén

$$\mu = -\frac{p_i}{D_i(x_0)}\ .$$

A levonható következtetések ugyanazok, mint az előző feladatok esetében. Egyrészt az optimális pontban minden minden termelési tényező parciális határtermékének és az árának az aránya megegyezik. Másrészt ebből adódik az is, hogy $\forall i, j = 1, ..., n$ esetén $\frac{D_i f(x_0)}{D_j(x_0)} = \frac{p_i}{p_j}$, amely szerint bármely két termelési tényező parciális határtermékének az aránya megegyezik az áraik arányával. Végül csak két terméket vizsgálva, ha az f függvényre teljesülnek az implicitfüggvény-tétel feltételei, akkor $g_0'(x_1^0)(= x_2'(x_1^0) = \frac{dx_2}{dx_1}) = -\frac{p_1}{p_2}$, azaz az optimális pontban a 2. terméknek az 1. termékre vonatkozó helyettesítési határrátája megegyezik árarányaik reciprokának az ellentettjével, amely szerint az optimális pontban a költségvetési egyenes érinti a hasznossági függvény közömbösségi görbéjét.

6.8.10 Megjegyzés. Meglepőnek tűnhet, hogy azt a roppant egyszerű rajzolgatást, ami a mikroökonómia egyik kiindulópontja, csak két félév matematika tanulás után lehet elmagyarázni (sőt a magyarázatunk egy kicsit hiányos is abban az értelemben, hogy mind az implicitfüggvény-tételt, mind a Lagrange-féle multiplikátor-tételt csak a speciális kétdimenziós esetben bizonyítottuk be). Vegyük észre azonban, hogy olyan egyszerűnek tűnő és szemléletes fogalomnak, mint a középiskolából jól ismert számegyenesnek a gondos bevezetése az analízis egyik legmélyebb területe, továbbá a szintén nagyon egyszerűnek tűnő, már az általános iskolában megismert érintő fogalmának, azaz a deriváltnak – amely az analízis egyik központi fogalma – a bevezetése rengeteg munkát igényel, különösen a többváltozós esetben.

1. Gyakorlatok, feladatok

1. Mutassuk meg, hogy az $\mathbf{R}^p$ tér egységgömbjének felszíne kompakt halmaz, és az $x \to |Ax|$ leképezés folytonos ezen a halmazon bármely $q \times p$ méretű A mátrix esetén.
2. Bizonyítsuk be, hogy a (6.1) és (6.2) egyenlőségek valóban normát definiálnak a lineáris leképezések terén.
3. Igazoljuk a norma következő karakterizációját:
$$\|A\| = \inf\{\lambda > 0 : |Ax| \le \lambda |x| \quad \forall x \in X\} .$$
4. Igaz marad-e a 6.1.4 Állítás, ha a normát a (6.1), vagy (6.2) alatti normák valamelyikére cseréljük?
5. Mutassuk meg, hogy minden ortogonális mátrix egységnyi normájú.
6. Bizonyítsuk be, hogy normális mátrixok esetében a norma megegyezik a sajátértékek abszolút értékeinek maximumával, azaz
$$\|A\| = \max\{|\lambda| : \lambda \text{ az } A \text{ sajátértéke}\} .$$
7. Jelölje λ az A mátrix domináns sajátértékét. Igazoljuk, hogy $|\lambda| \le \|A\|$, és az egyenlőség pontosan akkor teljesül, ha A diagonalizálható. (Vesd össze a 6. gyakorlattal.)
8. Közvetlenül a definíció alapján ellenőrizzük, hogy ha az $f : X \to Y$ függvény differenciálható az $x \in X$ pontban, akkor ott folytonos is.
9. Mutassuk meg, hogy ha az f koordinátafüggvényei $f_1, \ldots, f_q$, úgy f akkor és csak akkor differenciálható az x pontban, ha itt mindegyik koordinátafüggvénye is differenciálható.
10. Mutassuk meg, hogy a $g(t) = f(x + tv)$ differenciálhatósága bármely v mellett a 0 pontban nem feltétlenül jelenti az f differenciálhatóságát az x pontban. Tekintsük ugyanis az
$$f(x, y) = \begin{cases} 1 & \text{ha } y = x^2, \quad (x, y) \neq (0, 0) \\ 0 & \text{különben} \end{cases}$$
függvényt. Igazoljuk, hogy ekkor bármely $v \in \mathbb{R}^2$ mellett a $g(t) = f(tv)$ függvény differenciálható a 0 pontban, és $g'(0) = 0$. Azonban f még csak nem is folytonos az origóban, hiszen annak bármely környezetében egyaránt felveszi a 0 és az 1 értékeket is.
11. Határozzuk meg az
$$f(x, y) = \frac{1}{x} + \frac{2}{y} + 8xy$$
függvény szélsőértékeit.

12. Mutassuk meg, hogy a 6.6.4 Példában az $(1, 1, 1)$ pont kivételével egyetlen más kritikus pont sem szélsőérték.

13. Tekintsük az X valós euklideszi téren az $f(x) = \|x\|$ függvényt. Vizsgáljuk meg, hogy f milyen pontokban differenciálható, és adjuk meg a deriváltját.

14. Határozzuk meg az $f(x, y, z) = xyz$ függvénynek a maximumát az egységgömbből az $x + y + z = 0$ sík által kimetszett halmazon. Ebben az esetben a feltételt az
$$F(x, y, z) = \begin{bmatrix} x^2 + y^2 + z^2 - 1 \\ x + y + z \end{bmatrix}$$
függvény határozza meg.

15. Határozzuk meg az
$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$
mátrix normáját feltételes szélsőérték feladatként. Legyen $f(x) = \|Ax\|^2$, és $F(x) = \|x\|^2 - 1$, ahol $x \in \mathbb{R}^2$, és keressük meg a megfelelő feladat megoldását.

16. Az előző feladat gondolatmenetét felhasználva bizonyítsuk be a 6.1.5 Állítást a 6.8.3 Tétel segítségével. A normát az
$$\begin{aligned} &\|Ax\|^2 \to \max \\ &\|x\|^2 = 1 \end{aligned}$$
feltételes szélsőérték feladat megoldásaként állítsuk elő.

17. Irjunk egy adott körbe olyan háromszöget, amely oldalainak négyzetösszege maximális. Oldjuk meg a problémát feltételes szélsőérték feladatként.