

Lineáris algebra

ELTE-IK Programtervező Informatikus BSc, I. évfolyam

Szántó Ádám

2013-2014 őszi félév

Tartalomjegyzék

1. Logikai alapok és bizonyítások	3
1.1. Logika	3
1.2. Bizonyítások	4
1.2.1. Direkt bizonyítás	4
1.2.2. Indirekt bizonyítás	5
1.2.3. Teljes indukció	5
1.2.4. Halmazok egyenlőségének bizonyítása	6
1.2.5. Ekvivalencia bizonyítása	7
2. Vektorterek	8
2.1. Vektorok értelmezése	8
2.2. Alterek és bázisok	10
2.3. Az elemi bázistranszformáció	14
2.4. Kifeszített alterek, generátorrendszerek, rang	16
2.5. Lineáris egyenletrendszerek 1.	22
3. Mátrixok	27
3.1. Értelmezés és műveletek	27
3.2. Mátrixok szorzása	28
3.3. Particionálás és alkalmazásai	31
3.4. Rang és inverz	35
3.5. Rangtartó átalakítások és alkalmazásaik	38
3.6. Nevezetes mátrixok	39
3.7. Komplex számtest feletti mátrixok	40
4. Geometriai vektorok	41

5. Determináns	46
5.1. Értelmezés	46
5.2. Tulajdonságok	48
5.3. A determináns értékének kiszámítása elemi bázistranszformációval	51
5.4. Aldetermináns	53
5.5. Kapcsolat a lineáris egyenletrendszerekkel	57
6. Sajátérték és sajátvektor	59
6.1. Diagonalizálhatóság	59
6.2. Karakterisztikus polinom	62
7. Euklideszi terek	64
7.1. Bevezetés	64
7.2. Euklideszi norma	65
7.3. Ortonormált bázisok	67
7.4. Kvadratikus alak, definit	69
7.5. Saját ortonormált bázis keresése	71
8. Lineáris transzformációk	73
8.1. Vektortérhomomorfizmus	73
8.2. Képtér és magtér	75
8.3. Diagonalizálhatóság, sajátérték, sajátvektor	76
8.4. Általánosítás tetszőleges vektorterekre	78
9. Bilineáris függvények	84

1. fejezet

Logikai alapok és bizonyítások

A lineáris algebra tárgyalása során szinte minden állításunkat maradéktalanul bizonyítjuk is, alátámasztva ezzel a hitelességet. Elengedhetetlen tehát a matematikai logika alapszintű ismerete.

1.1. Logika

1. Definíció. *A predikátum a matematikában egy ún „definiálatlan alapfogalom”, melynek változói értékétől függően az értéke igaz, vagy hamis.*

Predikátum akár 0 változós is lehet. Nézzünk néhány nagyon egyszerű példát. Jelölje $E(x)$ azt, hogy x egy egyenes, $P(x)$ azt, hogy x egy pont, $I(x, y)$ pedig azt, hogy x pont, y egyenes és x illeszkedik y -ra. Látjuk, hogy a változók helyére konkrét dolgokat írva egyértelműen eldönthető azok igazságértéke. A predikátumok jelölése nagyon sokféle lehet, például az $=$ jel is felfogható egy kétváltozós predikátumnak, melyet általában a változók közé, zárójelek nélkül írjuk.

2. Definíció. *A logikai formulák az adott elmélet predikátumaiból épülnek fel ún. logikai műveletek segítségével, melyek közül a legtöbbet használtak a negáció (\neg , $\neg A$ pontosan akkor igaz, ha A hamis), konjunkció (\wedge , $A \wedge B$ pontosan akkor igaz, ha A és B is igaz), diszjunkció (\vee , $A \vee B$ pontosan akkor igaz, ha A vagy B igaz), implikáció (\implies , $A \implies B$ pontosan akkor igaz, ha $\neg A \vee B$ igaz), ekvivalencia (\iff , $A \iff B$ pontosan akkor igaz, ha $A \implies B$ és $B \implies A$ igaz). Ez utóbbit úgy is mondjuk, hogy A akkor és csak akkor, ha B . A logikai műveletek mellett használunk még ún. kvantorokat, melyek a \exists (egzisztenciális kvantor), és a \forall (univerzális kvantor). Ha A logikai formula, akkor $\exists x A$ is az, melyet úgy*

értelmezzük, hogy „létezik olyan x , melyre \mathcal{A} teljesül”. Továbbá, ha \mathcal{A} logikai formula, akkor $\forall x \mathcal{A}$ is az, melyet úgy értelmezzük, hogy „minden x esetén \mathcal{A} teljesül”.

Látjuk, hogy ha a predikátumokat logikai állításnak tekintjük, akkor a fenti műveletek véges sokszori alkalmazásával kapjuk az összes állítást. Nézzünk néhány példát: $P(x)$ jelentse azt, hogy x páros, $E(x)$ azt, hogy x egész. Nyilván $P(2)$, $P(14)$, $E(2)$ mind igazak, de pl. $P(3)$, $E(\pi)$ hamisak. Nyilván $P(2) \wedge E(2)$ is igaz. Igaz továbbá $\exists x \in \mathbb{R} : E(x)$ is, hiszen létezik olyan valós szám, mely egész is egyben, azonban $\forall x \in \mathbb{R} : E(x)$ már nem igaz, hiszen nem minden valós szám egész.

Sokszor szükségünk lesz egy állítás tagadására. A $\neg\neg\mathcal{A} = \mathcal{A}$ esete világos. A konjunkció és diszjunkció tagadása: $\neg(\mathcal{A} \wedge \mathcal{B}) = \neg\mathcal{A} \vee \neg\mathcal{B}$, $\neg(\mathcal{A} \vee \mathcal{B}) = \neg\mathcal{A} \wedge \neg\mathcal{B}$, ezekkel az implikáció és ekvivalencia is meghatározható, hisz ezekkel a műveletekkel értelmeztük azokat. A kvantoros állítások tagadása: $\neg(\exists x \mathcal{A}) = \forall x(\neg\mathcal{A})$, illetve $\neg(\forall x \mathcal{A}) = \exists x(\neg\mathcal{A})$. Ez szépen alkalmazható egymásba ágyazott esetekben is, tehát pl.

$$\neg(\forall x : \exists y : \forall z : \mathcal{A}) = \exists x : \forall y : \exists z : \neg\mathcal{A}.$$

Szokás még használni a $\exists!$ kvantort is, melynek jelentése az egyértelmű létezés, tehát pl. igaz a formula, mely szerint $\exists! x \in \mathbb{R} : 2^2 = x$, ez az érték egyértelműen 4.

1.2. Bizonyítások

Többféle bizonyítási módszer is létezik, mindegyiknek az a lényege, hogy az állításunk igaz vagy hamis voltát belássuk egyértelmű logikai következtetések segítségével.

1.2.1. Direkt bizonyítás

A legegyszerűbb módszer, lényege, hogy ha \mathcal{X} állítás igaz voltát szeretnénk belátni, akkor elindulunk valamely már ismert igaz állításból, majd következtetések segítségével próbálunk \mathcal{X} -hez eljutni. Nézzünk egy példát.

1. Tétel (Háromszög-egyenlőtlenség). $\forall a, b \in \mathbb{R} : |a + b| \leq |a| + |b|$.

Bizonyítás. Az abszolút érték fogalmát ismerve tudjuk, hogy $-|a| \leq a \leq |a|$ és $-|b| \leq b \leq |b|$ triviálisan teljesül. A két egyenletet összeadva

$$-(|a| + |b|) \leq a + b \leq |a| + |b|$$

adódik. Mivel minden x, y valós szám esetén $|x| \leq y$ egyenértékű azzal, hogy $-y \leq x \leq y$, ezért ennek felhasználásával az előbbiből azt kapjuk, hogy $|a + b| \leq |a| + |b|$, ami éppen a bizonyítandó állítás. \square

A tétel azért teljesül minden valós a, b esetén, mert a bizonyítás során ezekre semmilyen kikötést nem tettünk.

1.2.2. Indirekt bizonyítás

Ennek a típusnak a lényege, hogy amennyiben \mathcal{A} igaz voltát akarjuk belátni először feltezzük, hogy $\neg \mathcal{A}$ igaz, majd ezen feltétel felhasználásával direkt úton indulunk el. Ha el-
lentmondásra (lehetetlen állításra) jutunk, akkor az eredeti \mathcal{A} állításunk igaz volt.

2. Tétel. $\sqrt{2}$ irracionális, tehát nem írható fel egész számok hányadosaként.

Bizonyítás. Indirekt úton tegyük fel az állításunk ellentettjét, tehát hogy a $\sqrt{2}$ racionális. Ha $\sqrt{2}$ racionális, léteznek olyan $a, b \in \mathbb{Z}, b \neq 0$ egészek, mellyel $\sqrt{2} = a/b$. A törték bővíthetősége végett ez a felírás persze nem egyértelmű, ezért tegyük fel, hogy a/b a leegyszerűsített tört, tehát hogy a és b legnagyobb közös osztója 1 (ezt úgy is mondjuk, hogy a és b relatív prímek). Négyzetre emelve és átalakítva

$$\sqrt{2} = \frac{a}{b} \implies 2 = \frac{a^2}{b^2} \implies 2b^2 = a^2.$$

Azt a következtetést vonhatjuk le, hogy a^2 páros szám, hisz 2-nek többszöröse, innen pedig következik, hogy a is páros (könnyen belátható). Ha páros, akkor létezik egy olyan $c \in \mathbb{Z}$ szám, mellyel $a = 2c$. Helyettesítsük be ezt a fentibe:

$$2b^2 = (2c)^2 = 4c^2 \implies b^2 = 2c^2,$$

innen hasonló gondolatmenettel adódik, hogy b is páros. Azt kaptuk tehát, hogy a és b is párosak, ez viszont ellentmond annak, hogy relatív prímek, hisz ekkor a legnagyobb közös osztójuk legalább 2. Ezzel az ellentmondással adódik, hogy az eredeti állításunk, miszerint a $\sqrt{2}$ irracionális, igaz. \square

1.2.3. Teljes indukció

Ez az egyik legérdekesebb bizonyítási típus. Azokban az esetekben alkalmazható, amikor egy állítást minden természetes számra szeretnénk belátni. Az elv a következő: tegyük fel, hogy valamely n természetes számra adott egy $\mathcal{A}(n)$ állítás. Ha tudjuk, hogy $\mathcal{A}(1)$ igaz, illetve ha

feltesszük hogy valamely n -re $\mathcal{A}(n)$ igaz és ebben az esetben $\mathcal{A}(n+1)$ is igaz, akkor $\mathcal{A}(n)$ minden n természetes számra igaz.

3. Tétel. *Mutassuk meg, hogy minden $n \in \mathbb{N}$ -re fennáll, hogy*

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n \cdot (n+1)} = \frac{n}{n+1}.$$

Bizonyítás. Először az $n = 1$ esetet vizsgáljuk, hisz ha erre nem teljesül, akkor nyilván az sem hogy minden természetes számra teljesül. $n = 1$ esetén

$$\frac{1}{1 \cdot 2} = \frac{1}{2}$$

adódik, ami nyilván igaz. Most következik az ún. indukciós feltevés, feltesszük tehát, hogy valamely $n \in \mathbb{N}$ -re az állítás teljesül, tehát most az

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n \cdot (n+1)} = \frac{n}{n+1}$$

egyenlőséget feltesszük, és ennek segítségével próbáljuk igazolni az állítást $n+1$ -re, tehát a bizonyítandó:

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n \cdot (n+1)} + \frac{1}{(n+1) \cdot (n+2)} = \frac{n+1}{n+2}.$$

Az indukciós feltétel miatt a fenti összeg első n tagja átírható:

$$\frac{n}{n+1} + \frac{1}{(n+1) \cdot (n+2)} = \frac{n+1}{n+2}.$$

A bal oldalon a közös nevező $(n+1) \cdot (n+2)$, így

$$\frac{n \cdot (n+2) + 1}{(n+1) \cdot (n+2)} = \frac{n^2 + 2n + 1}{(n+1) \cdot (n+2)}$$

Mivel $n^2 + 2n + 1 = (n+1)^2$, így

$$\frac{n^2 + 2n + 1}{(n+1) \cdot (n+2)} = \frac{(n+1)^2}{(n+1) \cdot (n+2)} = \frac{n+1}{n+2},$$

és éppen ezt akartuk belátni, így tehát az eredeti állításunk minden n természetes számra igaz. □

1.2.4. Halmazok egyenlőségének bizonyítása

Azoknál az állításoknál, ahogy egy A és B halmaz egyenlőségét akarjuk belátni, legtöbbször a kétoldali tartalmazást látjuk be külön-külön, tehát $A \subseteq B$ és $B \subseteq A$ állításokat, mert ezek együttes teljesülése akkor és csak akkor áll fenn, ha $A = B$. Erre a generátorrendszerekről szóló részben találunk példát (ahol azt bizonyítjuk, hogy $\text{Span}(A) = W(A)$.)

1.2.5. Ekvivalencia bizonyítása

Az $\mathcal{A} \iff \mathcal{B}$ állításokat vagy ekvivalens átalakításokkal, vagy a kétirányú implikáció belátásával bizonyítunk, tehát külön-külön a $\mathcal{A} \implies \mathcal{B}$ és $\mathcal{B} \implies \mathcal{A}$ irányokkal. A jegyzetben erre is számos példát találunk, pl. az elemi bázistranszformáció bizonyításánál.

2. fejezet

Vektorterek

Középiskolában kaptunk egy képet a vektorokról. Megtanultuk hogy két fontos tulajdonsággal rendelkeznek: van hosszuk és irányuk. Többnyire csak a síkbeli esetekkel foglalkoztunk, összeadtuk őket, valós számmal szoroztuk és egy ún. skaláris szorzatot is értelmezni tudtunk. Most ennél tovább megyünk, és kellő részletességgel tárgyaljuk ezt a témát.

2.1. Vektorok értelmezése

Egy síkvektor nem más, mint egy valós számpár, melyben a számjegyek sorrendje mérvadó. Teljesen általános módon értelmezhetőek a térvektorok, ott valós számhármassokról beszélünk. Az összes síkvektor halmazát \mathbb{R}^2 jelöli, míg a térvektorokét \mathbb{R}^3 . Természetes módon definiáljuk két térvektor összegét, és skalárszorosát:

3. Definíció. Legyenek $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ a következő alakúak:

$$\mathbf{a} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

Legyen továbbá $\lambda \in \mathbb{R}$. Ekkor

$$\mathbf{a} + \mathbf{b} := \begin{bmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \alpha_3 + \beta_3 \end{bmatrix}, \quad \lambda \mathbf{a} := \begin{bmatrix} \lambda \alpha_1 \\ \lambda \alpha_2 \\ \lambda \alpha_3 \end{bmatrix}$$

A vektorokat vastagon szedett kisbetűkkel fogjuk jelölni, a vektor komponenseit pedig görög betűkkel. A fenti definíció természetes módon átvihető kétkomponensű, síkvektorokra is. A fenti definíciót most általánosítsuk:

4. Definíció. Legyen $1 \leq n \in \mathbb{N}$. Legyenek $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ a következő alakúak:

$$\mathbf{a} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

Legyen továbbá $\lambda \in \mathbb{R}$. Ekkor

$$\mathbf{a} + \mathbf{b} := \begin{bmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_n + \beta_n \end{bmatrix}, \quad \lambda \mathbf{a} := \begin{bmatrix} \lambda \alpha_1 \\ \lambda \alpha_2 \\ \vdots \\ \lambda \alpha_n \end{bmatrix}$$

Nyilván nem tudjuk vizualizálni az $n > 3$ eseteket úgy, mint a sík-, és térbeli vektorokat, ez ne is zavarjon minket, a cél az általánosítás. Vizsgáljuk meg a fent bevezetett összeadást algebrai szempontok szerint. Asszociativitást, illetve kommutativitást vizsgálunk:

$$\begin{aligned} \left(\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \right) + \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} &= \begin{bmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_n + \beta_n \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} = \begin{bmatrix} (\alpha_1 + \beta_1) + \gamma_1 \\ (\alpha_2 + \beta_2) + \gamma_2 \\ \vdots \\ (\alpha_n + \beta_n) + \gamma_n \end{bmatrix} = \\ &= \begin{bmatrix} \alpha_1 + (\beta_1 + \gamma_1) \\ \alpha_2 + (\beta_2 + \gamma_2) \\ \vdots \\ \alpha_n + (\beta_n + \gamma_n) \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} \beta_1 + \gamma_1 \\ \beta_2 + \gamma_2 \\ \vdots \\ \beta_n + \gamma_n \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \left(\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} \right) \end{aligned}$$

Szépen látszik, hogy az asszociativitás visszavezethető a valós számok asszociativitására. A kommutativitás teljesen hasonló módon mutatható meg.

5. Definíció. A nullvektort jelölje

$$\mathbf{0} := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n$$

Világos, hogy tetszőleges $\mathbf{a} \in \mathbb{R}^n$ -re $\mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}$. Észrevehetjük még, hogy ha egy vektort (-1) -gyel szorzunk, ellentetthez jutunk, tehát $\mathbf{a} + (-1)\mathbf{a} = \mathbf{0}$. Egyszerűbb $(-1)\mathbf{a}$ helyett annyit írunk, hogy $-\mathbf{a}$.

6. Definíció. Az \mathbb{R}^n -beli összeadás és skalárral való szorzás tulajdonságai:

1. Az összeadás művelet két vektorhoz rendel egy harmadik vektort, formálisan:

$$+ : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

Minden $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ -hez hozzá van rendelve egy $\mathbf{a} + \mathbf{b} \in \mathbb{R}^n$.

2. Kommutativitás: $\forall \mathbf{a}, \mathbf{b} \in \mathbb{R}^n : \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$.

3. Asszociativitás: $\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n : (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$.

4. Létezik nullvektor: $\exists \mathbf{0} \in \mathbb{R}^n : \forall \mathbf{a} \in \mathbb{R}^n : \mathbf{0} + \mathbf{a} = \mathbf{a}$.

5. Létezik ellentett: $\forall \mathbf{a} \in \mathbb{R}^n : \exists (-\mathbf{a}) \in \mathbb{R}^n : (-\mathbf{a}) + \mathbf{a} = \mathbf{0}$.

6. A skalárral való szorzás művelete egy valós számhoz és egy vektorhoz rendel egy vektort, formálisan:

$$(\lambda \cdot) : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

Minden $\lambda \in \mathbb{R}, \mathbf{a} \in \mathbb{R}^n$ -hez hozzá van rendelve egy $\lambda \mathbf{a} \in \mathbb{R}^n$.

7. Asszociativitás: $\forall \lambda, \mu \in \mathbb{R}, \mathbf{a} \in \mathbb{R}^n : (\lambda \mu) \mathbf{a} = \lambda(\mu \mathbf{a})$.

8. Első disztributivitás: $\forall \lambda, \mu \in \mathbb{R}, \mathbf{a} \in \mathbb{R}^n : (\lambda + \mu) \mathbf{a} = \lambda \mathbf{a} + \mu \mathbf{a}$.

9. Második disztributivitás: $\forall \lambda \in \mathbb{R}, \mathbf{a}, \mathbf{b} \in \mathbb{R}^n : \lambda(\mathbf{a} + \mathbf{b}) = \lambda \mathbf{a} + \lambda \mathbf{b}$.

10. $\forall \mathbf{a} \in \mathbb{R}^n : 1 \mathbf{a} = \mathbf{a}$.

Azt, hogy a fenti tulajdonságokkal \mathbb{R}^n rendelkezik, úgy nevezzük, hogy \mathbb{R}^n vektortér (vagy lineáris tér) \mathbb{R} felett.

2.2. Alterek és bázisok

Nem nehéz belátni, hogy \mathbb{R}^n rendelkezik a fenti tulajdonságokkal, viszont egy $W \subset \mathbb{R}^n$ részhalmazról eldönteni ezt, már nem feltétlen olyan egyszerű. Mit értünk pontosan az alatt, hogy egy részhalmaz olyan tulajdonságokkal rendelkezik, melyet eredetileg a teljes \mathbb{R}^n -re vezettünk be? Tegyük ezt rendbe:

7. Definíció. Legyen $\emptyset \neq W \subset \mathbb{R}^n$. Azt mondjuk W altere \mathbb{R}^n -nek, ha

$\forall \mathbf{a}, \mathbf{b} \in W : \mathbf{a} + \mathbf{b} \in W$, illetve $\forall \lambda \in \mathbb{R}, \mathbf{a} \in W : \lambda \mathbf{a} \in W$. Jelölésben: $W \leq \mathbb{R}^n$.

Mit is mond a definíció? Tetszőleges két vektort kiválasztva W -ből ezek összege is W -beli, illetve a skalárszorosra hasonlóan. Ezt úgy is mondjuk, hogy W „zárt” a fenti műveletekre. Később részletesebben is megismerkedünk az alterekkel, előbb bevezetünk néhány alapfogalmat.

8. Definíció. Legyen $k \geq 1$, $\mathbf{a}_1, \dots, \mathbf{a}_k \in \mathbb{R}^n$ egy vektorrendszer (tehát a vektorok között lehetnek egyenlők, akár mind egyenlő lehet). Legyenek továbbá $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ skalárok. Az $\mathbf{a}_1, \dots, \mathbf{a}_k$ vektorrendszer $\lambda_1, \dots, \lambda_k$ együtthatós lineáris kombinációja alatt a

$$\sum_{i=1}^k \lambda_i \mathbf{a}_i = \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \dots + \lambda_k \mathbf{a}_k \in \mathbb{R}^n$$

vektort értjük.

Nézzünk egy példát, legyenek

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathbb{R}^2, \quad \mathbf{a}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \in \mathbb{R}^2.$$

Ekkor például

$$2\mathbf{a}_1 + 4\mathbf{a}_2 = 2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 4 \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 12 \\ 16 \end{bmatrix} = \begin{bmatrix} 14 \\ 20 \end{bmatrix} \in \mathbb{R}^2$$

egy lineáris kombináció, ahol $\lambda_1 = 2$, $\lambda_2 = 4$. Triviális lineáris kombinációról akkor beszélünk, ha minden együttható 0. Könnyen láthatóan a triviális lineáris kombináció mindig a nullvektort adja.

Mivel lineáris kombinációkkal egy $\mathbf{a} = \lambda_1 \mathbf{a}_1 + \dots + \lambda_\ell \mathbf{a}_\ell$ alakú egyenletet kapunk, azt látjuk, hogy az $\mathbf{a}_1, \dots, \mathbf{a}_\ell$ vektorrendszer előállítja a \mathbf{a} vektort, magyarul léteznek olyan skalárok, mellyel az egyenlet teljesül. Ebből a megfontolásból ha tekintjük az „összeolvasztott” $\mathbf{a}, \mathbf{a}_1, \dots, \mathbf{a}_\ell$ vektorrendszert, ezt joggal nevezhetjük összefüggőnek, hisz van benne olyan vektor, mely kifejezhető a többi lineáris kombinációjával. A fenti egyenletet átírva

$$\mathbf{0} = \lambda_1 \mathbf{a}_1 + \dots + \lambda_\ell \mathbf{a}_\ell + (-1) \cdot \mathbf{a}$$

adódik, így ebben a felírásban létezik nem nulla együttható (-1) , az előállított vektor pedig a nullvektor, az új rendszer elemszáma pedig $k := \ell + 1$. Ezzel a megfontolással definiáljuk a lineáris összefüggőség fogalmát.

9. Definíció. Az $\mathbf{a}_1, \dots, \mathbf{a}_k \in \mathbb{R}^n$ vektorrendszer lineárisan összefüggő, ha léteznek olyan $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ nem mind 0 skalárok, tehát $\exists 1 \leq i \leq k : \lambda_i \neq 0$, melyre $\lambda_1 \mathbf{a}_1 + \dots + \lambda_k \mathbf{a}_k = \mathbf{0}$. A rendszert lineárisan függetlennek nevezzük, ha nem összefüggő, tehát csak triviális lineáris kombináció ad nullvektort.

Fentebb láttuk, hogy összefüggő esetben van olyan nem nulla együtthatójú vektor a rendszerben, melyet átvihetünk a másik oldalra, tehát a többi vektor előállítja azt lineáris kombinációjuként. Látjuk, hogy független esetben egyik vektor sem állítható elő a többivel. Mutatunk erre egy nagyon egyszerű példát. Legyenek

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n,$$

tehát \mathbf{e}_i -nek pontosan az i -ik koordinátája 1, a többi 0. Megmutatjuk, hogy ez a rendszer független, tehát a $\lambda_1 \mathbf{e}_1 + \dots + \lambda_n \mathbf{e}_n = \mathbf{0}$ egyenletből következik, hogy $\lambda_1 = \dots = \lambda_n = 0$.

$$\lambda_1 \mathbf{e}_1 + \dots + \lambda_n \mathbf{e}_n = \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \lambda_2 \\ \vdots \\ 0 \end{bmatrix} + \dots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix},$$

ez utóbbit pedig ha nullvektorral tesszük egyenlővé azonnal adódik, hogy minden $\lambda_i = 0$. Észrevehetünk még egy nagyon érdekes dolgot: a fenn bevezetett vektorrendszer \mathbb{R}^n összes vektorát előállítja lineáris kombinációval, ugyanis tetszőleges

$$\mathbf{a} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{R}^n$$

esetén $\alpha_1 \mathbf{e}_1 + \dots + \alpha_n \mathbf{e}_n = \mathbf{a}$. Az (első) definíciónkat a bázis fogalmára ez az észrevétel adja alapul (azért első, mert később precízebben is definiáljuk).

10. Definíció. Legyen $V \leq \mathbb{R}^n$ altér és $\mathbf{b}_1, \dots, \mathbf{b}_k \in V$. Azt mondjuk $\mathbf{b}_1, \dots, \mathbf{b}_k$ bázis V -ben, ha lineárisan független vektorrendszer és V minden vektorát előállítja lineáris kombinációként.

Nyilván \mathbb{R}^n is altér \mathbb{R}^n -ben, hisz teljesül rá az altér definíciója. A bázis fogalma viszont általánosabban, tetszőleges altérre definiálható.

Nézzünk egy példát. Tekintsük \mathbb{R}^3 -at mint a geometriai 3-dimenziós teret egy szokásos x, y, z Descartes koordináta rendszerben, ahol x, y, z a koordináta-tengelyeket jelölik. Tekintsük továbbá ebben a térben az x, y tengelyek által meghatározott síkot. Könnyen meggondolható,

hogy ez a sík éppen a

$$W = \left\{ \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ 0 \end{bmatrix} : \alpha_1, \alpha_2 \in \mathbb{R} \right\}$$

halmazzal reprezentálható, tehát azon vektorok halmazával, melyekben a harmadik, z -vel jelölt koordináta nulla. Elsőnek megmutatjuk, hogy ez a sík altér a térben, tehát $W \leq \mathbb{R}^3$. Nyilván W nem üres. Vegyünk két tetszőleges vektort W -ből. Ezek összege szintén egy olyan vektor, melynek utolsó komponense nulla, így az összeg is W -ben van. Nyilván bármely W -beli vektor bármely skalárszorosa is W -ben van, így ez valóban altér. A korábbiakhoz hasonló eszközökkel egyszerűen mutatható meg, hogy pl.

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

egy bázis W -ben.

A bázisokkal kapcsolatosan belátunk két egyszerű állítást.

4. Tétel. Legyen $\mathbf{b}_1, \dots, \mathbf{b}_k$ bázis $V \leq \mathbb{R}^n$ -ben, és $\mathbf{a} \in V$. Ekkor ezzel a bázissal \mathbf{a} egyértelműen állítható elő, tehát $\exists! \alpha_1, \dots, \alpha_k \in \mathbb{R} : \alpha_1 \mathbf{b}_1 + \dots + \alpha_k \mathbf{b}_k = \mathbf{a}$.

Bizonyítás. A létezés a bázis definíciójából adódik, csak az egyértelműséget kell belátni. Tegyük fel, hogy

$$\mathbf{a} = \alpha'_1 \mathbf{b}_1 + \dots + \alpha'_k \mathbf{b}_k = \alpha''_1 \mathbf{b}_1 + \dots + \alpha''_k \mathbf{b}_k$$

Mivel $\mathbf{a} - \mathbf{a} = \mathbf{0}$, így

$$\mathbf{0} = (\alpha'_1 - \alpha''_1) \mathbf{b}_1 + \dots + (\alpha'_k - \alpha''_k) \mathbf{b}_k.$$

Mivel a bázis lineárisan független, így minden zárójelben nullának kell legyen, következésképp minden i -re $\alpha'_i = \alpha''_i$, tehát az egyértelműséget megmutattuk. \square

5. Tétel. Legyen $V \leq \mathbb{R}^n$ és $\mathbf{b}_1, \dots, \mathbf{b}_k \in V$ olyan, hogy $\forall \mathbf{a} \in V : \exists! \alpha_1, \dots, \alpha_k \in \mathbb{R} : \mathbf{a} = \alpha_1 \mathbf{b}_1 + \dots + \alpha_k \mathbf{b}_k$. Ekkor $\mathbf{b}_1, \dots, \mathbf{b}_k$ bázis V -ben.

Bizonyítás. Ebben az esetben a bázis definíciójából csak a függetlenséget kell belátni, hisz az előállítást feltettük. $\mathbf{a} = \mathbf{0}$ választással a feltételek szerint egyértelműen állítjuk elő a nullvektort, és mivel a triviális lineáris kombináció mindig nullvektort ad, az egyértelműség végett szükségképpen csak így lehetséges az előállítás. \square

2.3. Az elemi bázistranszformáció

Mi is történik pontosan, amikor elemi bázistranszformációt hajtunk végre? Tekintsük \mathbb{R}^2 triviális $\mathbf{e}_1, \mathbf{e}_2$ bázisát, illetve legyen

$$\mathbf{x} := \begin{bmatrix} 3 \\ -5 \end{bmatrix} \in \mathbb{R}^2$$

adott vektor. Tudjuk, hogy ekkor $\mathbf{x} = 3\mathbf{e}_1 - 5\mathbf{e}_2$. Mi van akkor, ha az egyik bázisvektort (pl. \mathbf{e}_2 -t) ki akarjuk cserélni egy másik vektorra? Legyen ez a vektor pl.

$$\mathbf{b} := \begin{bmatrix} 2 \\ 4 \end{bmatrix} \in \mathbb{R}^2.$$

Megtehetjük ezt egyáltalán? A kapott \mathbf{e}_1, \mathbf{b} rendszer valóban bázis? Próbáljuk meg \mathbf{x} -et előállítani ebben az új rendszerben, tehát az $\mathbf{x} = y_1\mathbf{e}_1 + y_2\mathbf{b}$ egyenletből keressük az együtthatókat, y_1 -et és y_2 -t.

$$\begin{bmatrix} 3 \\ -5 \end{bmatrix} = y_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y_2 \begin{bmatrix} 2 \\ 4 \end{bmatrix} \implies \begin{cases} y_1 + 2y_2 = 3 \\ 4y_2 = -5. \end{cases} \implies y_2 = -\frac{5}{4}, \quad y_1 = \frac{11}{2}.$$

Ellenőrizzünk!

$$y_1\mathbf{e}_1 + y_2\mathbf{b} = \frac{11}{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{5}{4} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{11}{2} - \frac{5}{2} \\ -5 \end{bmatrix} = \begin{bmatrix} 3 \\ -5 \end{bmatrix} = \mathbf{x},$$

így tehát sikerrel jártunk, egy egyszerű egyenletrendszert kellett megoldanunk. Felmerülnek a kérdések: milyen vektorok „vihetők be a bázisba”? Mi a feltétel erre vonatkozóan? Hogyan állít elő egy tetszőleges vektort az új bázis? Erről szól az elemi bázistranszformáció tétele, és bizonyítása.

Tétel. Legyen $V \leq \mathbb{R}^n$, $\mathbf{b}_1, \dots, \mathbf{b}_k$ bázis V -ben, $\mathbf{a} \in V$, $\mathbf{a} = \alpha_1\mathbf{b}_1 + \dots + \alpha_k\mathbf{b}_k$. Legyen továbbá $1 \leq i \leq k$ rögzített. Ekkor \mathbf{a} akkor és csak akkor vihető be a bázisba \mathbf{b}_i helyére, ha $\alpha_i \neq 0$. Formálisan

$$\mathbf{b}_1, \dots, \mathbf{b}_{i-1}, \mathbf{a}, \mathbf{b}_{i+1}, \dots, \mathbf{b}_k \text{ bázis } V\text{-ben} \iff \alpha_i \neq 0.$$

Bizonyítás. Mindkét irányt bizonyítanunk kell. Tekintsük először a \Rightarrow irányt és indirekt úton tegyük fel, hogy $\alpha_i = 0$, tehát $\mathbf{a} = \alpha_1\mathbf{b}_1 + \dots + 0\mathbf{b}_i + \dots + \alpha_k\mathbf{b}_k$. Innen átrendezéssel kapjuk, hogy $\mathbf{0} = \alpha_1\mathbf{b}_1 + \dots + (-1)\mathbf{a} + \dots + \alpha_k\mathbf{b}_k$ ami pont azt jelenti, hogy

$$\mathbf{b}_1, \dots, \mathbf{b}_{i-1}, \mathbf{a}, \mathbf{b}_{i+1}, \dots, \mathbf{b}_k$$

összefüggő, ami ellentmondás, hisz feltétel volt hogy bázis, így tehát $\alpha_i = 0$. Tekintsük most $\mathbf{a} \Leftarrow$ irányt. Itt kettő dolgot kell bizonyítanunk: a lineáris függetlenséget, illetve hogy az új rendszer is előállít minden vektort V -ben. Az utóbbival kezdjük. Tekintsünk egy tetszőleges $\mathbf{x} \in V$ vektort. Ennek az előállítása az eredeti bázisban

$$\mathbf{x} = x_1 \mathbf{b}_1 + \cdots + x_i \mathbf{b}_i + \cdots + x_k \mathbf{b}_k.$$

Keressük az új bázisbeli előállítást, tehát olyan y_1, \dots, y_k együtthatókat, melyekkel

$$\mathbf{x} = y_1 \mathbf{b}_1 + \cdots + y_{i-1} \mathbf{b}_{i-1} + y_i \mathbf{a} + y_{i+1} \mathbf{b}_{i+1} + \cdots + y_k \mathbf{b}_k.$$

Fejezzük ki \mathbf{a} -t az eredeti bázisbeli előállításával:

$$\begin{aligned} \mathbf{x} &= y_1 \mathbf{b}_1 + \cdots + y_{i-1} \mathbf{b}_{i-1} + y_i (\alpha_1 \mathbf{b}_1 + \cdots + \alpha_k \mathbf{b}_k) + y_{i+1} \mathbf{b}_{i+1} + \cdots + y_k \mathbf{b}_k = \\ &= (y_1 + y_i \alpha_1) \mathbf{b}_1 + \cdots + (y_i \alpha_i) \mathbf{b}_i + \cdots + (y_k + y_i \alpha_k) \mathbf{b}_k \end{aligned}$$

Mivel a bázisbeli előállítás egyértelmű, vessük össze a kapott alakot az eredeti bázisbeli előállítással: minden együtthatónak meg kell egyeznie, így meg is kaphatjuk az új együtthatókat.

Tekintsük y_i -t: $x_i = y_i \alpha_i \implies \boxed{y_i = \frac{x_i}{\alpha_i}}$. Tekintsük most a többi y_t -t ($t \neq i$):

$$x_t = y_t + y_i \alpha_t \implies y_t = x_t - y_i \alpha_t \implies \boxed{y_t = x_t - \frac{x_i}{\alpha_i} \alpha_t}$$

A fenti műveletek elvégzésére az egyetlen feltétel, hogy $\alpha_i \neq 0$, de ezt feltettük. Már csak azt kell bizonyítani, hogy az új rendszer lineárisan független-e. Legyenek $\mu_1, \dots, \mu_k \in \mathbb{R}$ és vizsgáljuk:

$$\mu_1 \mathbf{b}_1 + \cdots + \mu_{i-1} \mathbf{b}_{i-1} + \mu_i \mathbf{a} + \mu_{i+1} \mathbf{b}_{i+1} + \cdots + \mu_k \mathbf{b}_k = \mathbf{0}.$$

Írjuk be \mathbf{a} előállítását:

$$\begin{aligned} \mu_1 \mathbf{b}_1 + \cdots + \mu_{i-1} \mathbf{b}_{i-1} + \mu_i (\alpha_1 \mathbf{b}_1 + \cdots + \alpha_k \mathbf{b}_k) + \mu_{i+1} \mathbf{b}_{i+1} + \cdots + \mu_k \mathbf{b}_k &= \mathbf{0} \implies \\ \implies (\mu_1 + \mu_i \alpha_1) \mathbf{b}_1 + \cdots + (\mu_i \alpha_i) \mathbf{b}_i + \cdots + (\mu_k + \mu_i \alpha_k) \mathbf{b}_k &= \mathbf{0}. \end{aligned}$$

Mivel $\mathbf{b}_1, \dots, \mathbf{b}_k$ független, így minden zárójelben 0 van:

$$t = i : \mu_i \alpha_i = 0 \xrightarrow{\alpha_i \neq 0} \mu_i = 0.$$

$$t \neq i : \mu_t + \mu_i \alpha_t = 0 \implies \mu_t = -\mu_i \alpha_t \stackrel{\mu_i=0}{=} 0,$$

és ezzel készen vagyunk.

Kényelmes lehet táblázatot használni a bázistranszformáció elvégzéséhez:

$$\begin{array}{c|cc|c|cc} & \mathbf{a} & \mathbf{x} & & \mathbf{a} & \mathbf{x} \\ \hline \mathbf{b}_1 & \alpha_1 & x_1 & \mathbf{b}_1 & 0 & x_1 - \frac{x_i}{\alpha_i} \alpha_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{b}_i & \alpha_i & x_i & \mathbf{a} & 1 & \frac{x_i}{\alpha_i} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{b}_k & \alpha_k & x_k & \mathbf{b}_k & 0 & x_k - \frac{x_i}{\alpha_i} \alpha_k \end{array}$$

Most írjuk fel ebben a formában a kezdő példánkat:

$$\begin{array}{c|cc|c|cc} & \mathbf{b} & \mathbf{x} & & \mathbf{b} & \mathbf{x} \\ \hline \mathbf{e}_1 & 2 & 3 & \mathbf{e}_1 & 0 & \frac{11}{2} \\ \mathbf{e}_2 & 4 & -5 & \mathbf{b} & 1 & -\frac{5}{4} \end{array}$$

2.4. Kifeszített alterek, generátorrendszerek, rang

11. Definíció. Az $A, \subseteq \mathbb{R}^n$ jelentse azt, hogy A véges vagy végtelen \mathbb{R}^n -beli vektorrendszer (tehát egy vektor többször is előfordulhat benne).

Bevezetünk egy új fogalmat, melyet lényegében már ismerünk, csak új elnevezést adunk rá, ez pedig nem más, mint a lineáris függés.

12. Definíció. Legyen $\emptyset \neq A, \subseteq \mathbb{R}^n$ és $\mathbf{v} \in \mathbb{R}^n$. Azt mondjuk, \mathbf{v} lineárisan függ A -tól, ha \mathbf{v} előáll A -beli vektorok (véges) lineáris kombinációjaként.

Ezt már ismertük, hisz amikor lineáris kombinációval előállítunk egy vektort, az előállított vektor lineárisan függ azoktól, melyek kombinációival felírtuk. Emlékezzünk vissza a lineáris összefüggés definíciójára. Szóban említettük, hogy ebben az esetben van olyan vektor a rendszerben, melyet „átvihetünk az egyenlet másik oldalára”, tehát létezik olyan vektor, mely a többivel kifejezhető. A lineáris függés fogalmával ezt precízen megfogalmazhatjuk, és bizonyíthatjuk.

6. Tétel. Legyen $2 \leq k \in \mathbb{N}$, $\mathbf{a}_1, \dots, \mathbf{a}_k \in \mathbb{R}^n$. Ekkor az $\mathbf{a}_1, \dots, \mathbf{a}_k$ vektorrendszer akkor és csak akkor lineárisan összefüggő, ha létezik köztük olyan vektor, mely előáll a többi lineáris kombinációjaként. Formálisan:

$$\begin{aligned} & \mathbf{a}_1, \dots, \mathbf{a}_k \text{ lineárisan összefüggő} \iff \\ & \iff \exists \mathbf{a}_i \text{ ami lineárisan függ az } \mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_k \text{ vektorrendszertől.} \end{aligned}$$

Bizonyítás. Nézzük a \Rightarrow irányt. Ha feltettük hogy $\mathbf{a}_1, \dots, \mathbf{a}_k$ összefüggő, akkor tudjuk hogy vannak olyan $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ nem mind nulla számok, mellyel $\lambda_1 \mathbf{a}_1 + \dots + \lambda_k \mathbf{a}_k = \mathbf{0}$. Legyen ezek közül $\lambda_i \neq 0$. Ekkor

$$\lambda_1 \mathbf{a}_1 + \dots + \lambda_{i-1} \mathbf{a}_{i-1} + \lambda_{i+1} \mathbf{a}_{i+1} + \dots + \lambda_k \mathbf{a}_k = -\lambda_i \mathbf{a}_i.$$

Mindkét oldalt $-\lambda_i$ -vel leosztva megkaptuk \mathbf{a}_i előállítását az $\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_k$ vektorok lineáris kombinációjaként.

Nézzük most a \Leftarrow irányt. Most azt feltételezzük, hogy létezik \mathbf{a}_i , mely a többitől lineárisan függ, tehát

$$\mathbf{a}_i = \lambda_1 \mathbf{a}_1 + \dots + \lambda_{i-1} \mathbf{a}_{i-1} + \lambda_{i+1} \mathbf{a}_{i+1} + \dots + \lambda_k \mathbf{a}_k$$

Ha átvisszük \mathbf{a}_i -t a másik oldalra, akkor

$$\mathbf{0} = \lambda_1 \mathbf{a}_1 + \dots + \lambda_{i-1} \mathbf{a}_{i-1} + (-1) \mathbf{a}_i + \lambda_{i+1} \mathbf{a}_{i+1} + \dots + \lambda_k \mathbf{a}_k$$

adódik, mely egy nemtriviális előállítása a nullvektornak (hisz -1 biztosan nem 0), így $\mathbf{a}_1, \dots, \mathbf{a}_k$ lineárisan összefüggő. \square

Belátunk egy másik összefüggést is, mely a függetlenség fogalmával is kapcsolatot teremt.

7. Tétel. *Legyenek $\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{b} \in \mathbb{R}^n$ és tegyük fel, hogy $\mathbf{a}_1, \dots, \mathbf{a}_k$ lineárisan független, de $\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{b}$ már lineárisan összefüggő. Ebben az esetben \mathbf{b} lineárisan függ $\mathbf{a}_1, \dots, \mathbf{a}_k$ -től.*

Bizonyítás. Indirekt úton tegyük fel, hogy a $\lambda_1 \mathbf{a}_1 + \dots + \lambda_k \mathbf{a}_k + \lambda_{k+1} \mathbf{b} = \mathbf{0}$ egyenletben $\lambda_{k+1} = 0$. Tudjuk, hogy a $\lambda_1 \mathbf{a}_1 + \dots + \lambda_k \mathbf{a}_k = \mathbf{0}$ egyenletben minden $\lambda_i = 0$, hisz a rendszer független, ám ha a \mathbf{b} -vel kiegészített rendszer összefüggő, akkor $\lambda_{k+1} \neq 0$ jöhet csak szóba, mert ekkor kell léteznie nem nulla λ -nak. Innen pedig adódik, hogy \mathbf{b} lineárisan függ a többitől, hiszen át tudjuk vinni a másik oldalra. \square

Vezessük be adott vektorrendszer esetén az összes tőle függő vektorok halmazát:

13. Definíció. *Adott $\emptyset \neq A, \subseteq \mathbb{R}^n$ esetén legyen*

$$W(A) := \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} \text{ lineárisan függ } A\text{-től.}\}$$

Most megteremtjük a kapcsolatot az alterek és lineáris kombinációk között, megmutatjuk, hogy a fenti halmaz bármely nemüres vektorrendszer esetén mindig altér, valamint azt, hogy az altér tartalmazza A -t magát. Később meg fogjuk mutatni ráadásul azt is, hogy ez a halmaz a legszűkebb, A -t tartalmazó altér.

8. Tétel. *Legyen $\emptyset \neq A, \subseteq \mathbb{R}^n$ tetszőleges. Ekkor $W(A) \leq \mathbb{R}^n$, valamint $A, \subseteq W(A)$.*

Bizonyítás. Elsőnek belátjuk, hogy $W(A)$ nem üres. Ez elég nyilvánvaló, sőt még az is észrevehető, hogy $\mathbf{a} \in A$ esetén $1\mathbf{a} \in W(A)$, hisz $1\mathbf{a} = \mathbf{a}$ is egy lineáris kombináció, így ezzel $A, \subseteq W(A)$ állítást be is láttuk.

Most azt kell megmutatnunk, hogy bármely két $W(A)$ -beli vektor összege is $W(A)$ -beli. Legyenek $\mathbf{a}, \mathbf{b} \in W(A)$. Ezek nyilván A -beli vektorok lineáris kombinációi, sőt el tudjuk érni, hogy ugyanazoké a vektoroké. Hogyan? Ha egy adott vektor csak az egyik lineáris kombinációban fordul elő, a másikba is beírhatjuk 0 együtthatóval. Így azt kapjuk, hogy

$$\mathbf{a} = \alpha_1 \mathbf{a}_1 + \cdots + \alpha_k \mathbf{a}_k$$

$$\mathbf{b} = \beta_1 \mathbf{a}_1 + \cdots + \beta_k \mathbf{a}_k$$

alakúak, összegük pedig $\mathbf{a} + \mathbf{b} = (\alpha_1 + \beta_1) \mathbf{a}_1 + \cdots + (\alpha_k + \beta_k) \mathbf{a}_k$, ahol minden zárójelben valós szám áll, $\mathbf{a}_1, \dots, \mathbf{a}_k$ pedig A -beliek, így az összeg $W(A)$ -ban van. A skalárszoros is hasonlóan szépen belátható. \square

Egy későbbi még fontosabb tétel kimondásához megmutatjuk, hogy kettő (sőt bármennyi) altér metszete is mindig altér. Unió esetén ez általában nem teljesül.

9. Tétel. $W_1, W_2 \leq \mathbb{R}^n \implies W_1 \cap W_2 \leq \mathbb{R}^n$.

Bizonyítás. A nullvektor minden altérnek eleme, hiszen bármely W altér esetén $\mathbf{a} \in W$ -re $\lambda \mathbf{a} \in W$ és $\lambda = 0$ esete miatt állíthatjuk ezt biztosan. Ezek szerint a nullvektor W_1 és W_2 -nek, így a metszetnek is eleme, tehát a metszet nem üres.

Ha $\mathbf{a}, \mathbf{b} \in W_1 \cap W_2$, akkor $\mathbf{a}, \mathbf{b} \in W_1$ és $\mathbf{a}, \mathbf{b} \in W_2$ is teljesül (hisz a metszetben is benne voltak), és mivel W_1 és W_2 alterek, így az összegvektor benne lesz mindkettőben, így a metszetben is.

Ha $\lambda \in \mathbb{R}$, $\mathbf{a} \in W_1 \cap W_2$, akkor hasonlóan, \mathbf{a} mindkettő altérben benne van. Mivel ezek alterek, így a λ -szoros is benne lesz mindkettőben, így a metszetben is. \square

Bevezetjük a generált (kifeszített) altér fogalmát.

14. Definíció. Legyen $\emptyset \neq A, \subseteq \mathbb{R}^n$ tetszőleges, és tekintsük \mathbb{R}^n összes, A -t tartalmazó alterének metszetét, tehát a

$$\bigcap_{\substack{W \leq \mathbb{R}^n \\ A \subseteq W}} W$$

halmazt. Ezt a halmazt az A által generált (kifeszített) altérnek nevezzük, jelölése $\text{Span}(A)$, vagy $\langle A \rangle$. Ha A véges és $\mathbf{a}_1, \dots, \mathbf{a}_k$ elemekből áll, akkor inkább a $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_k)$ vagy $\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$ jelöléseket használjuk.

Világos (a definícióból), hogy $\text{Span}(A)$ a legszűkebb A -t tartalmazó altér. Megmutatjuk, hogy ez a halmaz éppen $W(A)$.

10. Tétel. $\emptyset \neq A, \subseteq \mathbb{R}^n$ esetén $\text{Span}(A) = W(A)$.

Bizonyítás. Halmazok egyenlőségét úgy szokás bizonyítani, hogy belátjuk mindkét oldali tartalmazást, tehát első körben megmutatjuk, hogy $\text{Span}(A) \subseteq W(A)$. Azt tudjuk, hogy $W(A)$ altér, illetve $A, \subseteq W(A)$, tehát A elemei a $W(A)$ altérből valók, ez utóbbi tehát benne van a metszendők között, a legszűkebb metszet tehát ennél csak kisebb vagy egyenlő lehet mely még mindig tartalmazza A -t, tehát $\text{Span}(A) \subseteq W(A)$. Nézzük most a másik $W(A) \subseteq \text{Span}(A)$ tartalmazást. Legyen $W' \leq \mathbb{R}^n$ egy A -t tartalmazó altér, tehát a metszendők egyike, és legyen $\mathbf{v} \in W(A)$. Ekkor $\mathbf{v} = \alpha_1 \mathbf{a}_1 + \dots + \alpha_k \mathbf{a}_k$ alakú, ahol $\mathbf{a}_1, \dots, \mathbf{a}_k \in A$ az α -k pedig valóságok. Nyilván $\mathbf{a}_1, \dots, \mathbf{a}_k \in W'$ is teljesül, hisz W' tartalmazza magát az A -t. Mivel W' altér, így bármely W' -beli vektorrendszer lineáris kombinációja is benne van W' -ben, tehát $\mathbf{v} \in W'$. Ezzel beláttuk, hogy $W(A) \subseteq W'$. Mivel W' -re semmi kikötésünk nem volt azon kívül, hogy a metszendők egyike, így beláttuk, hogy $W(A)$ a metszendők **bármelyikében** benne van, így a metszetben is. \square

A generált altér tehát éppen az adott vektorrendszer összes lineáris kombinációinak halmaza. Ha a generált altér éppen a teljes vizsgált altér, akkor a rendszerre a generátorrendszer kifejezést fogjuk használni.

15. Definíció. Legyen $V \leq \mathbb{R}^n$. Azt mondjuk, G generátorrendszer V -ben, ha $\emptyset \neq G, \subseteq V$, valamint $\text{Span}(G) = V$.

Íly módon a bázis fogalmát is precízebben megfogalmazhatjuk.

16. Definíció. Bázis alatt lineárisan független generátorrendszert értünk.

11. Tétel. Legyen $\{0\} \neq V \leq \mathbb{R}^n$. Ha V -ben van véges generátorrendszer, akkor létezik bázis V -ben, sőt bármely véges generátorrendszerből kiválasztható bázis.

Bizonyítás. Legyen a feltételezett generátorrendszer $\mathbf{g}_1, \dots, \mathbf{g}_m$. Ha ez lineárisan független, akkor bázis is, és készen vagyunk. Ha összefüggő, akkor létezik benne \mathbf{g}_i , mely lineárisan függ a többitől, tehát felírható $\mathbf{g}_i = \lambda_1 \mathbf{g}_1 + \dots + \lambda_{i-1} \mathbf{g}_{i-1} + \lambda_{i+1} \mathbf{g}_{i+1} + \dots + \lambda_m \mathbf{g}_m$ alakban (nyilván csak $m - 1$ db λ van az előbbi felírásban, csak az indexelés könnyítéséhez írtuk fel így). Vegyük észre, hogy ez elhagyható a generáláshoz: ha $\mathbf{v} \in V$ tetszőleges, előállítása $\mathbf{v} = \alpha_1 \mathbf{g}_1 + \dots + \alpha_i \mathbf{g}_i + \dots + \alpha_m \mathbf{g}_m$, akkor az előzőek miatt

$$\mathbf{v} = \alpha_1 \mathbf{g}_1 + \dots + \alpha_i \underbrace{(\lambda_1 \mathbf{g}_1 + \dots + \lambda_{i-1} \mathbf{g}_{i-1} + \lambda_{i+1} \mathbf{g}_{i+1} + \dots + \lambda_m \mathbf{g}_m)}_{\mathbf{g}_i} + \dots + \alpha_m \mathbf{g}_m$$

amely felírásban már csak $\mathbf{g}_1, \dots, \mathbf{g}_{i-1}, \mathbf{g}_{i+1}, \dots, \mathbf{g}_m$ vektorok szerepeltek. Ezt a kiválasztást folytathatjuk addig, amíg független rendszert nem kapunk. Ha csak egy vektorunk marad, akkor az lesz bázis, mert az nem lehet a nullvektor (ha az lenne, akkor $V = \{\mathbf{0}\}$ lenne, amit kizártunk). \square

12. Tétel (Kicserélési tétel). *Legyen $V \leq \mathbb{R}^n$, $\mathbf{a}_1, \dots, \mathbf{a}_k$ lineárisan független rendszer V -ben, $\mathbf{b}_1, \dots, \mathbf{b}_m$ generátorrendszer V -ben. Ekkor*

- 1) *van olyan \mathbf{b}_j vektor a generátorrendszerben, mellyel $\mathbf{b}_j, \mathbf{a}_2, \dots, \mathbf{a}_k$ lineárisan független;*
- 2) *$k \leq m$, tehát a lineárisan független rendszerek elemszáma nem nagyobb, mint a generátorrendszerek elemszáma.*

Bizonyítás. Az állításokat sorban bizonyítjuk.

- 1) $k = 1$ esetén egyszerű a helyzet, hisz bármely nemnulla generátorvektorra kicserélhető, mert könnyen látszik, hogy bármely nemnulla vektor önmagában lineárisan független. A $k \geq 2$ esetben indirekt úton tegyük fel, hogy nem létezik az említett vektor, tehát minden $j \in \{1, \dots, m\}$ -re $\mathbf{b}_j, \mathbf{a}_2, \dots, \mathbf{a}_k$ összefüggő rendszer. Vegyük észre, hogy mivel $\mathbf{a}_1, \dots, \mathbf{a}_k$ független, ezért $\mathbf{a}_2, \dots, \mathbf{a}_k$ is az (nem nehéz belátni, hogy független rendszerből vektort elhagyva nem kaphatunk összefüggő rendszert). Korábbi ismereteink alapján tudjuk, hogy ekkor \mathbf{b}_j lineárisan függ $\mathbf{a}_2, \dots, \mathbf{a}_k$ -tól minden $j \in \{1, \dots, m\}$ esetén. Azt kaptuk tehát, hogy $\mathbf{a}_2, \dots, \mathbf{a}_k$ -val ki tudjuk fejezni az összes generátorvektort, a generátorvektorokkal pedig \mathbf{a}_1 -et, ezek szerint \mathbf{a}_1 is lineárisan függ $\mathbf{a}_2, \dots, \mathbf{a}_k$ -tól, ez viszont ellentmond annak, hogy $\mathbf{a}_1, \dots, \mathbf{a}_k$ lineárisan független.
- 2) A tétel első állítását használjuk ki. Nyilván az állítás nem csak az első vektor kicserélésére érvényes, egymás után tudjuk alkalmazni a cserét. Indirekt úton tegyük fel, hogy $k > m$. Ez azt jelenti, hogy az összes generátorvektort ki tudtuk cserélni, és így kapunk a tétel szerint egy $\mathbf{b}_1, \dots, \mathbf{b}_m, \mathbf{a}_{m+1}, \dots, \mathbf{a}_k$ lineárisan független rendszert. Ez azonban nem lehetséges, hisz a benn maradt összes \mathbf{a}_ℓ vektor ($\ell \in \{m+1, \dots, k\}$) lineárisan függ $\mathbf{b}_1, \dots, \mathbf{b}_m$ -től.

\square

A kiválasztási tétel után kissé port kavarhat az elménkben a fenti tétel. Lényegében arról szól, hogy adott független rendszer esetén bármelyik vektort kicserélhetjük alkalmas generátorvektorra, az a generátorvektor nem fog lineárisan függeni a többitől. Most a bázisokkal kapcsolatban fogunk belátni egy fontos és hasznos állítást. A jelölésről: ha azt írjuk, hogy pl. B egy bázis, akkor csak a bázisvektorok leírását rövidítjük, tehát $B = \mathbf{b}_1, \dots, \mathbf{b}_k$.

17. Definíció. Egy végtelen elemszámú vektorrendszer lineárisan független, ha bármely véges részrendszere lineárisan független.

13. Tétel. Legyenek B_1, B_2 bázisok $V \leq \mathbb{R}^n$ -ben, és $|B_1| = k$ az elemszám. Ekkor B_2 is véges, és $|B_2| = k$.

Bizonyítás. Ha B_2 végtelen volna, akkor ki tudunk venni belőle egy $k + 1$ elemű független részrendszert. Tudjuk, hogy a független rendszerek elemszáma legfeljebb egy bázis elemszáma lehet, nagyobb nem (hisz a bázist bármely vektorral bővítve az új vektor biztosan lineárisan függne a többitől, hiszen a bázis előállít mindent), így ellentmondás adódik $k + 1 \leq k$ -val. Alkalmazzuk a kicserélési tétel 2) pontját trükkösen: mivel a bázis lineárisan független és generátorrendszer is, így $|B_1| \leq |B_2|$ adódik, majd a „felcserélt szerepekkel” $|B_2| \leq |B_1|$, ami már implikálja a kívánt egyenlőséget. \square

A fenti tétel megkönnyíti az életünket, beláttuk hogy amennyiben létezik véges bázis (ebben a tárgyban leginkább csak véges bázisok szerepelnek majd, léteznek azonban bizonyos lineáris terekben végtelen elemszámú bázisok, lásd a valós polinomok $\mathbb{R}[x]$ vektorterét), abban az esetben minden bázis elemszáma ugyanennyi lesz. Emlékszünk a korábbi „sík a térben” példákra? Mutattunk benne egy 2 elemű bázist. Ezzel a tétellel igazoltuk, hogy abban az altérben csak kételemű bázisok léteznek. Ezt a mennyiséget fogjuk dimenzió néven említeni.

18. Definíció. Tetszőleges $V \leq \mathbb{R}^n$ esetén a tér dimenziója

$$\dim V = \begin{cases} 0 & \text{ha } V = \{\mathbf{0}\} \\ \text{egy tetszőleges bázis elemszáma} & \text{ha } V \neq \{\mathbf{0}\}. \end{cases}$$

Az előző példánkban tehát a háromdimenziós térben az említett sík egy kétdimenziós altér. Tegyük most rendbe az eddigi kusza gondolatokat. \mathbb{R}^n -ben ismerünk n elemű generátorrendszert (hisz ilyen pl. a korábban említett triviális $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis). Tudjuk már, hogy lineárisan független rendszer \mathbb{R}^n -ben ezek szerint legfeljebb csak n elemű lehet, ami nyilván \mathbb{R}^n bármely alterére is igaz. A $\{\mathbf{0}\}$ nulltérben a $\mathbf{0}$ egy véges generátorrendszer. Amennyiben $\{\mathbf{0}\} \neq V \leq \mathbb{R}^n$, úgy nyilván létezik $\mathbf{0} \neq \mathbf{a} \in V$. Ez a vektor önmagában lineárisan független, tehát egészen biztosan létezik független rendszer, és mivel minden független rendszer elemszáma legfeljebb n , ezért létezik maximális elemszámú független rendszer V -ben, legyen ez $\mathbf{a}_1, \dots, \mathbf{a}_k$. Ekkor bármely $\mathbf{b} \in V$ esetén $\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{b}$ már összefüggő, hiszen a maximális k -nál több elemű. Tudjuk azt is, hogy a fenti esetben \mathbf{b} lineárisan függ a többitől, és mivel \mathbf{b} -re semmi kikötést nem tettünk, bármelyik vektor lehet, ezért $\mathbf{a}_1, \dots, \mathbf{a}_k$ nem csak független, hanem generátorrendszer is, tehát bázis. Ezt egy tételben összefoglaljuk:

14. Tétel. $V \leq \mathbb{R}^n$ esetén létezik V -ben véges generátorrendszer, sőt V -ben bármely lineárisan független rendszer kiegészíthető V bázisává.

Legyen $V \leq \mathbb{R}^n$ és $\dim V = k > 0$. Tudjuk, hogy $k \leq n$, illetve ha G tetszőleges generátorrendszer, B tetszőleges bázis és L tetszőleges független rendszer V -ben, akkor $|B| = k$, $|L| \leq k$ és $|G| \geq k$. Bázisok esetén tudjuk, hogy $|L| = |G| = k$, de még meg kell mutatnunk, hogy ez csak bázisokra igaz. Ha $|G| = k$, akkor tudjuk, hogy G -ből kiválasztható bázis, ami csak k elemű lehet, így $G = B$. Hasonlóan, ha $|L| = k$, akkor tudjuk, hogy L kiegészíthető bázissá, mely csak k elemű lehet, így $L = B$. Általánosan $|L| \leq |B| \leq |G|$, és pontosan a bázisokra teljesülnek az egyenlőségek.

19. Definíció. $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ esetén az $\mathbf{a}_1, \dots, \mathbf{a}_m$ vektorrendszer rangja az általuk generált altér dimenziója, melyet ρ -val, ρ -val vagy r -rel jelölünk, tehát

$$\rho(\mathbf{a}_1, \dots, \mathbf{a}_m) = \dim \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m).$$

A rang felderítésekor tehát bázist kell keresnünk a vektorok által kifeszített altérben, és ennek a bázisnak az elemszáma adja a rangot. Kicsit végiggondolva, azért lényegesen egyszerűbb ennek a vizsgálata, mint ahogy elsőre tűnhet: Nyilván $\mathbf{a}_1, \dots, \mathbf{a}_m$ generátorrendszer $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m)$ -ben, hisz utóbbi az $\mathbf{a}_1, \dots, \mathbf{a}_m$ vektorok összes lineáris kombinációinak halmaza, más szóval minden $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m)$ -beli vektor előáll az $\mathbf{a}_1, \dots, \mathbf{a}_m$ vektorokkal. Nincs más dolgunk innentől kezdve mint a generátorrendszerből kiválasztani bázist. Ezt megtehetjük úgy, hogy $\mathbf{a}_1, \dots, \mathbf{a}_m$ vektorok közül keresünk olyat, melyet a többi előállít, és ha találunk ilyet, azt elhagyjuk, és kezdjük előlről. Ha több ilyen nincs, bázist kapunk, és annak elemszáma fogja a rangot adni.

2.5. Lineáris egyenletrendszerek 1.

Általánosan egy m egyenletből álló, n ismeretlenű lineáris egyenletrendszer a következő alakú:

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

Adottak tehát az a_{ij} , b_i valós számok, és keressük az x_j megoldásokat. Eddigi ismereteink segítségével az adatokat szervezzük vektorokba a triviális bázis szerint. Legyen tehát

$\mathbf{e}_1, \dots, \mathbf{e}_m$ a triviális bázis \mathbb{R}^m -ben. A vektorok a következő alakúak (az elnevezésüket alatta indokoljuk):

$$[\mathbf{b}]_{\mathbf{e}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, [\mathbf{a}_1]_{\mathbf{e}} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, [\mathbf{a}_2]_{\mathbf{e}} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, [\mathbf{a}_n]_{\mathbf{e}} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Az elnevezés azt jelenti, hogy ezeket a vektorokat a triviális $\mathbf{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ bázisban írtuk fel. Ez az általános felírás, de a bázistranszformációnál láttuk, hogy más bázisban „nem természetes” a vektorok alakja. Más szóval, pl $[\mathbf{b}]_{\mathbf{e}}$ azt jelenti, hogy $\mathbf{b} = b_1\mathbf{e}_1 + \dots + b_m\mathbf{e}_m$. Az egyenletrendszer most felírható a következő alakban (ha megengedjük az $\mathbf{a}_j = [\mathbf{a}_j]_{\mathbf{e}}$ jelölést):

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{b}.$$

Így tehát maga a megoldhatóság, tehát az hogy léteznek-e egyáltalán az adott $x_1, \dots, x_n \in \mathbb{R}$ számok azzal ekvivalens, hogy \mathbf{b} benne van-e $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$ -ben. Ha ugyanis benne van, akkor bizony léteznek olyan számok, melyekkel \mathbf{b} kifejezhető a többi segítségével.

Ismerünk egy módszert báziscserére, az elemi bázistranszformációt. Hogy segít ez rajtunk? Gondoljunk bele: Ha elérnénk, hogy $\mathbf{a}_1, \dots, \mathbf{a}_n$ bázisként viselkedjen, akkor kifejezhető lenne biztosan segítségükkel \mathbf{b} . Itt persze sokféle eset lehetséges: lehet, hogy nincs megoldás, lehet hogy végtelen sok van, de az is lehet hogy pontosan egy megoldás van. Vizsgálódjunk, legyen a transzformáció kezdeti táblázata a következő:

	\mathbf{a}_1	\mathbf{a}_2	\dots	\mathbf{a}_n	\mathbf{b}
\mathbf{e}_1	a_{11}	a_{12}	\dots	a_{1n}	b_1
\mathbf{e}_2	a_{21}	a_{22}	\dots	a_{2n}	b_2
\vdots	\vdots	\vdots		\vdots	\vdots
\mathbf{e}_m	a_{m1}	a_{m2}	\dots	a_{mn}	b_m

Hogy olvasnánk ezt a táblázatot? Roppant egyszerűen, adott oszlopát úgy, pl. \mathbf{a}_2 -es oszlopot, fentről lefelé, azt olvassuk, hogy $\mathbf{a}_2 = a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + \dots + a_{m2}\mathbf{e}_m$. Célunk tehát $\mathbf{a}_1, \dots, \mathbf{a}_n$ közül a lehető legtöbbet bevinni a bázisba, a szaggatott vonal azt jelzi hogy \mathbf{b} -t nem akarjuk bevinni, neki csupán csak az értékeit akarjuk majd leolvasni.

A módszer ezek után az, hogy adott \mathbf{a}_\bullet -t próbálunk bevinni $\mathbf{e}_{\bullet\bullet}$ helyére (a pöttyök azt jelentik, hogy az index lehet azonos, de lehet különböző is). Tudjuk, hogy \mathbf{a}_j akkor vihető be

\mathbf{e}_i helyére, ha $a_{ij} \neq 0$, tehát abban az esetben nem tudunk elindulni sem, ha minden $a_{ij} = 0$ lenne. Ez azt jelentené, hogy mindegyik $\mathbf{a}_1, \dots, \mathbf{a}_n$ a nullvektor, ekkor ha $\mathbf{b} = \mathbf{0}$ is teljesül akkor minden valós szám n -es megoldás, ha utóbbi nem teljesül, akkor nincs megoldás. Ha az eljárás elindul, akkor adott \mathbf{a}_j -t viszünk be \mathbf{e}_i helyére. Ez előbb utóbb leáll. Vagy elfogynak az \mathbf{e} -k, vagy elfogynak az \mathbf{a} -k, vagy minden maradék lehetséges cserhelyen 0 áll. Általánosan a következőképp lehet leírni a végeredményt:

	$\mathbf{a}_{i_{r+1}}$	\dots	\mathbf{a}_{i_n}	\mathbf{b}
\mathbf{a}_{i_1}	d_{11}	\dots	$d_{1,n-r}$	d_1
\vdots	\vdots		\vdots	\vdots
\mathbf{a}_{i_r}	d_{r1}	\dots	$d_{r,n-r}$	d_r
\mathbf{e}'_{r+1}	0	\dots	0	\bullet
\vdots	\vdots		\vdots	\vdots
\mathbf{e}'_m	0	\dots	0	\bullet

Ezt az ábrát most vegyük alaposan szemügyre: ez az az állapot, amikor az $\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_r}$ vektorokat be tudtuk vinni a bázisba, az új index azt jelzi, hogy adott \mathbf{e}_i helyére nem feltétlen \mathbf{a}_i -t visszük be, a sorrend tetszőleges lehet. Ezek az $\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_r}$ vektorok bármelyikek lehetnek a $\mathbf{a}_1, \dots, \mathbf{a}_n$ közül, annyit látunk csak, hogy r darabot sikerült bevinni. A megmaradt bázisvektorokat azért láttuk el vesszővel, mert a fentiek miatt ezek indexelése is eltérhet, tehát lehet, hogy $\mathbf{e}'_m \neq \mathbf{e}_m$. A nullák jelzik, hogy onnan már semmit sem tudunk bevinni (hisz ha lenne ott nem nulla elem, azt még be tudnánk vinni). Milyenek esetek fordulhatnak elő? Hogyan is kell olvasni a táblázatot? A fejlécben szereplő vektorok előállítását láthatjuk, az alatta szereplő értékek a lineáris kombináció együtthatói, vektorai pedig a bal oszlopban lévő vektorok. A jobb érthetőség végett írjuk fel pl. \mathbf{a}_{i_n} -et a táblázat alapján:

$$\mathbf{a}_{i_n} = d_{11}\mathbf{a}_{i_1} + \dots + d_{r1}\mathbf{a}_{i_r} + 0 \cdot \mathbf{e}'_{r+1} + \dots + 0 \cdot \mathbf{e}'_m.$$

Az összeg végén lévő nullás tagokat már ki sem kell írni, mi most a teljesség igényével megtettük. Vegyük észre, hogy a fenti egyenlőség azt jelenti, hogy

$$\mathbf{a}_{i_n} \in \text{Span}(\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_r}).$$

Az egyenletrendszer szempontjából nekünk a \mathbf{b} előállítása az érdekes. Vegyük észre, hogy ha minden $\bullet = 0$, akkor biztosan létezik megoldás, hiszen ekkor \mathbf{b} előáll csupa $\mathbf{a}_1, \dots, \mathbf{a}_n$ -beli vektor lineáris kombinációjaként. Akkor is hasonló a helyzet, ha $r = m$, hiszen ekkor minden bázisvektort ki tudunk cserélni. Az eddigi ismereteink alapján azt is megállapíthatjuk, hogy

a fenti általános esetben $\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_r}$ bázis $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$ -ben, továbbá $\varrho(\mathbf{a}_1, \dots, \mathbf{a}_n) = r$. Miért? Beszéltük, hogy generátorrendszerből kiválasztható bázis, ekkor azokat a vektorokat hagyjuk el, melyeket a többivel ki tudunk fejezni, és látjuk, hogy $\mathbf{a}_{i_{r+1}}, \dots, \mathbf{a}_{i_n}$ előállítható csupán $\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_r}$ vektorokkal. Emiatt kapunk bázist, hiszen több vektort már nem tudtunk bevinni, ha pedig van egy bázisunk, annak elemszáma lesz az altér dimenziója, ami pedig a rang.

Lehetséges továbbá, hogy $r = n$, ekkor minden vektort bevittünk, de maradhattak még ezen felüli bázisvektorok. Ekkor egész biztosan 0 vagy 1 megoldás van. 0 akkor, ha a maradék bázisvektorok helyén van nullától különböző elem is, tehát \mathbf{b} előállításához szükség van eredeti bázisvektorra is. Másik esetben egyértelmű a megoldás, hiszen \mathbf{b} -t az $\mathbf{a}_1, \dots, \mathbf{a}_n$ vektorok egyértelműen állítják elő. Amennyiben $r < n$, akkor maradtak be nem vitt vektorok. Lehet hogy ekkor is nulla megoldás van pontosan a fenn említett ok miatt, de ekkor lehet végtelen is. Miért? Ekkor a teljes $\mathbf{a}_1, \dots, \mathbf{a}_n$ összefüggő, hiszen nem tudtuk az összes vektort bevinni. Erre nézünk egy példát is. Legyen a megoldandó rendszer a következő:

$$x_1 + x_2 + 2x_3 + 4x_4 = -3$$

$$x_1 + 3x_2 + 3x_3 - x_4 = -1$$

$$2x_1 + 6x_2 + 6x_3 - 5x_4 = 3$$

$$x_1 + 3x_2 + 3x_3 + 2x_4 = -6$$

A korábbi jelölésekkel ekkor

$$\mathbf{b} = \begin{bmatrix} -3 \\ -1 \\ 3 \\ -6 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 3 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 2 \\ 3 \\ 6 \\ 3 \end{bmatrix}, \mathbf{a}_4 = \begin{bmatrix} 4 \\ -1 \\ -5 \\ 2 \end{bmatrix}$$

Ezzel a következő egyenlet adódik:

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 + x_4\mathbf{a}_4 = \mathbf{b}.$$

Végezzük el a bázistranszformációt, kerettel jelöljük, hogy melyik lépésben mit viszünk be a bázisba (ha ezt eddig nem említettük volna, ez a sorrend lényegtelen).

$$\begin{array}{c|cccc|c} & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{b} \\ \hline \mathbf{e}_1 & \boxed{1} & 1 & 2 & 4 & -3 \\ \mathbf{e}_2 & 1 & 3 & 3 & -1 & -1 \\ \mathbf{e}_3 & 2 & 6 & 6 & -5 & 3 \\ \mathbf{e}_4 & 1 & 3 & 3 & 2 & -6 \end{array} \left\| \begin{array}{c|ccc|c} & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{b} \\ \hline \mathbf{a}_1 & 1 & 2 & 4 & -3 \\ \mathbf{e}_2 & 2 & 1 & -5 & 2 \\ \mathbf{e}_3 & 4 & 2 & -13 & 9 \\ \mathbf{e}_4 & 2 & \boxed{1} & -2 & -3 \end{array} \right\| \left\| \begin{array}{c|cc|c} & \mathbf{a}_2 & \mathbf{a}_4 & \mathbf{b} \\ \hline \mathbf{a}_1 & -3 & 8 & 3 \\ \mathbf{e}_2 & 0 & \boxed{-3} & 5 \\ \mathbf{e}_3 & 0 & -9 & 15 \\ \mathbf{a}_3 & 2 & -2 & -3 \end{array} \right\| \left\| \begin{array}{c|c|c} & \mathbf{a}_2 & \mathbf{b} \\ \hline \mathbf{a}_1 & -3 & 49/3 \\ \mathbf{a}_4 & 0 & -5/3 \\ \mathbf{e}_3 & 0 & 0 \\ \mathbf{a}_3 & 2 & -19/3 \end{array} \right\|$$

Látjuk, hogy csak az \mathbf{a}_2 -t nem tudtuk bevinni, mert az osztó helyen 0 van. Ha $\mathbf{e}_3 - \mathbf{b}$ cellában nem nulla volna, akkor nem létezne megoldás, így viszont végtelen sok létezik. A következőképp kell az adatokat kiolvasni:

$$x_1 = \frac{49}{3} + 3x_2,$$

$$x_4 = -\frac{5}{3},$$

$$x_3 = -\frac{19}{3} - 2x_2.$$

A be nem vitt vektorhoz tartozó változó az ún. szabad változó. Látjuk, hogy tetszőleges $x_2 \in \mathbb{R}$ esetén léteznek a fenti értékek. Nézzünk egy példát ellenőrzésképp, legyen pl. $x_2 = 4$. Ekkor $x_1 = 85/3$, $x_3 = -43/3$ és $x_4 = -5/3$, mely utóbbi mindig ennyi.

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 3 \\ 6 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 3 \\ 6 \\ 3 \end{bmatrix} + x_4 \begin{bmatrix} 4 \\ -1 \\ -5 \\ 2 \end{bmatrix} = \begin{bmatrix} 85/3 + 4 - 86/3 - 20/3 \\ 85/3 + 12 - 129/3 + 5/3 \\ 170/3 + 24 - 258/3 + 25/3 \\ 85/3 + 12 - 129/3 - 10/3 \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \\ 3 \\ -6 \end{bmatrix} = \mathbf{b}$$

tehát valóban megoldást kaptunk. A mátrix fogalmának bevezetésével precízebben fogjuk tudni kezelni a lineáris egyenletrendszereket, jobb megoldhatósági feltételeket adunk, és a későbbi félévekben (Numerikus módszerek c. tárgyan) hatékony algoritmusokat is tanulunk akár nagyon nagy rendszerek megoldásaira.

3. fejezet

Mátrixok

A mátrixokkal az eddig megismert vektor fogalmát általánosítjuk.

3.1. Értelmezés és műveletek

20. Definíció. Legyenek m, n pozitív egész számok adottak, továbbá minden $i = 1, \dots, m$ és $j = 1, \dots, n$ értékekre az a_{ij} módon jelölt valós számok (világos, hogy $m \cdot n$ darab ilyen van).

Az

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

táblázatot egy \mathbb{R} feletti mátrixnak nevezzük, s (ebben a jegyzetben) vastagon szedett nagybetűkkel jelöljük őket. Ha a fentit \mathbf{A} -val jelöljük, akkor $\mathbf{A} \in \mathbb{R}^{m \times n}$, mely utóbbi halmaz az összes m sorból és n oszlopból álló valós számok feletti mátrixok halmaza. Magát a mátrixot jelölhetjük még $\mathbf{A} = [a_{ij}]_{m \times n}$ módon is. Adott i . sor j . elemét a_{ij} -vel, vagy ${}_i[\mathbf{A}]_j$ -vel fogjuk jelölni.

Az \mathbf{A} és \mathbf{B} mátrixokat akkor tekintjük egyenlőnek, ha $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ és minden szóba jövő i, j párra ${}_i[\mathbf{A}]_j = {}_i[\mathbf{B}]_j$. Az összeadást és számmal való szorzást ugyanolyan módon definiáljuk, mint a vektorok esetében:

21. Definíció. Legyenek $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ és $\lambda \in \mathbb{R}$. Ekkor $\mathbf{A} + \mathbf{B} \in \mathbb{R}^{m \times n}$, melyre minden szóba jövő i, j esetén ${}_i[\mathbf{A} + \mathbf{B}]_j = {}_i[\mathbf{A}]_j + {}_i[\mathbf{B}]_j$. Hasonlóan, $\lambda \mathbf{A} \in \mathbb{R}^{m \times n}$ úgy, hogy minden i, j esetén ${}_i[\lambda \mathbf{A}]_j = \lambda {}_i[\mathbf{A}]_j$.

Így már világos, hogy miért beszélhetünk a vektorok általánosításáról, hisz vegyük észre, hogy $\mathbb{R}^m = \mathbb{R}^{m \times 1}$. Érdekeség továbbá, hogy a fenti műveletekkel $\mathbb{R}^{m \times n}$ vektortér, így itt is

értelmezhetőek a függetlenség, összefüggés, bázis, generátorrendszer fogalmak, nézzünk erre egy példát. Könnyen láthatóan $\mathbb{R}^{2 \times 2}$ -ben a következő mátrixnégyes bázis:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

3.2. Mátrixok szorzása

Egy teljesen új dologgal fogunk most megismerkedni, mégpedig a szorzással. Első gondolatra az volna természetes, ha két ugyanannyi komponensből álló mátrixot komponensenként szoroznánk, azonban gyakorlati szempontból ez nem célravezető. **A természetesség helyett most az alkalmazhatóság a cél!** Tekintsük ismét az általános lineáris egyenletrendszert:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = b_m$$

Rendeljük most a rendszer elemeihez a következő mátrixokat / vektorokat:

$$\mathbf{A} := \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}, \quad \mathbf{b} := \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m, \quad \mathbf{x} := \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

Látszólag szépen szeparáltuk az adatokat a már ismert struktúrákba, azonban a dimenziója mindegyiknek más. Célunk a szorzás definiálásakor, hogy az egyenletrendszert fel tudjuk írni $\mathbf{Ax} = \mathbf{b}$. alakban. De mégis hogyan szoroznánk egy $\mathbb{R}^{m \times n}$ -beli mátrixot egy $\mathbb{R}^{n \times 1}$ -beli mátrixszal (vektorral), hogy eredményképp egy $\mathbb{R}^{m \times 1}$ -beli mátrixot (vektort) kapjunk? A definícióból minden szépen világossá válik:

22. Definíció. Legyenek $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times k}$. Legyen ekkor $\mathbf{AB} \in \mathbb{R}^{m \times k}$ úgy, hogy minden $i = 1, \dots, m$ és $j = 1, \dots, k$ esetén

$$i[\mathbf{AB}]_j = \sum_{\ell=1}^n i[\mathbf{A}]_{\ell\ell}[\mathbf{B}]_j$$

Na ez most micsoda? Válasszunk egy tetszőleges i, j párt, és a definícióbeli összeget írjuk ki részletesen:

$$i[\mathbf{AB}]_j = i[\mathbf{A}]_{11}[\mathbf{B}]_j + i[\mathbf{A}]_{22}[\mathbf{B}]_j + \cdots + i[\mathbf{A}]_{nn}[\mathbf{B}]_j$$

Mátrix szorzásánál tehát adott i, j indexű elemre a szorzatmátrix bal mátrixának i . sora kötött, jobb mátrixának pedig j -edik oszlopa, majd a bal mátrix adott során balról jobbra, a jobb mátrix adott oszlopán fentről lefelé haladva páronként szorzunk, ezeket a szorzatokat pedig összeadjuk. Világos, hogy pont ugyanannyi elem van az adott sorban ill. oszlopban, hisz ez feltétel volt a definícióban. Egyből felmerülhet a gyanú, hogy a mátrixszorzás nem lesz kommutatív művelet. Amennyiben $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times k}$, akkor $\mathbf{AB} \in \mathbb{R}^{m \times k}$, viszont \mathbf{BA} nem is értelmes, ha $k \neq m$. Ha ez az egyenlőség fennáll, tehát $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$, akkor ugyan \mathbf{AB}, \mathbf{BA} léteznek, csak épp, $\mathbf{AB} \in \mathbb{R}^{m \times m}$ és $\mathbf{BA} \in \mathbb{R}^{n \times n}$, melyek szintén nem egyezhetnek meg, ha $m \neq n$, így tehát már csak az $m = n$ eset maradt. Könnyen adható példa az egyenlőség nem teljesülésére, legyenek pl

$$\mathbf{A} := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{B} := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Ekkor

$$\mathbf{AB} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{BA} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Látjuk továbbá érdekességgként, hogy két nem nullmátrix szorzata adhatja eredményül a nullmátrixot.

Könnyen látható, hogy ezzel a szorzásfogalommal a lineáris egyenletrendszer mátrixos alakjában már teljesül, hogy $\mathbf{Ax} = \mathbf{b}$. Később a lineáris transzformációk vizsgálatakor a különböző mozgások leírására is mátrixszorzást fogunk alkalmazni, ahogy az a modern számítógépes grafikában is így van (lásd: Számítógépes grafika c. tárgy).

23. Definíció. Az $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ mátrixot $n \times n$ -es egységmátrixnak nevezzük, ha a szóba jövő i, j -k esetén

$${}_i[\mathbf{I}_n]_j = \delta_{ij} = \begin{cases} 1 & \text{ha } i = j \\ 0 & \text{ha } i \neq j. \end{cases}$$

A δ_{ij} csak leíráskönnyítő ez esetben, ún. Kronecker-szimbólum.

Az egységmátrix tehát a következő alakú:

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

Megmutatjuk, hogy jogos az egységmátrix elnevezés, ugyanis a vele való szorzás minden mátrixot helyben hagy, tehát nem változtat:

15. Tétel. $\forall \mathbf{A} \in \mathbb{R}^{m \times n} : \mathbf{I}_m \mathbf{A} = \mathbf{A} \mathbf{I}_n = \mathbf{A}$.

Bizonyítás. Az $\mathbf{I}_m \mathbf{A} = \mathbf{A}$ esetet mutatjuk meg, a másik teljesen analóg módon bizonyítható. Úgy bizonyítunk, hogy tetszőleges i, j párra látjuk be az állítást:

$${}_i[\mathbf{I}_m \mathbf{A}]_j = \sum_{\ell=1}^m {}_i[\mathbf{I}_m]_{\ell\ell} [\mathbf{A}]_j = {}_i[\mathbf{I}_m]_{ii} [\mathbf{A}]_j = 1 \cdot {}_i[\mathbf{A}]_j = {}_i[\mathbf{A}]_j.$$

Az összegből a többi tag azért tűnik el, mert $i \neq \ell$ esetén ${}_i[\mathbf{I}_m]_{\ell\ell} = 0$. □

Bevezetünk egy új műveletet, az ún. transzponálást:

24. Definíció. $\mathbf{A} \in \mathbb{R}^{m \times n}$ esetén az \mathbf{A} mátrix transzponáltja $\mathbf{A}^\top \in \mathbb{R}^{n \times m}$ úgy, hogy minden szóba jövő i, j -re ${}_i[\mathbf{A}^\top]_j = {}_j[\mathbf{A}]_i$.

Szerencsére a transzponálás az eddigi műveleteinkel egész szépen működik együtt:

16. Tétel. Teljesülnek a következők:

1. $\forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n} : (\mathbf{A} + \mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$,
2. $\forall \mathbf{A} \in \mathbb{R}^{m \times n}, \lambda \in \mathbb{R} : (\lambda \mathbf{A})^\top = \lambda \mathbf{A}^\top$,
3. $\forall \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{n \times k} : (\mathbf{A} \mathbf{B})^\top = \mathbf{B}^\top \mathbf{A}^\top$.

Bizonyítás. Szépen sorban, mechanikusan, tetszőleges i, j -re:

$$\begin{aligned} {}_i[(\mathbf{A} + \mathbf{B})^\top]_j &= {}_j[\mathbf{A} + \mathbf{B}]_i = {}_j[\mathbf{A}]_i + {}_j[\mathbf{B}]_i = {}_i[\mathbf{A}^\top]_j + {}_i[\mathbf{B}^\top]_j = {}_i[\mathbf{A}^\top + \mathbf{B}^\top]_j, \\ {}_i[(\lambda \mathbf{A})^\top]_j &= {}_j[\lambda \mathbf{A}]_i = \lambda {}_j[\mathbf{A}]_i = \lambda {}_i[\mathbf{A}^\top]_j = {}_i[\lambda \mathbf{A}^\top]_j, \\ {}_i[(\mathbf{A} \mathbf{B})^\top]_j &= {}_j[\mathbf{A} \mathbf{B}]_i = \sum_{\ell=1}^n {}_j[\mathbf{A}]_{\ell\ell} [\mathbf{B}]_i = \\ &= \sum_{\ell=1}^n {}_\ell[\mathbf{A}^\top]_{ji} [\mathbf{B}^\top]_\ell = \sum_{\ell=1}^n {}_i[\mathbf{B}^\top]_{\ell\ell} [\mathbf{A}^\top]_j = {}_i[\mathbf{B}^\top \mathbf{A}^\top]_j, \end{aligned}$$

és ezzel készen vagyunk. □

Könnyen látható, hogy amennyiben $\mathbf{A}_1, \dots, \mathbf{A}_n$ ebben a sorrendben összeszorozható mátrixok, akkor

$$(\mathbf{A}_1 \mathbf{A}_2 \cdots \mathbf{A}_n)^\top = \mathbf{A}_n^\top \mathbf{A}_{n-1}^\top \cdots \mathbf{A}_1^\top.$$

Láttuk, hogy a szorzás nem kommutatív művelet, ellenben most megmutatjuk, hogy az asszociativitás, sőt még a disztributivitás is teljesül. Ennek persze szükséges feltétele az, hogy minden szorzás elvégezhető legyen (a megfelelő dimenzióértékek megegyezzenek). Ezeket fogjuk most megmutatni.

17. Tétel. *Legyenek $\mathbf{A} \in \mathbb{R}^{m \times n_1}$, $\mathbf{B} \in \mathbb{R}^{n_2 \times k_2}$, $\mathbf{C} \in \mathbb{R}^{k_3 \times s}$. Akkor és csak akkor létezik $(\mathbf{AB})\mathbf{C}$ és $\mathbf{A}(\mathbf{BC})$, ha $n_1 = n_2$ és $k_2 = k_3$. Ha ez teljesül, akkor $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.*

Bizonyítás. A mátrixszorzás definíciójából a létezés feltétele egyértelmű. Könnyen látjuk, hogy ekkor $(\mathbf{AB})\mathbf{C}, \mathbf{A}(\mathbf{BC}) \in \mathbb{R}^{m \times s}$, így az egyezőséghez a dimenziók rendben vannak. Legyen $n_1 = n_2 =: n$ és $k_2 = k_3 =: k$. Tetszőleges i, j -re bizonyítunk:

$$\begin{aligned} {}_i[(\mathbf{AB})\mathbf{C}]_j &= \sum_{p=1}^k {}_i[\mathbf{AB}]_{pp}[\mathbf{C}]_j = \sum_{p=1}^k \left(\sum_{q=1}^n {}_i[\mathbf{A}]_{qq}[\mathbf{B}]_p \right) {}_p[\mathbf{C}]_j = \\ &= \sum_{q=1}^n \sum_{p=1}^k {}_i[\mathbf{A}]_{qq}[\mathbf{B}]_{pp}[\mathbf{C}]_j = \sum_{q=1}^n {}_i[\mathbf{A}]_q \left(\sum_{p=1}^k {}_p[\mathbf{B}]_{pp}[\mathbf{C}]_j \right) = \sum_{q=1}^n {}_i[\mathbf{A}]_{qq}[\mathbf{BC}]_j = {}_i[\mathbf{A}(\mathbf{BC})]_j, \end{aligned}$$

és ezzel kész vagyunk. □

18. Tétel. *Hasonlóan értelmezett dimenziófeltételek mellett $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$.*

19. Tétel. $\lambda \in \mathbb{R}, \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{n \times k} \implies \lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B} = \mathbf{A}(\lambda\mathbf{B})$.

A fentiek teljesen hasonló módon bizonyíthatók.

3.3. Partícionálás és alkalmazásai

Legyen $\mathbf{A} \in \mathbb{R}^{m \times n}$ és tegyük fel, hogy $m \geq 2$, $n \geq 3$. Ekkor \mathbf{A} felírható pl.

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}$$

alakban, ahol a különböző \mathbf{A}_{ij} -k maguk is mátrixok. Ezt az egyszerű eljárást nevezzük partícionálásnak, vagy blokkosításnak. Világos, hogy ilyenkor az egy sorban lévő blokkok (mátrixok) ugyanannyi sort, míg az egy oszlopban lévők ugyanannyi oszlopot tartalmaznak. A legérdekesebb az, hogy megfelelő partícionálás mellett úgy tudunk mátrixokat szorozni, mintha

a blokkok maguk is számok volnának. Legyenek pl. $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times k}$ a blokkosítás pedig legyen a következő:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \\ \mathbf{B}_{31} & \mathbf{B}_{32} \end{bmatrix}$$

Amennyiben a blokkok dimenziói megfelelők (szorozhatók), akkor a szorzatmátrix a következő blokkmátrix lesz:

$$\mathbf{AB} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} + \mathbf{A}_{13}\mathbf{B}_{31} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} + \mathbf{A}_{13}\mathbf{B}_{32} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} + \mathbf{A}_{23}\mathbf{B}_{31} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} + \mathbf{A}_{23}\mathbf{B}_{32} \end{bmatrix}$$

A leggyakoribb alkalmazása a partícionálásnak az oszlopvektorokra partícionálás. Amennyiben $\mathbf{A} = (a_{ij})_{m \times n} \in \mathbb{R}^{m \times n}$ akkor felírhatjuk \mathbf{A} -t $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ alakban, ahol $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ az oszlopokból nyert vektorok. Az imént a szorzásnál említettek miatt ekkor tetszőleges $\mathbf{x} \in \mathbb{R}^n$ esetén $\mathbf{Ax} = \mathbf{a}_1x_1 + \dots + \mathbf{a}_nx_n$, ahol azt is látjuk, hogy \mathbf{Ax} (ami ugye egy vektor, a dimenziókból is adódik) benne van az $\mathbf{a}_1, \dots, \mathbf{a}_n$ vektorok által kifeszített altérben, hisz lineáris kombinációjuk előállítja azt. Most kicsit pontosabban foglalkozunk azzal az esettel, amikor az $\mathbf{Ax} = \mathbf{b}$ lineáris egyenletrendszernek végtelen sok megoldása van. Az ottani jelölésekkel tegyük fel, hogy létezik megoldás, és $r < n$, tehát van olyan vektor, amit nem tudunk bevinni a bázisba. Indexcserével elérhetjük, hogy a bevitt vektorok $\mathbf{a}_1, \dots, \mathbf{a}_r$ legyenek (tehát \mathbf{A} oszlopainak cseréjével, és egyúttal az ismeretlenek cseréjével). A végén persze ezeket vissza kell majd cserélni. Eddig a transzformációkor általában nem írtuk ki a már bevitt vektorokat, mert azok előállítása úgyis egyértelmű (ha \mathbf{a}_i -t bevittük a bázisba, akkor annak előállítása $1 \cdot \mathbf{a}_i$ és a többi bázisvektor 0-szor). Most azonban kiírjuk ezeket is:

	\mathbf{a}_1	\dots	\mathbf{a}_r	\mathbf{a}_{r+1}	\dots	\mathbf{a}_n	\mathbf{b}
\mathbf{a}_1	\mathbf{I}_r			\mathbf{D}			\mathbf{d}
\vdots							
\mathbf{a}_r							
\mathbf{e}'_{r+1}	$\mathbf{0}$			$\mathbf{0}$			$\mathbf{0}$
\vdots							
\mathbf{e}'_m							

Világos, hogy bal felül egységmátrixot kapunk, az imént említettek miatt. Az is világos, hogy alul nullmátrixok és nullvektor van, ez biztosítja hogy leállt a vektorok bevitele, a nullvektor pedig azt, hogy létezik megoldás. A \mathbf{D} az valamilyen $\mathbb{R}^{r \times (n-r)}$ -beli mátrix. Tekintsük a

következő partícionálásokat: $\mathbf{A}_1 := [\mathbf{a}_1, \dots, \mathbf{a}_r]$, $\mathbf{A}_2 := [\mathbf{I}_r, \mathbf{D}]$. A táblázatot ismerve világos, hogy $\mathbf{A}_1 \mathbf{d} = \mathbf{b}$. Vegyük észre, hogy $\mathbf{A} = \mathbf{A}_1 \mathbf{A}_2$, ugyanis

$$\mathbf{A}_1 \mathbf{A}_2 = [\mathbf{a}_1, \dots, \mathbf{a}_r] \cdot [\mathbf{I}_r, \mathbf{D}] = [\mathbf{A}_1 \mathbf{I}_r, \mathbf{A}_1 \mathbf{D}] = [\mathbf{A}_1, \mathbf{A}_1 \mathbf{D}] = [\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{a}_{r+1}, \dots, \mathbf{a}_n] = \mathbf{A}.$$

Ezeket felhasználva adott $\mathbf{x} \in \mathbb{R}^n$ esetén

$$\mathbf{A} \mathbf{x} = \mathbf{b} \iff \mathbf{A}_1 \mathbf{A}_2 \mathbf{x} = \mathbf{b} \iff \mathbf{A}_1 \mathbf{A}_2 \mathbf{x} = \mathbf{A}_1 \mathbf{d} \iff \mathbf{A}_1 (\mathbf{A}_2 \mathbf{x} - \mathbf{d}) = \mathbf{0}$$

Megmutatjuk, hogy az utolsó egyenlőség éppen azzal ekvivalens, hogy $\mathbf{A}_2 \mathbf{x} - \mathbf{d} = \mathbf{0}$. Ez természetesnek tűnhet, de nem az, hiszen láttuk, hogy két nem nullmátrix szorzata is lehet nullvektor. A bizonyítandó ekvivalencia \Rightarrow iránya onnan következik, hogy \mathbf{A}_1 oszlopai lineárisan független rendszert alkotnak. Ezt most egy lemmaként belátjuk:

1. Lemma. *Legyenek \mathbf{A}, \mathbf{b} ebben a sorrendben összeszorozható valós mátrix és vektor, és tegyük fel, hogy \mathbf{A} oszlopai lineárisan független rendszert alkotnak, és $\mathbf{A} \mathbf{b} = \mathbf{0}$. Ekkor $\mathbf{b} = \mathbf{0}$.*

Bizonyítás. Legyen $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n]$. Ekkor $\mathbf{A} \mathbf{b}$ felírható a következőképp:

$$\mathbf{a}_1 b_1 + \dots + \mathbf{a}_n b_n = \mathbf{0},$$

ez az egyenlet pedig csak úgy lehet igaz, ha $\mathbf{b} = \mathbf{0}$, hiszen $\mathbf{a}_1, \dots, \mathbf{a}_n$ független. \square

A fenti lemma gyakorlatilag a már jól ismert függetlenség mátrixos következménye. Most partícionáljuk \mathbf{x} -et a következő módon:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}'_r \\ \mathbf{x}''_{n-r} \end{bmatrix} \in \mathbb{R}^n.$$

Az indexek arra utalnak, hogy mettől-meddig partícionálunk, a vesszők meg arra az esetre vannak, ha netán $r = n - r$ előfordula. Folytatva a korábbi gondolatmenetet:

$$\mathbf{A} \mathbf{x} = \mathbf{b} \iff \mathbf{A}_2 \mathbf{x} = \mathbf{d} \iff [\mathbf{I}_r, \mathbf{D}] \cdot \begin{bmatrix} \mathbf{x}'_r \\ \mathbf{x}''_{n-r} \end{bmatrix} = \mathbf{d} \iff \mathbf{x}'_r = \mathbf{d} - \mathbf{D} \mathbf{x}''_{n-r}.$$

Ez most érthetetlen lehet elsőre, de azt kell észrevennünk, hogy az általános megoldást kaptuk meg. Azt látjuk, hogy \mathbf{x} felső része, tehát \mathbf{x}'_r függ az alsó résztől, \mathbf{x}''_{n-r} -től, mégpedig a fenti módon. Emlékszünk amikor előző fejezetben milyen módon olvastuk ki a táblázatból az egyenlet általános megoldásait? Szépen látszik, hogy éppen a most kapott egyenlőséget

alkalmaztuk. Ezt most azzal a korábbi példával meg is mutatjuk, emlékezzünk a bázistranszformációra:

$$\begin{array}{c|ccccc|c} & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{b} \\ \hline \mathbf{e}_1 & \boxed{1} & 1 & 2 & 4 & -3 \\ \mathbf{e}_2 & 1 & 3 & 3 & -1 & -1 \\ \mathbf{e}_3 & 2 & 6 & 6 & -5 & 3 \\ \mathbf{e}_4 & 1 & 3 & 3 & 2 & -6 \end{array} \parallel \begin{array}{c|ccccc|c} & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{b} \\ \hline \mathbf{a}_1 & 1 & 2 & 4 & -3 \\ \mathbf{e}_2 & 2 & 1 & -5 & 2 \\ \mathbf{e}_3 & 4 & 2 & -13 & 9 \\ \mathbf{e}_4 & 2 & \boxed{1} & -2 & -3 \end{array} \parallel \begin{array}{c|cc|c} & \mathbf{a}_2 & \mathbf{a}_4 & \mathbf{b} \\ \hline \mathbf{a}_1 & -3 & 8 & 3 \\ \mathbf{e}_2 & 0 & \boxed{-3} & 5 \\ \mathbf{e}_3 & 0 & -9 & 15 \\ \mathbf{a}_3 & 2 & -2 & -3 \end{array} \parallel \begin{array}{c|cc|c} & \mathbf{a}_2 & \mathbf{b} \\ \hline \mathbf{a}_1 & -3 & 49/3 \\ \mathbf{a}_4 & 0 & -5/3 \\ \mathbf{e}_3 & 0 & 0 \\ \mathbf{a}_3 & 2 & -19/3 \end{array} \parallel$$

Sorcsérékkel és a bevitt vektorok kiírásával alakítsuk át a legvégül kapott táblázatot, ahogy azt általánosan is írtuk:

$$\begin{array}{c|ccccc|c} & \mathbf{a}_1 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_2 & \mathbf{b} \\ \hline \mathbf{a}_1 & 1 & 0 & 0 & -3 & 49/3 \\ \mathbf{a}_3 & 0 & 1 & 0 & 2 & -19/3 \\ \mathbf{a}_4 & 0 & 0 & 1 & 0 & -5/3 \\ \mathbf{e}_3 & 0 & 0 & 0 & 0 & 0 \end{array} \xRightarrow{\text{Indexcsere}} \begin{array}{c|ccccc|c} & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{b} \\ \hline \mathbf{a}_1 & 1 & 0 & 0 & -3 & 49/3 \\ \mathbf{a}_2 & 0 & 1 & 0 & 2 & -19/3 \\ \mathbf{a}_3 & 0 & 0 & 1 & 0 & -5/3 \\ \mathbf{e}_3 & 0 & 0 & 0 & 0 & 0 \end{array}$$

Az indexcsere valóban csak indexcsere, hisz látjuk az értékeken nem változtattunk. Nyilván az ismeretlenek \mathbf{x} vektorában is cserélődnek az indexek ilyenkor. Látjuk a táblázatban kapott egységmatrixot, továbbá

$$\mathbf{D} = \begin{bmatrix} -3 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} 49/3 \\ -19/3 \\ -5/3 \end{bmatrix}$$

A \mathbf{D} mátrix most éppen egy vektor, de nyilván lehet többoszlopos mátrix is, amikor egynél több vektort nem tudunk a bázisba vinni. Mivel most $r = 3$ és $n = 4$, az \mathbf{x} -et így módon partícionáljuk:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}'_3 \\ \mathbf{x}''_1 \end{bmatrix} \in \mathbb{R}^4$$

ahol $\mathbf{x}'_3 \in \mathbb{R}^3$ és $\mathbf{x}''_1 \in \mathbb{R}^1 = \mathbb{R}$. Arra jutottunk, hogy $\mathbf{x}'_3 = \mathbf{d} - \mathbf{D}\mathbf{x}''_1$ tehát

$$\mathbf{x}'_3 = \begin{bmatrix} 49/3 \\ -19/3 \\ -5/3 \end{bmatrix} - \begin{bmatrix} -3 \\ 2 \\ 0 \end{bmatrix} \cdot \mathbf{x}''_1 \implies \mathbf{x} = \begin{bmatrix} 49/3 + 3\mathbf{x}''_1 \\ -19/3 - 2\mathbf{x}''_1 \\ -5/3 \\ \mathbf{x}''_1 \end{bmatrix}$$

És ez pontosan az általános megoldás amit kaptunk, miután az eredeti indexeket helyreállítjuk, tehát a tényleges eredmény:

$$\mathbf{x} = \begin{bmatrix} 49/3 + 3\mathbf{x}_1'' \\ \mathbf{x}_1'' \\ -19/3 - 2\mathbf{x}_1'' \\ -5/3 \end{bmatrix},$$

ahol $\mathbf{x}_1'' \in \mathbb{R}$ egy szabad paraméter.

3.4. Rang és inverz

25. Definíció. Az $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n] \in \mathbb{R}^{m \times n}$ mátrix oszloprangja $\varrho_{\mathcal{O}}(\mathbf{A}) = \varrho(\mathbf{a}_1, \dots, \mathbf{a}_n)$ tehát az oszlopvektorok rendszerének rangja, sorrangja pedig $\varrho_S(\mathbf{A}) = \varrho_{\mathcal{O}}(\mathbf{A}^\top)$.

A vektorrendszerek rangját ismerve a fenti definíció teljesen természetes értelmezése a rang fogalmának mátrixokon. Meg fogjuk mutatni, hogy az oszlop-, és sorrang mindig megegyezik, de ennek belátásához először bizonyítunk egy segédtelet.

20. Tétel. Legyenek $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_n]$, $\mathbf{D} = [\mathbf{d}_1, \dots, \mathbf{d}_k]$ ebben a sorrendben összeszorozható valós mátrixok. Ekkor $\varrho_{\mathcal{O}}(\mathbf{CD}) \leq \varrho_{\mathcal{O}}(\mathbf{C})$

Bizonyítás. A szorzást írjuk fel a következőképp: $\mathbf{CD} = \mathbf{C}[\mathbf{d}_1, \dots, \mathbf{d}_k] = [\mathbf{Cd}_1, \dots, \mathbf{Cd}_k]$. Tudjuk, hogy $\mathbf{Cd}_i \in \text{Span}(\mathbf{c}_1, \dots, \mathbf{c}_n)$ minden i -re, hiszen \mathbf{Cd}_i a $\mathbf{c}_1, \dots, \mathbf{c}_n$ vektorok egy lineáris kombinációja, hiszen

$$\begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{m1} & \dots & c_{mn} \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = \begin{bmatrix} c_{11}d_1 + \dots + c_{1n}d_n \\ \vdots \\ c_{m1}d_1 + \dots + c_{mn}d_n \end{bmatrix} = d_1 \begin{bmatrix} c_{11} \\ \vdots \\ c_{m1} \end{bmatrix} + \dots + d_n \begin{bmatrix} c_{1n} \\ \vdots \\ c_{mn} \end{bmatrix} = d_1 \mathbf{c}_1 + \dots + d_n \mathbf{c}_n.$$

Ezek miatt teljesül az is, hogy $\text{Span}(\mathbf{Cd}_1, \dots, \mathbf{Cd}_k) \subseteq \text{Span}(\mathbf{c}_1, \dots, \mathbf{c}_n)$, innen meg már következik az állítás, hisz ha egy altér egy másik altér részhalmaza (csúnyán mondva egy altér egy másik altér altere...) akkor annak dimenziója legfeljebb a bővebb altér dimenziója lehet, vagy annál kisebb. \square

21. Tétel. Bármely $\mathbf{A} \in \mathbb{R}^{m \times n}$ esetén $\varrho_{\mathcal{O}}(\mathbf{A}) = \varrho_S(\mathbf{A})$.

Bizonyítás. A $\mathbf{A} = \mathbf{0}$ eset nyilvánvaló, legyen tehát $\mathbf{0} \neq \mathbf{A} \in \mathbb{R}^{m \times n}$. Bizonyítás nélkül közöljük, hogy az elemi bázistranszformáció során a mátrix sorrangja és oszloprangja nem

változik (még oszlop/sorcserektől sem). Legyen tehát $\mathbf{A}_1 := [\mathbf{a}_1, \dots, \mathbf{a}_r]$ és $\mathbf{A}_2 = [\mathbf{I}_r, \mathbf{D}]$. Ekkor mint láttuk $\mathbf{A} = \mathbf{A}_1 \mathbf{A}_2$ és

$$\begin{aligned} \varrho_S(\mathbf{A}) &= \varrho_S(\mathbf{A}_1 \mathbf{A}_2) = \varrho_O((\mathbf{A}_1 \mathbf{A}_2)^\top) = \varrho_O(\mathbf{A}_2^\top \mathbf{A}_1^\top) \leq \\ &\leq \varrho_O(\mathbf{A}_2^\top) = r = \varrho_O(\mathbf{A}). \end{aligned}$$

Tetszőleges mátrixra beláttuk tehát, hogy $\varrho_S(\mathbf{A}) \leq \varrho_O(\mathbf{A})$. Mivel ez minden mátrixra fennáll, így \mathbf{A}^\top -ra is, tehát $\varrho_S(\mathbf{A}^\top) \leq \varrho_O(\mathbf{A}^\top)$, tehát $\varrho_O(\mathbf{A}) \leq \varrho_S(\mathbf{A})$, ami pedig pont az egyenlőséget implikálja. \square

A következőkben az $\mathbf{A}\mathbf{x} = \mathbf{b}$ feladat általánosítását fogjuk vizsgálni, az $\mathbf{A}\mathbf{X} = \mathbf{B}$ feladatot, ahol $\mathbf{A}, \mathbf{B}, \mathbf{X}$ megfelelő méretű mátrixok. Legyen tehát $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n] \in \mathbb{R}^{m \times n}$, $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_k] \in \mathbb{R}^{m \times k}$ és az $\mathbf{A}\mathbf{X} = \mathbf{B}$ egyenlet megoldását keressük $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_k] \in \mathbb{R}^{n \times k}$ alakban:

$$\mathbf{A}\mathbf{X} = \mathbf{B} \iff \mathbf{A}[\mathbf{x}_1, \dots, \mathbf{x}_k] = [\mathbf{b}_1, \dots, \mathbf{b}_k] \iff [\mathbf{A}\mathbf{x}_1, \dots, \mathbf{A}\mathbf{x}_k] = [\mathbf{b}_1, \dots, \mathbf{b}_k],$$

ez utóbbi pedig azzal ekvivalens, hogy minden szóbjövő i -re $\mathbf{A}\mathbf{x}_i = \mathbf{b}_i$ lineáris egyenletrendszerek megoldásait keressük, szükséges tehát, hogy mindegyik rendszer megoldható legyen, tehát $\mathbf{b}_1, \dots, \mathbf{b}_k \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$ teljesüljön. Az ilyen mátrixegyenletek legfontosabb alkalmazása az, amikor \mathbf{B} egységmátrix.

26. Definíció. Legyen $\mathbf{A} \in \mathbb{R}^{m \times n}$. Ekkor

1. $\mathbf{A}^{(j)} \in \mathbb{R}^{n \times m}$ jobb oldali inverze \mathbf{A} -nak, ha $\mathbf{A}\mathbf{A}^{(j)} = \mathbf{I}_m$,
2. $\mathbf{A}^{(b)} \in \mathbb{R}^{n \times m}$ bal oldali inverze \mathbf{A} -nak, ha $\mathbf{A}^{(b)}\mathbf{A} = \mathbf{I}_n$,
3. \mathbf{A}^{-1} kétoldali inverze (vagy csak egyszerűen inverze) \mathbf{A} -nak, ha baloldali és jobboldali inverze is \mathbf{A} -nak.

Tekintsük pl. a jobboldali esetet. Ez az $\mathbf{A}\mathbf{X} = \mathbf{I}_m$ mátrixegyenlet megoldását követeli, tehát az $\mathbf{A}\mathbf{x}_i = \mathbf{e}_i$ egyenletrendszereket. Szükséges tehát, hogy $\mathbf{e}_1, \dots, \mathbf{e}_m \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$ teljesüljön. Szerencsénkre $\mathbf{e}_1, \dots, \mathbf{e}_m$ bázis \mathbb{R}^m -ben, így $\text{Span}(\mathbf{e}_1, \dots, \mathbf{e}_m) = \mathbb{R}^m$, így a fenti azzal ekvivalens, hogy

$$\mathbb{R}^m \subseteq \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$$

Vegyük észre, hogy $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) \subseteq \mathbb{R}^m$, hiszen $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$, így a fenti tartalmazás már azzal ekvivalens, hogy $\mathbb{R}^m = \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$, így a dimenziójuknak is meg kell egyezzen, és mivel $\dim \mathbb{R}^m = m$, így azt kaptuk, hogy a megoldhatóság szükséges és elégséges feltétele

$$\dim \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \varrho(\mathbf{A}) = m.$$

Ezzel a most következő fontos tétel első állítását már be is láttuk:

22. Tétel. Legyen $\mathbf{A} \in \mathbb{R}^{m \times n}$. Ekkor

$$1. \exists \mathbf{A}^{(j)} \iff \varrho(\mathbf{A}) = m,$$

$$2. \exists \mathbf{A}^{(b)} \iff \varrho(\mathbf{A}) = n,$$

3. ha léteznek $\mathbf{A}^{(j)}, \mathbf{A}^{(b)}$, akkor $n = m$ és $\mathbf{A}^{-1} = \mathbf{A}^{(j)} = \mathbf{A}^{(b)}$ és ez az inverz egyértelmű.

A második állítás következik az elsőből, hiszen $\exists \mathbf{A}^{(b)} \iff \exists (\mathbf{A}^\top)^{(j)}$. A harmadik állításból az inverzek egyenlőségét kell még belátnunk.

$$\mathbf{A}^{(j)} = \mathbf{I}_n \mathbf{A}^{(j)} = (\mathbf{A}^{(b)} \mathbf{A}) \mathbf{A}^{(j)} = \mathbf{A}^{(b)} (\mathbf{A} \mathbf{A}^{(j)}) = \mathbf{A}^{(b)} \mathbf{I}_m = \mathbf{A}^{(b)},$$

így a dimenziók egyezése is világos. Az egyértelműséghez legyenek \mathbf{A}_1^{-1} és \mathbf{A}_2^{-1} a mátrix inverzei. Ekkor $\mathbf{A}_1^{-1} = \mathbf{A}_1^{-1} \mathbf{I}_n = \mathbf{A}_1^{-1} \mathbf{A} \mathbf{A}_2^{-1} = \mathbf{I}_n \mathbf{A}_2^{-1} = \mathbf{A}_2^{-1}$, ezek tehát szükségképpen megegyeznek.

Legyen most $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n] \in \mathbb{R}^{n \times n}$. Tudjuk, hogy akkor létezik inverz, ha $\varrho(\mathbf{A}) = n$. Emlékszünk, hogy most az $\mathbf{A} \mathbf{x}_i = \mathbf{e}_i$ azonos alakú lineáris egyenletrendszereket kell megoldanunk, és ezt tehetjük egy lépésben is, az elemi bázistranszformáció segítségével. A kiinduló táblázat a következő:

$$\begin{array}{c} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_n \end{array} \left| \begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_n \\ & \mathbf{A} & \\ & & \end{array} \right| \begin{array}{ccc} \mathbf{e}_1 & \dots & \mathbf{e}_n \\ & \mathbf{I}_n & \end{array} \parallel$$

Tegyük fel tehát, hogy $\varrho(\mathbf{A}) = n$, máskülönben nem minden vektort tudunk bevinni a bázisba és mint láttuk, nincs inverz sem. Ha a mátrix teljes rangú, akkor mindent be tudunk vinni, de nem biztos hogy az eredeti sorrendben! Sorcsérékkel kell elérnünk, hogy az indexek jó helyen legyenek. Ezek után kapjuk a következőt:

$$\begin{array}{c} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{array} \left| \begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_n \\ & \mathbf{I}_n & \\ & & \end{array} \right| \begin{array}{ccc} \mathbf{e}_1 & \dots & \mathbf{e}_n \\ & \mathbf{A}^{-1} & \end{array} \parallel$$

így a már jól ismert módszerrel tudunk invertálni. Érdekes itt is kiírni a már bevitt vektorok koordinátáit, hogy a sorcsérékről utólag ne feledkezzünk el.

3.5. Rangtartó átalakítások és alkalmazásai

Említettük már, hogy egy mátrix oszlopainak cseréjétől a rangja nem változik. Ez világos, hiszen egy vektorrendszerben a vektorok sorrendje nem mérvadó. Nyilván a vektorrendszer által generált altér sem változik, az sem függött a vektorok sorrendjétől. Vektorok cseréje tehát ún. rangtartó átalakítás. Akad még jó pár ilyen, ezt egy tételben összefoglaljuk.

23. Tétel. *Legyenek $\mathbf{a}_1, \dots, \mathbf{a}_k \in \mathbb{R}^n$, $0 \neq \lambda \in \mathbb{R}$ és $k \geq 2$. Ekkor*

1. $\varrho(\mathbf{a}_2, \mathbf{a}_1, \dots, \mathbf{a}_k) = \varrho(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$,
2. $\varrho(\lambda \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k) = \varrho(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$,
3. $\varrho(\mathbf{a}_1 + \mathbf{a}_2, \mathbf{a}_2, \dots, \mathbf{a}_k) = \varrho(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$,
4. $\varrho(\mathbf{a}_1 + \lambda \mathbf{a}_2, \mathbf{a}_2, \dots, \mathbf{a}_k) = \varrho(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$.

A fentiek felhasználásával egy adott nullmátrixtól különböző mátrixot véges sok rangtartó átalakítással speciális alakra hozhatunk. Tegyük a következőt: válasszunk ki egy nemnulla elemet (ilyen legalább egy biztos van, hisz feltettük hogy a nullmátrixtól eltérő a mátrixunk), majd ha szükséges sor és oszlopcserékkel a bal felső sarokba visszük. A tétel 2-es állítása végette megtehetjük, hogy az így kapott első oszlopot megszorozzuk a bal felső szám reciprokával, így bal fenn biztosan 1 lesz. Ezzel az egyessel megtehetjük, hogy az első sor minden elemét a második elemtől kezdve kinullázzuk úgy, hogy a második oszloptól kezdve hozzáadjuk az első sor alkalmas konstansszorosát. Ha például az egyes mellett 2-es szerepel, vesszük az első oszlop (-2) -szeresét, majd ezt hozzáadjuk a második oszlophoz, így a kettes helyén már nulla lesz, és így tovább. Fordítva ugyanez megtehető az első oszlop elemeivel a másodiktól kezdve. Folytatjuk az eljárást az első sor és oszlop elhagyásával kapott mátrixon addig, amíg ez nullmátrix, vagy ilyen már nincs is. Jelölje $\mathbf{A} \rightsquigarrow \mathbf{B}$ azt, hogy az \mathbf{A} mátrixból véges sok rangtartó átalakítás segítségével megkapható \mathbf{B} . Most a következőt csináltuk:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & a'_{22} & a'_{23} & \dots & a'_{2n} \\ 0 & a'_{32} & a'_{33} & \dots & a'_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a'_{m2} & a'_{m3} & \dots & a'_{mn} \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & a''_{33} & \dots & a''_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a''_{m3} & \dots & a''_{mn} \end{bmatrix} \rightsquigarrow \dots$$

A vesszők arra utalnak, hogy az ottani érték nagy valószínűséggel nem egyezik a korábbival. A fenti gondolatmenetet követve világos, hogy csak akkor áll le a művelet, ha a „maradék”

mátrixban már csak nulla elem van, mert akkor a reciprokkal szorzás nem elvégezhető, vagy akkor ha ez a rész már nincs is. Így tehát a következőt kapjuk:

$$\mathbf{A} \rightsquigarrow \begin{bmatrix} \mathbf{I}_? & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

A kérdőjel arra utal, hogy nem tudjuk hogy hányszor végezhető el a művelet, mekkora lesz a kapott egységmátrix. Egy tételben megmutatjuk, hogy ezt az értéket már jól ismerjük:

24. Tétel. Legyen $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\varrho(\mathbf{A}) = r \geq 1$. Ekkor

$$\mathbf{A} \rightsquigarrow \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

A tétel valóban értelmes dolgot állít, hisz mivel a sorrang és oszloprang megegyezik, ezért $\varrho(\mathbf{A}) \leq \min(m, n)$ Megmutatható, hogy ezek az átalakítások mátrixszorzásként is felírhatók.

3.6. Nevezetes mátrixok

27. Definíció. Egy $\mathbf{D} \in \mathbb{R}^{n \times n}$ mátrixot diagonálisnak nevezünk, ha $i \neq j$ esetén $d_{ij} = 0$, tehát a mátrix csak a főátlójában tartalmazhat nemnulla elemet:

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix}$$

Ha tudjuk hogy a mátrix diagonális, megadásához elegendő elemeinek felsorolása:

$$\mathbf{D} = \text{diag}(d_1, \dots, d_n).$$

Diagonális mátrixok szorzása egyszerű, mert ebben az esetben csak a megfelelő diagonális elemeket kell összeszorozni.

28. Definíció. Egy mátrix felsőháromszögmátrix, ha a főátló alatti elemek mind nullák. Alsőháromszögmátrix esetén a főátló felett vannak csupa nullák. A mátrix tridiagonális, ha csak a diagonálisban és közvetlen szomszédjokban lehetnek nemnulla elemek:

$$\begin{bmatrix} \beta_1 & \gamma_1 & 0 & \dots & 0 & 0 \\ \alpha_1 & \beta_2 & \gamma_2 & \dots & 0 & 0 \\ 0 & \alpha_2 & \beta_3 & \dots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{n-1} & \beta_n \end{bmatrix}$$

Egy \mathbf{A} mátrix szimmetrikus, ha $\mathbf{A} = \mathbf{A}^\top$, antiszimmetrikus, ha $-\mathbf{A} = \mathbf{A}^\top$. Egy \mathbf{Q} mátrix ortogonális, ha $\mathbf{Q}^\top = \mathbf{Q}^{-1}$ (ezekkel foglalkozunk még sokat). \mathbf{A} projektor, ha $\mathbf{A}^2 = \mathbf{A}$, nilpotens, ha van olyan k egész, hogy $\mathbf{A}^k = \mathbf{0}$, valamint invertálható, ha létezik (kétoldali) inverze.

3.7. Komplex számtest feletti mátrixok

Eddig $\mathbb{R}^{m \times n}$ -beli mátrixokkal foglalkoztunk, most azonban bevezetés gyanánt megismerkedünk a $\mathbb{C}^{m \times n}$ -beli mátrixokkal is. Egy $\lambda \in \mathbb{C}$ -beli számot az algebrai $a + bi$ alakjában reprezentáljuk, ahol $a, b \in \mathbb{R}$. Emlékeztetőül a konjugálás művelete: $\overline{a + bi} = a - bi$. Az eddigi alpműveleteket (összeadás, skalárral szorzás, mátrixszorzás) teljesen analóg módon definiáljuk, hisz ezek a műveletek komplex számokkal is elvégezhetők. Egy új művelettel, az adjungálással ismerkedünk meg:

29. Definíció. Az $\mathbf{A} \in \mathbb{C}^{m \times n}$ adjungáltja $\mathbf{A}^* \in \mathbb{C}^{n \times m}$ úgy, hogy minden szóbjövő j, k -ra $j[\mathbf{A}^*]_k = \overline{k[\mathbf{A}]_j}$, tehát a mátrixot transzponáljuk, majd elemeit konjugáljuk.

Mivel a konjugálás minden valós számot helyben hagy, ezért ez tekinthető a transzponálás általánosításának. Bizonyítás nélkül megismerkedünk a konjugálás és a mátrixműveletek kapcsolatával, mely nem sokban tér el a transzponálásnál látottaktól:

25. Tétel. Teljesülnek a következők:

1. $\forall \mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n} : (\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$,
2. $\forall \mathbf{A} \in \mathbb{C}^{m \times n}, \lambda \in \mathbb{C} : (\lambda \mathbf{A})^* = \bar{\lambda} \mathbf{A}^*$,
3. $\forall \mathbf{A} \in \mathbb{C}^{m \times n}, \mathbf{B} \in \mathbb{C}^{n \times k} : (\mathbf{AB})^* = \mathbf{B}^* \mathbf{A}^*$.

4. fejezet

Geometriai vektorok

Ebben a rövid fejezetben kimondottan a geometriából ismert \mathbb{R}^3 geometriai térrel foglalkozunk. Emlékezzünk, hogy geometriai vektoroknak értelmeztük a hajlásszögét, azaz a kettejük által bezárt szöget, melyet \mathbf{a}, \mathbf{b} vektorok esetén $\gamma(\mathbf{a}, \mathbf{b})$ -vel jelölünk. Világos, hogy ez éppen $[0, \pi]$ -beli érték. Megmutatható néhány egyszerű állítás:

26. Tétel. \mathbf{a}_1 önmagában akkor és csak akkor összefüggő, ha a nullvektor. \mathbf{a}_1 és \mathbf{a}_2 akkor és csak akkor összefüggőek, ha párhuzamosak. Három vektor esetében pedig pontosan akkor, ha egysíkúak (létezik olyan origón átmenő sík, melyen a vektorok elhelyezhetők). \mathbf{a}_1 önmagában akkor és csak akkor független, ha nem a nullvektor. \mathbf{a}_1 és \mathbf{a}_2 akkor és csak akkor függetlenek, ha nem párhuzamosak. Három vektor esetében pedig pontosan akkor, ha nem egysíkúak.

Most bevezetünk egy ismerősen csengő fogalmat, a skaláris szorzatot geometriai vektorokra. Később ezt a fogalmat általánosítjuk \mathbb{R}^n -re és megmutatjuk azt is, hogy léteznek még más skaláris szorzatnak nevezett műveletek is.

30. Definíció. \mathbf{a} és \mathbf{b} geometriai vektorok skaláris szorzata $\mathbf{ab} = |\mathbf{a}||\mathbf{b}| \cos \gamma(\mathbf{a}, \mathbf{b})$.

27. Tétel (Tulajdonságok). Legyenek $\mathbf{a}, \mathbf{b}, \mathbf{c}$ geometriai vektorok, $\lambda \in \mathbb{R}$. Ekkor

1. $\mathbf{ab} = \mathbf{ba}$,
2. $\mathbf{ab} = 0$ akkor és csak akkor, ha \mathbf{a} és \mathbf{b} merőlegesek,
3. $\lambda(\mathbf{ab}) = (\lambda\mathbf{a})\mathbf{b} = \mathbf{a}(\lambda\mathbf{b})$,
4. $\mathbf{a}(\mathbf{bc}) \stackrel{\text{ált.}}{=} (\mathbf{ab})\mathbf{c}$,
5. $\mathbf{cc} = |\mathbf{c}|^2 \geq 0$,

6. $\mathbf{a}(\mathbf{b} + \mathbf{c}) = \mathbf{ab} + \mathbf{ac}$, $(\mathbf{a} + \mathbf{b})\mathbf{c} = \mathbf{ac} + \mathbf{bc}$

A fenti tulajdonságok segítségével nem szükséges ismernünk a bezárt szöget, a skaláris szorzatot pusztán a koordinátákból megkaphatjuk. Legyenek \mathbf{i} , \mathbf{j} , \mathbf{k} páronként merőleges egységvektorok (tehát hosszuk 1), könnyen láthatóan éppen ilyenek az \mathbb{R}^3 -beli kanonikus bázisvektorok. Ez nyilván bázis a geometriai vektorok körében. Tetszőleges \mathbf{a}, \mathbf{b} geometriai vektort véve ebben a bázisban ezek a vektorok felírhatók a következő alakban:

$$[\mathbf{a}]_{\mathbf{i}, \mathbf{j}, \mathbf{k}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \quad [\mathbf{b}]_{\mathbf{i}, \mathbf{j}, \mathbf{k}} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}.$$

Az a tény, hogy ebben a bázisban írtuk fel a vektorokat azt jelenti, hogy

$$\mathbf{a} = \alpha_1 \mathbf{i} + \alpha_2 \mathbf{j} + \alpha_3 \mathbf{k} \text{ és } \mathbf{b} = \beta_1 \mathbf{i} + \beta_2 \mathbf{j} + \beta_3 \mathbf{k}.$$

Tegyük fel, hogy egyik vektor sem nullvektor. A geometriai koszinusztétel segítségével számolhatjuk a bezárt szöget. A távolságot ismerjük:

$$|\mathbf{b} - \mathbf{a}| = \sqrt{(\beta_1 - \alpha_1)^2 + (\beta_2 - \alpha_2)^2 + (\beta_3 - \alpha_3)^2}$$

A koszinusztételt alkalmazva

$$\begin{aligned} \left(\sqrt{(\beta_1 - \alpha_1)^2 + (\beta_2 - \alpha_2)^2 + (\beta_3 - \alpha_3)^2} \right)^2 &= |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}| \cos \gamma(\mathbf{a}, \mathbf{b}) \implies \\ \implies (\beta_1 - \alpha_1)^2 + (\beta_2 - \alpha_2)^2 + (\beta_3 - \alpha_3)^2 &= \\ = \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \beta_1^2 + \beta_2^2 + \beta_3^2 - 2|\mathbf{a}||\mathbf{b}| \cos \gamma(\mathbf{a}, \mathbf{b}) \end{aligned}$$

Innen kifejezhető a skaláris szorzat, először az egyenlet bal oldalát írjuk fel:

$$\begin{aligned} (\beta_1 - \alpha_1)^2 + (\beta_2 - \alpha_2)^2 + (\beta_3 - \alpha_3)^2 &= \\ \beta_1^2 + \beta_2^2 + \beta_3^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 - 2(\beta_1\alpha_1 + \beta_2\alpha_2 + \beta_3\alpha_3) \end{aligned}$$

A négyzetes tagok mindkét oldalon szerepelnek, így azok kiesnek:

$$\begin{aligned} -2(\beta_1\alpha_1 + \beta_2\alpha_2 + \beta_3\alpha_3) &= -2|\mathbf{a}||\mathbf{b}| \cos \gamma(\mathbf{a}, \mathbf{b}) \implies \\ \implies \beta_1\alpha_1 + \beta_2\alpha_2 + \beta_3\alpha_3 &= |\mathbf{a}||\mathbf{b}| \cos \gamma(\mathbf{a}, \mathbf{b}) \implies \mathbf{ab} = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3, \end{aligned}$$

így mindjárt könnyebben számolhatjuk a skaláris szorzatot. A fentiből kiderült az is, hogy

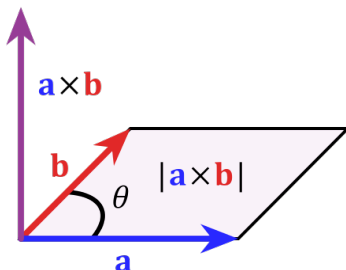
$$\cos \gamma(\mathbf{a}, \mathbf{b}) = \frac{\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3}{\sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} \sqrt{\beta_1^2 + \beta_2^2 + \beta_3^2}}$$

31. Definíció. Az $\mathbf{a}, \mathbf{b}, \mathbf{c}$ nem egysíkú geometriai vektorok ebben a sorrendben jobbrendszer alkotnak, ha közös kezdőpontból felrajzolva őket az \mathbf{a} és \mathbf{b} által alkotott síkra emelt merőlegesen a \mathbf{c} -t tartalmazó féltérből nézve pozitív irányú, 0° és 180° közötti forgatással vihetjük át $\mathbf{a}/|\mathbf{a}|$ -t $\mathbf{b}/|\mathbf{b}|$ -re.

A leosztás csak azért volt szükséges, hogy az átvitel mint transzformáció, értelmes legyen, hisz tetszőleges \mathbf{a} geometriai vektor esetén $|\mathbf{a}|/|\mathbf{a}| = 1$, az irány meg természetesen nem változik. Ez az ún. „normálás” egy speciális esete, később még találkozunk vele.

32. Definíció. Két geometriai vektor vektoriális szorzatát tulajdonságaival definiáljuk. Az $\mathbf{a} \times \mathbf{b}$ vektort az \mathbf{a} és \mathbf{b} vektorok vektoriális szorzatának nevezzük, ha

1. $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin \gamma(\mathbf{a}, \mathbf{b})$,
2. $\mathbf{a} \times \mathbf{b}$ merőleges \mathbf{a} -ra és \mathbf{b} -re is,
3. $|\mathbf{a} \times \mathbf{b}| \neq 0$ esetén $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$ jobbrendszer.



4.1. ábra. A vektoriális szorzat

Már csupán ennyi alapján belátható néhány jó tulajdonság.

28. Tétel (Tulajdonságok). Legyenek $\mathbf{a}, \mathbf{b}, \mathbf{c}$ geometriai vektorok, $\lambda \in \mathbb{R}$. Ekkor

1. $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ akkor és csak akkor, ha \mathbf{a} és \mathbf{b} párhuzamosak;
2. $\mathbf{b} \times \mathbf{a} = -\mathbf{a} \times \mathbf{b}$ ún. antikommutativitás;
3. $\lambda(\mathbf{a} \times \mathbf{b}) = (\lambda\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda\mathbf{b})$;
4. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$, $(\mathbf{b} \times \mathbf{c}) \times \mathbf{a} = \mathbf{b} \times \mathbf{a} + \mathbf{c} \times \mathbf{a}$

A vektoriális szorzat ismeretében gondoljuk végig milyen kapcsolatban áll az \mathbf{i} , \mathbf{j} , \mathbf{k} egységvektorokkal. Tegyük fel, hogy ezek ebben a sorrendben alkotnak jobbrendszer. Ekkor

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \mathbf{j} \times \mathbf{i} = -\mathbf{k}, \mathbf{j} \times \mathbf{k} = \mathbf{i}, \mathbf{k} \times \mathbf{j} = -\mathbf{i}, \mathbf{k} \times \mathbf{i} = \mathbf{j}, \mathbf{i} \times \mathbf{k} = -\mathbf{j},$$

Továbbá mindegyik vektor önmagával vett vektoriális szorzata a nullvektor. Most (a skaláris szorzatnál látottakhoz hasonlóan) vizsgáljuk meg, hogy kapható meg a vektoriális szorzat a koordinátákból. Legyen ismét

$$[\mathbf{a}]_{\mathbf{i},\mathbf{j},\mathbf{k}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \quad [\mathbf{b}]_{\mathbf{i},\mathbf{j},\mathbf{k}} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}.$$

$$\mathbf{a} = \alpha_1 \mathbf{i} + \alpha_2 \mathbf{j} + \alpha_3 \mathbf{k} \text{ és } \mathbf{b} = \beta_1 \mathbf{i} + \beta_2 \mathbf{j} + \beta_3 \mathbf{k}.$$

Ekkor

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= (\alpha_1 \mathbf{i} + \alpha_2 \mathbf{j} + \alpha_3 \mathbf{k}) \times (\beta_1 \mathbf{i} + \beta_2 \mathbf{j} + \beta_3 \mathbf{k}) = \\ &= \underbrace{\alpha_1 \beta_1 (\mathbf{i} \times \mathbf{i})}_{\mathbf{0}} + \underbrace{\alpha_1 \beta_2 (\mathbf{i} \times \mathbf{j})}_{\alpha_1 \beta_2 \mathbf{k}} + \underbrace{\alpha_1 \beta_3 (\mathbf{i} \times \mathbf{k})}_{-\alpha_1 \beta_3 \mathbf{j}} + \\ &+ \underbrace{\alpha_2 \beta_1 (\mathbf{j} \times \mathbf{i})}_{-\alpha_2 \beta_1 \mathbf{k}} + \underbrace{\alpha_2 \beta_2 (\mathbf{j} \times \mathbf{j})}_{\mathbf{0}} + \underbrace{\alpha_2 \beta_3 (\mathbf{j} \times \mathbf{k})}_{\alpha_2 \beta_3 \mathbf{i}} + \\ &+ \underbrace{\alpha_3 \beta_1 (\mathbf{k} \times \mathbf{i})}_{\alpha_3 \beta_1 \mathbf{j}} + \underbrace{\alpha_3 \beta_2 (\mathbf{k} \times \mathbf{j})}_{-\alpha_3 \beta_2 \mathbf{i}} + \underbrace{\alpha_3 \beta_3 (\mathbf{k} \times \mathbf{k})}_{\mathbf{0}} = \\ &= (\alpha_2 \beta_3 - \alpha_3 \beta_2) \mathbf{i} - (\alpha_1 \beta_3 - \alpha_3 \beta_1) \mathbf{j} + (\alpha_1 \beta_2 - \alpha_2 \beta_1) \mathbf{k} \end{aligned}$$

A kiemelés magyarázatra szorulhat, nézzük az első tényezőt hogy kaptuk: a disztributivitást ismerve van olyan tag, hogy $(\alpha_1 \mathbf{i}) \times (\beta_1 \mathbf{i})$, a skalár kiemelhetőségét alkalmazva a második tagra ez éppen $\beta_1((\alpha_1 \mathbf{i}) \times \mathbf{i})$, innen az első tagra pedig $\alpha_1 \beta_1 (\mathbf{i} \times \mathbf{i})$. A fenti eszközökkel könnyen belátható két fontos tétel:

29. Tétel (Kifejtési tétel). *Tetszőleges $\mathbf{a}, \mathbf{b}, \mathbf{c}$ geometriai vektorokra*

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{ac})\mathbf{b} - (\mathbf{bc})\mathbf{a}.$$

30. Tétel (Felcserélési tétel). *Tetszőleges $\mathbf{a}, \mathbf{b}, \mathbf{c}$ geometriai vektorokra $(\mathbf{a} \times \mathbf{b})\mathbf{c} = \mathbf{a}(\mathbf{b} \times \mathbf{c})$.*

33. Definíció. *Az $\mathbf{a}, \mathbf{b}, \mathbf{c}$ geometriai vektorok vegyesszorzata: „ \mathbf{abc} ” = $(\mathbf{a} \times \mathbf{b})\mathbf{c}$.*

Fejtsük ki a vegyesszorzatot a komponensekkel. A következő fejezetben ezt majd kapcsolatba hozzuk egy új fogalommal, az ún. determinánssal. Legyen $\mathbf{A} \in \mathbb{R}^{3 \times 3}$,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Számoljuk ki az oszlopvektorok vegyesszorzatát, tehát a következő vektorokét:

$$[\mathbf{a}]_{i,j,k} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}, \quad [\mathbf{b}]_{i,j,k} = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix}, \quad [\mathbf{c}]_{i,j,k} = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}.$$

Megjegyzés: a mátrix ekkor felírható $\mathbf{A} = [[\mathbf{a}]_{i,j,k}, [\mathbf{b}]_{i,j,k}, [\mathbf{c}]_{i,j,k}]$ alakban. A következő eredményt kapjuk:

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

5. fejezet

Determináns

A determináns egy nagyon fontos, ám nem könnyen értelmezhető fogalom (legalábbis elsőre) a lineáris algebrában. Szeretném ismertetni (egyik) lehetséges értelmezését, mely az eredeti előadásjegyzetben nem szerepel.

5.1. Értelmezés

Tekintsük a jól ismert két egyenletből álló kétismeretlenes lineáris egyenletrendszert:

$$a_{11}x_1 + a_{12}x_2 = b_1$$

$$a_{21}x_1 + a_{22}x_2 = b_2$$

Szorozzuk meg az első egyenletet a_{22} -vel, a másodikat a_{12} -vel:

$$a_{11}a_{22}x_1 + a_{12}a_{22}x_2 = b_1a_{22}$$

$$a_{21}a_{12}x_1 + a_{22}a_{12}x_2 = b_2a_{12}$$

majd vonjuk ki az első egyenletből a másodikat. Az x_2 így kiesik:

$$(a_{11}a_{22} - a_{21}a_{12})x_1 = b_1a_{22} - b_2a_{12} \implies x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}}$$

Hasonlóan ha most x_1 -et akarjuk „eltüntetni”, akkor az első egyenletet a_{21} -el, másodikat a_{11} -el szorozva, majd kivonva (most a második egyenletből az első):

$$a_{11}a_{21}x_1 + a_{12}a_{21}x_2 = b_1a_{21}$$

$$a_{21}a_{11}x_1 + a_{22}a_{11}x_2 = b_2a_{11}$$

$$\implies (a_{22}a_{11} - a_{12}a_{21})x_2 = b_2a_{11} - b_1a_{21} \implies x_2 = \frac{b_2a_{11} - b_1a_{21}}{a_{22}a_{11} - a_{12}a_{21}}$$

A három egyenletből álló háromismeretlenes esetről teljesen hasonló de jóval hosszabb számolás adja, hogy $x_1 = p/q$, ahol

$$p = b_1a_{22}a_{33} + b_3a_{12}a_{23} + b_2a_{13}a_{32} - b_3a_{13}a_{22} - b_1a_{23}a_{32} - b_2a_{12}a_{33},$$

$$q = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Hasonló képlet adható a többi ismeretlenre is. Vegyük észre, hogy a nevező pontosan a korábban definiált vegyesszorozattal egyezik meg. Vizsgáljuk meg a nevezőket, vegyük észre a kapcsolatot a tényezők indexei között. Ha a háromismeretlenes esetet nézzük szépen látszik, hogy egy tényezőben az indexek első számai az 1,2,3, a második számok pedig ezeknek valamilyen permutációja, jól látszik hogy az összes permutáció megjelenik. A szorzatok előjelére még magyarázatot kell adni:

34. Definíció. *Vegyük az $1, 2, \dots, n$ számok valamely i_1, i_2, \dots, i_n permutációját. Azt mondjuk, ebben a permutációban az i_μ és i_ν elemek inverzióban vannak egymással, ha $\mu < \nu$ és $i_\mu > i_\nu$. $I(i_1, \dots, i_n)$ jelölje az i_1, \dots, i_n permutációban az inverzióban lévő párok számát, ezt a számot inverziószámnak hívjuk.*

Vizsgáljuk ismét a fenti p/q hányadosban a kitevőt. Vegyük észre, hogy akkor pozitív egy tényező előjele, ha a második indexek inverziószáma páros, tehát ha az egyik tényező

$$a_{1i_1}a_{2i_2}a_{3i_3}$$

alakú ahol $I(i_1, i_2, i_3)$ páros szám. Negatív az előjel, ha ez a szám páratlan. Például az első tagnál $i_1 = 1, i_2 = 2$ és $i_3 = 3$, jól láthatóan ebben a permutációban nincs inverzió, így az inverziószám 0, az előjel tehát pozitív. A második tagnál $i_1 = 2, i_2 = 3, i_3 = 1$, ebben a permutációban a 2 és a 3 is inverzióban állnak az 1-el, így az inverziószám 2, ami szintén páros. A negyedik tagnál $i_1 = 3, i_2 = 2, i_3 = 1$, itt a három inverzióban van a 2-vel és az 1-el is, valamint a 2 is inverzióban van az 1-el, így az inverziószám 3, az előjel negatív (úgy kell tehát egy szorzatban felírni a tényezők sorrendjét, hogy az első indexek mindig növekvőek legyenek). A fenn tárgyalt előjel éppen $(-1)^{I(i_1, i_2, i_3)}$, tehát egy tényező $(-1)^{I(i_1, i_2, i_3)}a_{1i_1}a_{2i_2}a_{3i_3}$ alakú, és az összegben az összes lehetséges permutáció megjelenik. Ha ezt az értéket egy négyzetes mátrix oszlopvektorai nyomán szeretnénk kiszámolni, úgy jutunk el a determináns fogalmához:

35. Definíció. Legyen $\mathbf{A} \in \mathbb{R}^{n \times n}$ a következő alakú:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

Az \mathbf{A} mátrix determinánása az alább definiált szám:

$$\sum_{\substack{i_1, \dots, i_n \\ (1, \dots, n)}} (-1)^{I(i_1, \dots, i_n)} a_{1i_1} a_{2i_2} \cdots a_{ni_n}.$$

Jelölésben $\det(\mathbf{A})$ vagy $|\mathbf{A}|$, vagy ha szükséges, magát a mátrixot tesszük egyenes vonalak közé:

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \in \mathbb{R}$$

$\mathbf{A} \in \mathbb{R}^{2 \times 2}$ esetén $|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}$. Írjuk fel a 3×3 -as esetet, és kicsit írjuk át:

$$\begin{aligned} |\mathbf{A}| &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} = \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \end{aligned}$$

Vegyük észre a szerkezeti hasonlóságot a vektoriális szorzat kiszámítására adott képlettel. Képezzük az alábbi blokkosított mátrixot:

$$\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{bmatrix}$$

A fenti mátrix bár nem 3×3 -as, mégis blokkosítva „tud úgy viselkedni”, mintha az lenne. A determináns fenti kifejtéséből pontosan az adódik, hogy

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{vmatrix}$$

5.2. Tulajdonságok

Nyilván $|\mathbf{0}| = 0$, továbbá megmutatható, hogy $|\mathbf{A}^\top| = |\mathbf{A}|$. Az utolsó észrevétel felhasználásával azonnal észrevehetjük, hogy a determináns értéke akkor is nulla, ha az egyik sor (vagy

egyik oszlop) minden tagja nulla, hisz ekkor minden szorzatban az egyik tényező biztosan 0. Könnyen megmutatható a determináns és a skalárral szorzás kapcsolata is: A (λA) mátrixban minden tagnak van egy λ -szorosa, így az n -tagú szorzatokból λ^n kiemelhető, ezek pedig az összeg elé vihetők, így $|\lambda A| = \lambda^n |A|$. Ezt a teljesség kedvéért formálisan is megmutatjuk:

$$\begin{aligned} |\lambda A| &= \sum_{\substack{i_1, \dots, i_n \\ (1, \dots, n)}} (-1)^{I(i_1, \dots, i_n)} \lambda a_{1i_1} \lambda a_{2i_2} \cdots \lambda a_{ni_n} = \\ &= \sum_{\substack{i_1, \dots, i_n \\ (1, \dots, n)}} (-1)^{I(i_1, \dots, i_n)} \lambda^n a_{1i_1} a_{2i_2} \cdots a_{ni_n} = \lambda^n \left(\sum_{\substack{i_1, \dots, i_n \\ (1, \dots, n)}} (-1)^{I(i_1, \dots, i_n)} a_{1i_1} a_{2i_2} \cdots a_{ni_n} \right) = \lambda^n |A|. \end{aligned}$$

Hasonlóan mutatható meg, hogy ha csak egy sort (vagy oszlopot) szorzunk λ -val, akkor a determináns csak λ -szoros lesz. Sajnos általában $|A + B| \neq |A| + |B|$. Ennek ellenére megmutatható, hogy ha egy mátrixban az egyik sorban minden elemet kéttagú összegként írunk fel, akkor a determináns „jól” viselkedik:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_{k1} + c_{k1} & b_{k2} + c_{k2} & \cdots & b_{kn} + c_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Bizonyítás nélkül közöljük az inverziószám és determináns néhány kapcsolatát.

31. Tétel. *Legyen $n \geq 2$.*

1. *Ha az $1, \dots, n$ számok egy i_1, \dots, i_n permutációjában két elemet felcserélünk, akkor az inverziószám páratlan számmal változik;*
2. *Ha az $A \in \mathbb{R}^{n \times n}$ mátrix két sorát felcseréljük, a determináns (-1) -szeresére változik.*

Bizonyítjuk viszont a fentieknek két fontos következményét:

32. Tétel. *Legyen $n \geq 2$. Ha az $A \in \mathbb{R}^{n \times n}$ mátrixnak van két megegyező sora, akkor a determinánása 0.*

Bizonyítás. Tudjuk, hogy sorok cseréjétől a determináns (-1) -szeres lesz, másrészt az azonos sorok cseréje a mátrixot nem változtatja, tehát $|A| = -|A|$, ez pedig csak úgy lehetséges, ha $|A| = 0$. □

33. Tétel. Legyen $n \geq 2$ és $\lambda \in \mathbb{R}$. Ha az $\mathbf{A} \in \mathbb{R}^{n \times n}$ mátrix egyik sorához hozzáadjuk egy másik sor λ -szorosát, akkor a determináns értéke nem változik, ez a művelet tehát azon felül hogy rangtartó, még determinánstartó is.

Bizonyítás. Jelölje ${}_k[\mathbf{A}]$ az \mathbf{A} mátrix k -adik sorát. Nézzük azt az esetet, amikor az első sorhoz adjuk a második sor λ -szorosát. Ekkor

$$\begin{vmatrix} {}_1[\mathbf{A}] + \lambda {}_2[\mathbf{A}] \\ {}_2[\mathbf{A}] \\ \vdots \\ {}_k[\mathbf{A}] \end{vmatrix} = \begin{vmatrix} {}_1[\mathbf{A}] \\ {}_2[\mathbf{A}] \\ \vdots \\ {}_k[\mathbf{A}] \end{vmatrix} + \begin{vmatrix} \lambda {}_2[\mathbf{A}] \\ {}_2[\mathbf{A}] \\ \vdots \\ {}_k[\mathbf{A}] \end{vmatrix} = \begin{vmatrix} {}_1[\mathbf{A}] \\ {}_2[\mathbf{A}] \\ \vdots \\ {}_k[\mathbf{A}] \end{vmatrix} + \lambda \begin{vmatrix} {}_2[\mathbf{A}] \\ {}_2[\mathbf{A}] \\ \vdots \\ {}_k[\mathbf{A}] \end{vmatrix} = \begin{vmatrix} {}_1[\mathbf{A}] \\ {}_2[\mathbf{A}] \\ \vdots \\ {}_k[\mathbf{A}] \end{vmatrix} + \lambda \cdot \mathbf{0} = |\mathbf{A}|,$$

és ezzel készen vagyunk. Az általános eset is teljesen hasonló. \square

A determináns felírása már egy kicsit nagyobb mátrix esetén is fárasztó művelet. Hamarosan eljutunk odáig, hogy viszonylag könnyen felírhatjuk ezt a képletet igen nagy mátrixokra is, ehhez azonban még egy kicsit vizsgálódnunk kell. Tudjuk, hogy ha egy sorban végig nulla van, akkor a determináns értéke is nulla. Tekintsük most azt a speciális esetet, amikor egy sorban egy elem kivételével az összes többi nulla, és erre írjuk fel a definíciót:

$$\begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & a_{1n} \\ \vdots & & \vdots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \cdots & 0 & a_{nn} \end{vmatrix} =$$

$$= \sum_{\substack{i_1, \dots, i_{n-1}, i_n \\ (1, \dots, n-1, n)}} (-1)^{I(i_1, i_2, \dots, i_{n-1}, i_n)} a_{1i_1} a_{2i_2} \cdots a_{n-1, i_{n-1}} a_{ni_n}.$$

Látjuk, hogy a mátrix n -edik sorában a_{nn} kivételével minden tag nulla, így a fenti a következőképp is írható a kieső tagok miatt:

$$\sum_{\substack{i_1, \dots, i_{n-1}, i_n \\ (1, \dots, n-1, n)}} (-1)^{I(i_1, i_2, \dots, i_{n-1}, i_n)} a_{1i_1} a_{2i_2} \cdots a_{n-1, i_{n-1}} a_{nn}.$$

Ekkor viszont a szorzat utolsó tagja mindig ugyanaz, független a permutációtól, így kiemelhető az összegből és elég $n-1$ -ig permutálni:

$$a_{nn} \sum_{\substack{i_1, \dots, i_{n-1} \\ (1, \dots, n-1)}} (-1)^{I(i_1, i_2, \dots, i_{n-1})} a_{1i_1} a_{2i_2} \cdots a_{n-1, i_{n-1}}$$

Az összeg most éppen a bal felső $(n-1) \times (n-1)$ -es részmátrix determinánsa. Ezzel beláttuk a következő tételt:

34. Tétel.

$$\begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & a_{1n} \\ \vdots & & \vdots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \cdots & 0 & a_{nn} \end{vmatrix} = a_{nn} \cdot \begin{vmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{vmatrix}$$

Ha ezt egy felsőháromszög mátrixra alkalmazzuk rekurzív módon ameddig csak tudjuk azt a következményt kapjuk, hogy ilyen mátrixok determinánsa éppen a főátlóban lévő elemek szorzata:

35. Tétel.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix} = a_{11}a_{22} \cdot \dots \cdot a_{nn}$$

Mivel a transzponálás a determináns értékén nem változtat, a fenti állítás érvényes alsóháromszög mátrixokra is, továbbá nyilván diagonális mátrixokra is, így tehát $|\mathbf{I}_n| = 1$.

5.3. A determináns értékének kiszámítása elemi bázis-transzformációval

Legyen $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n] \in \mathbb{R}^{n \times n}$. Tudjuk, hogy ha $\varrho(\mathbf{A}) < n$, akkor $\mathbf{a}_1, \dots, \mathbf{a}_n$ összefüggő, más szóval létezik olyan \mathbf{a}_j , mely lineárisan függ a többitől, tehát

$$\mathbf{a}_j = \alpha_1 \mathbf{a}_1 + \cdots + \alpha_{j-1} \mathbf{a}_{j-1} + \alpha_{j+1} \mathbf{a}_{j+1} + \cdots + \alpha_n \mathbf{a}_n.$$

A j -edik oszlophoz ha hozzáadjuk az első oszlop $-\alpha_1$ -szeresét, majd a második oszlop $-\alpha_2$ -szeresét és így tovább, a determináns értéke nem változik, ám a j -edik oszlopban csupa nulla lesz, így a determináns értéke is nulla. Ezzel megmutattuk, hogy $\varrho(\mathbf{A}) < n$ esetén $|\mathbf{A}| = 0$. Amennyiben $\varrho(\mathbf{A}) = n$, úgy minden vektort be tudunk vinni a bázisba, ám természetesen lehet, hogy nem az eredeti sorrendben. Nézzük a módszert, legyen az eredeti mátrix

$$\begin{bmatrix} \beta_{11} & \cdots & \beta_{1j} & \cdots & \beta_{1n} \\ \vdots & & \vdots & & \vdots \\ \beta_{i1} & \cdots & \beta_{ij} & \cdots & \beta_{in} \\ \vdots & & \vdots & & \vdots \\ \beta_{n1} & \cdots & \beta_{nj} & \cdots & \beta_{nn} \end{bmatrix}$$

Tegyük fel, hogy \mathbf{a}_j -t akarjuk bevinni \mathbf{e}_i helyére. Tudjuk hogy ez lehetséges, ha $\beta_{ij} \neq 0$. Ezt az elemet generáloelemnek fogjuk nevezni, és a k -adik lépésben g_k -val fogjuk jelölni, most tehát $g_1 = \beta_{ij}$. Képezzük a következő segédmátrixot: az eredeti mátrix i -edik sorát szorozzuk β_{ij}^{-1} -gyel, vagyis minden elemet osszunk le β_{ij} -vel az i -edik sorban:

$$\begin{bmatrix} \beta_{11} & \cdots & \beta_{1j} & \cdots & \beta_{1n} \\ \vdots & & \vdots & & \vdots \\ \beta_{i1}/\beta_{ij} & \cdots & 1 & \cdots & \beta_{in}/\beta_{ij} \\ \vdots & & \vdots & & \vdots \\ \beta_{n1} & \cdots & \beta_{nj} & \cdots & \beta_{nn} \end{bmatrix}$$

Tudjuk azt, hogy ennek az új mátrixnak a determinánsa az eredeti mátrix determinánsának $1/\beta_{ij}$ -szerese, más szóval az eredeti mátrix determinánsa az új mátrix determinánsának β_{ij} -szerese. Ezután a mátrix j -edik oszlopában az egyesen kívüli elemek kinullázhatók, ha minden k -adik sorhoz ($k \neq i$) hozzáadjuk az i -edik sor $-\beta_{kj}$ -szeresét. Ezzel a determináns értéke nem változik.

$$\begin{bmatrix} \beta_{11} - \beta_{1j} \cdot \beta_{i1}/\beta_{ij} & \cdots & 0 & \cdots & \beta_{1n} - \beta_{1j} \cdot \beta_{in}/\beta_{ij} \\ \vdots & & \vdots & & \vdots \\ \beta_{i1}/\beta_{ij} & \cdots & 1 & \cdots & \beta_{in}/\beta_{ij} \\ \vdots & & \vdots & & \vdots \\ \beta_{n1} - \beta_{nj} \cdot \beta_{i1}/\beta_{ij} & \cdots & 0 & \cdots & \beta_{nn} - \beta_{nj} \cdot \beta_{in}/\beta_{ij} \end{bmatrix}$$

Mivel a determináns most nem változott, ezért az eredeti mátrix determinánsa még mindig ennek az újnak a β_{ij} -szerese, tehát g_1 -szerese. Mivel $\varrho(\mathbf{A}) = n$, így ezt a lépést n -szer elvégezhetjük, az eredeti mátrix determinánsa a legvégül kapott mátrix determinánsának $g_1 g_2 \cdots g_n$ -szerese lesz. Ha nem sikerült az eredeti sorrendben bevinni a vektorokat, akkor sorcserékkel elérhetjük, hogy a kapott mátrix \mathbf{I}_n legyen, melynek determinánsa 1, itt azonban vigyáznunk kell a sorok cseréjével a determináns előjele miatt. Jelölje \mathbf{U} a kapott mátrixot. Ennek determinánsa 1 vagy -1 , hiszen ez az egységmátrix sorainak cseréjével kapható. Ha a vektorok sorrendje a bázisban $\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_n}$, tekintsük a $(-1)^{I(i_1, \dots, i_n)} |\mathbf{U}|$ szorzatot. Ekkor ha két sort megcserélünk, akkor az inverziószám páratlan változása végett $(-1)^{I(i_1, \dots, i_n)}$ előjele megváltozik, ám nyilván $|\mathbf{U}|$ előjele is, a fenti szorzat értéke tehát nem változik, tehát ez az érték megegyezik $(-1)^{I(1, \dots, n)} |\mathbf{I}_n|$ -nel, mely utóbbi éppen 1, vagyis $|\mathbf{U}| = (-1)^{I(i_1, \dots, i_n)}$. Mivel az eredeti mátrix determinánsa ennek a $g_1 \cdots g_n$ -szerese, így

$$|\mathbf{A}| = (-1)^{I(i_1, \dots, i_n)} g_1 \cdots g_n.$$

Fontos és könnyítő észrevétel továbbá, hogy a fenti átalakítások éppen a transzformáció lépései, így tehát csak a generáloelemeket kell feljegyeznünk minden lépésben, és az előjelre

figyelni. Nézzünk egy példát, legyen

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{3 \times 3}.$$

Végezzük el a transzformációt, szándékosan „nem szép” sorrendben cserélve a vektorokat. Minden lépésben bekeretezzük a generálóelemet.

$$\begin{array}{c} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{array} \left| \begin{array}{ccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ 1 & 2 & \boxed{1} \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{array} \right\| \begin{array}{c} \mathbf{a}_3 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{array} \left| \begin{array}{ccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ 1 & 2 & 1 \\ 0 & -3 & 0 \\ \boxed{1} & 0 & 0 \end{array} \right\| \begin{array}{c} \mathbf{a}_3 \\ \mathbf{e}_2 \\ \mathbf{a}_1 \end{array} \left| \begin{array}{ccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ 0 & 2 & 1 \\ 0 & \boxed{-3} & 0 \\ 1 & 0 & 0 \end{array} \right\| \begin{array}{c} \mathbf{a}_3 \\ \mathbf{a}_2 \\ \mathbf{a}_1 \end{array} \left| \begin{array}{ccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{array} \right\|$$

A generálóelemek $g_1 = 1$, $g_2 = 1$ és $g_3 = -3$. A bázisvektorok sorrendje $\mathbf{a}_3, \mathbf{a}_2, \mathbf{a}_1$, innen pedig $I(3,2,1) = 3$, tehát

$$|\mathbf{A}| = (-1)^3 \cdot 1 \cdot 1 \cdot (-3) = 3.$$

Egy nagyon fontos következményt kapunk:

$$|\mathbf{A}| \neq 0 \iff \varrho(\mathbf{A}) = n \iff \exists \mathbf{A}^{-1}.$$

5.4. Aldetermináns

Emlékezzünk vissza arra az esetre, amikor az utolsó sorban az utolsó elem kivételével mindenhol nulla volt. Ezt az esetet az eddigi ismeretekkel általánosíthatjuk arra az esetre, amikor a mátrix i -edik sorában végig nulla van, kivéve a j -edik elemet. Sorcserékkel ezt a sort átvihetjük az utolsóba, mely éppen $n - i$ db cserét jelent, így a determináns értéke az eredeti $(-1)^{n-i}$ -szerese lesz. Ezek után oszlopok cseréjével ezt a nemnulla elemet kivihetjük a bal alsó sarokba, mely $n - j$ db oszlopcserét jelent, így a determináns $(-1)^{n-i} \cdot (-1)^{n-j}$ -szerese lesz az eredetinek, ami pedig:

$$(-1)^{n-i} \cdot (-1)^{n-j} = (-1)^{2n-i-j} = \underbrace{(-1)^{2n}}_1 \cdot (-1)^{-i-j} = \underbrace{(-1)^{-2(i+j)}}_1 \cdot (-1)^{i+j} = (-1)^{i+j}$$

A következő felírásban $*$ jelöli az olyan egy oszlopos blokkokat, melyek lényegtelenek az eredmény szempontjából:

$$\left| \begin{array}{cccccc} \mathbf{E} & * & \mathbf{F} & & & \\ 0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \\ & \mathbf{G} & * & \mathbf{H} & & & \end{array} \right| = (-1)^{n-i} \left| \begin{array}{cccccc} \mathbf{E} & * & \mathbf{F} & & & \\ & \mathbf{G} & * & \mathbf{H} & & \\ 0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \end{array} \right| =$$

$$= (-1)^{n-i} \cdot (-1)^{n-j} \begin{vmatrix} \mathbf{E} & \mathbf{F} & * \\ \mathbf{G} & \mathbf{H} & * \\ \mathbf{0} & \mathbf{0} & a_{ij} \end{vmatrix} = a_{ij} \cdot (-1)^{i+j} \begin{vmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{vmatrix}$$

Természetesen a fenti eredményt ismerve nem kell a cseréket végrehajtani, ha felismerjük a mátrixban a $\mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{H}$ blokkokat, és összezsúsztatjuk, pl.:

$$\begin{vmatrix} 1 & 2 & 8 & 2 & 5 \\ 5 & 3 & 6 & 4 & 2 \\ 0 & 0 & 0 & 2 & 0 \\ 4 & 5 & 7 & 1 & 3 \\ 8 & 2 & 5 & 4 & 1 \end{vmatrix} = 2 \cdot (-1)^{3+4} \cdot \begin{vmatrix} 1 & 2 & 8 & 5 \\ 5 & 3 & 6 & 2 \\ 4 & 5 & 7 & 3 \\ 8 & 2 & 5 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & 8 & 5 \\ 5 & 3 & 6 & 2 \\ 4 & 5 & 7 & 3 \\ 8 & 2 & 5 & 1 \end{vmatrix}$$

Ilyen blokkos megközelítéssel definiáljuk az aldetermináns fogalmát:

36. Definíció. Legyen $\mathbf{A} \in \mathbb{R}^{n \times n}$ a következő alakú:

$$\mathbf{A} = \begin{bmatrix} \mathbf{E} & * & \mathbf{F} \\ a_{i1} & \cdots & a_{i,j-1} & a_{ij} & a_{i,j+1} & \cdots & a_{in} \\ \mathbf{G} & * & \mathbf{H} \end{bmatrix} \in \mathbb{R}^{n \times n}$$

Ekkor a mátrix a_{ij} eleméhez tartozó aldeterminánsa

$$\text{aldet}_{ij}(\mathbf{A}) = \begin{vmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{vmatrix}$$

Ugyanehhez az elemhez tartozó előjelezett aldetermináns pedig

$$\mathbf{A}_{ij} = (-1)^{i+j} \text{aldet}_{ij}(\mathbf{A}) = (-1)^{i+j} \begin{vmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{vmatrix}$$

Eljött az idő, hogy kimondjuk a determináns felírásának megkönnyítésére szolgáló tételt:

36. Tétel (Kifejtési tétel). Legyen $n \geq 2$, $\mathbf{A} \in \mathbb{R}^{n \times n}$. Ekkor

a) Tetszőleges $1 \leq i \leq n$ esetén

$$|\mathbf{A}| = \sum_{j=1}^n a_{ij} \mathbf{A}_{ij}$$

b) Tetszőleges $1 \leq j \leq n$ esetén

$$|\mathbf{A}| = \sum_{i=1}^n a_{ij} \mathbf{A}_{ij}.$$

Bizonyítás. Csak az a) esetet bizonyítjuk, mert a másik vele teljesen analóg. Legyen $1 \leq i \leq n$ rögzített. Ekkor

$$\begin{aligned}
& \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = \\
& = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} + 0 + \cdots + 0 & & 0 + \cdots + a_{ij} + \cdots + 0 & & 0 + \cdots + 0 + a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = \\
& = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & a_{ij} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} + \cdots \\
& \quad \cdots + \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}
\end{aligned}$$

A fenti alakot már ismerjük, ki tudjuk fejezni az i -edik sorban szereplő nemnulla elemek segítségével, melyet általánosan így írhatunk fel:

$$\begin{aligned}
& \sum_{j=1}^n \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & a_{ij} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = \\
& = \sum_{j=1}^n |\mathbf{B}_j| = \sum_{j=1}^n a_{ij} (-1)^{i+j} \text{aldet}_{ij}(\mathbf{B}_j) = \sum_{j=1}^n a_{ij} (-1)^{i+j} \text{aldet}_{ij}(\mathbf{A}) = \sum_{j=1}^n a_{ij} \mathbf{A}_{ij}.
\end{aligned}$$

$\text{aldet}_{ij}(\mathbf{B}_j) = \text{aldet}_{ij}(\mathbf{A})$, hiszen csak az i . sor és j . oszlopot leszámítva a két mátrix megegyezik. Ezzel készen is vagyunk. \square

A fenti tétel alkalmazását sakktábla-szabálynak is szokás hívni az előjelek váltakozása miatt. Írjunk fel egy mátrix determinánsát a tétel segítségével az első sor szerint:

$$\begin{vmatrix} 1 & 2 & 8 \\ 5 & 3 & 6 \\ 5 & 4 & 7 \end{vmatrix} = 1 \cdot \begin{vmatrix} 3 & 6 \\ 4 & 7 \end{vmatrix} - 2 \cdot \begin{vmatrix} 5 & 6 \\ 5 & 7 \end{vmatrix} + 8 \cdot \begin{vmatrix} 5 & 3 \\ 5 & 4 \end{vmatrix} =$$

$$= 1 \cdot (3 \cdot 7 - 6 \cdot 4) - 2 \cdot (5 \cdot 7 - 6 \cdot 5) + 8 \cdot (5 \cdot 4 - 3 \cdot 5) = 27.$$

Bizonyítás nélkül kimondjuk a nevezetes Cramer-szabályt:

37. Tétel (Cramer-szabály). Legyen $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n] \in \mathbb{R}^{n \times n}$, $|\mathbf{A}| \neq 0$ és $\mathbf{b} \in \mathbb{R}^n$. Ekkor egyrészt egyértelműen létezik olyan $\mathbf{x} \in \mathbb{R}^n$, melyre $\mathbf{Ax} = \mathbf{b}$, másrészt az \mathbf{x} j -edik komponense a következőképp számolható ($j = 1, \dots, n$):

$$x_j = \frac{\det([\mathbf{a}_1, \dots, \mathbf{b}, \dots, \mathbf{a}_n])}{\det([\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n])}.$$

A későbbi félévekben találkozunk majd az ún. Vandermonde-mátrixszal, melynek sajátos alakja mellett a determinánsának kifejtése is említést érdemel:

38. Tétel (Vandermonde-determináns). Legyen $n \geq 2$, $a_1, \dots, a_n \in \mathbb{R}$. Ekkor a Vandermonde-mátrix determinánsa a következő:

$$\mathbf{V}_n(a_1, \dots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{n \geq i > j \geq 1} (a_i - a_j).$$

Bizonyítás. $n = 2$ -re a következőt kapjuk:

$$\begin{vmatrix} 1 & a_1 \\ 1 & a_2 \end{vmatrix} = a_2 - a_1,$$

ez tehát rendben van. $n \geq 3$ esetén tegyük azt, hogy $\mathbf{V}_n(a_1, \dots, a_n)$ definíció szerinti alakjában vonjuk ki jobbról bal felé haladva minden oszlopból az öt megelőző oszlop a_1 -szeresét. A determináns jobb alsó sarkában jelezzük az aktuális mátrix dimenzióját:

$$\mathbf{V}_n(a_1, \dots, a_n) = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a_2 - a_1 & (a_2 - a_1)a_2 & \cdots & (a_2 - a_1)a_2^{n-2} \\ 1 & a_3 - a_1 & (a_3 - a_1)a_3 & \cdots & (a_3 - a_1)a_3^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n - a_1 & (a_n - a_1)a_n & \cdots & (a_n - a_1)a_n^{n-2} \end{vmatrix}_n =$$

$$= \begin{vmatrix} a_2 - a_1 & (a_2 - a_1)a_2 & \cdots & (a_2 - a_1)a_2^{n-2} \\ a_3 - a_1 & (a_3 - a_1)a_3 & \cdots & (a_3 - a_1)a_3^{n-2} \\ \vdots & \vdots & & \vdots \\ a_n - a_1 & (a_n - a_1)a_n & \cdots & (a_n - a_1)a_n^{n-2} \end{vmatrix}_{n-1}$$

Most azt látjuk, hogy a $\mathbf{V}_{n-1}(a_1, \dots, a_{n-1})$ mátrix első sorát $(a_2 - a_1)$ -gyel, második sorát $(a_3 - a_1)$ -gyel stb. szoroztuk, ismerjük már, hogy ez miképp hat a determinánsra, így az egyenlet tovább folytatható a következőképp:

$$\begin{aligned} & (a_2 - a_1) \cdot \dots \cdot (a_n - a_1) \mathbf{V}_{n-1}(a_1, \dots, a_{n-1}) = \\ & = (a_2 - a_1) \cdot \dots \cdot (a_n - a_1) \prod_{n-1 \geq i > j \geq 1} (a_i - a_j) = \prod_{n \geq i > j \geq 1} (a_i - a_j) \end{aligned}$$

és készen vagyunk. □

5.5. Kapcsolat a lineáris egyenletrendszerekkel

39. Tétel. Legyen $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\rho(\mathbf{A}) = r \geq 1$. Ekkor \mathbf{A} -nak van olyan $r \times r$ -es részmátrixa, melynek determinánsa nem nulla, viszont az összes $(r+1) \times (r+1)$ -es részmátrix determinánsa 0.

Bizonyítás. Mivel a rang r , így van r db lineárisan független oszlopvektor. Hagyjuk meg ezeket az oszlopokat, a többi „dobjuk el”, így kapunk egy $n \times r$ -es részmátrixot. Ennek a részmátrixnak is r a rangja, és így persze a sorrangja is, tehát kiválaszthatunk belőle egy $r \times r$ -es részmátrixot, melynek rangja szintúgy r , ekkor pedig ennek a determinánsa nem lehet 0. Ha most veszünk egy tetszőleges $(r+1) \times (r+1)$ -es részmátrixot tudjuk, hogy ezeknek oszlopai összefüggő rendszert alkotnak, hisz a rang r , ekkor viszont a determináns 0. □

Most a homogén lineáris egyenletrendszerek nemtriviális megoldásának létezésére adunk szükséges és elégséges feltételt a determináns segítségével.

40. Tétel. Legyen $\mathbf{A} \in \mathbb{R}^{n \times n}$. Az $\mathbf{Ax} = \mathbf{0}$ ún. homogén lineáris egyenletrendszernek akkor és csak akkor létezik nemtriviális (nullvektortól különböző) megoldása, ha $|\mathbf{A}| = 0$.

Bizonyítás. Ha a determináns 0, akkor a rang nem lehet teljes, \mathbf{A} oszlopai pedig összefüggő rendszert alkotnak, ami pedig pont azt jelenti, hogy létezik olyan $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$, melyre $\mathbf{Ax} = \mathbf{0}$. □

Következzen egy számolást nagyban könnyítő tétel:

41. Tétel (Szorzástétel). Ha $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$, akkor $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$.

Bizonyítás.

$$\begin{aligned}
 |\mathbf{AB}| &= \left| \begin{array}{ccc} \sum_{k_1=1}^n a_{1k_1} b_{k_1 1} & \cdots & \sum_{k_1=1}^n a_{1k_1} b_{k_1 n} \\ \vdots & & \vdots \\ \sum_{k_n=1}^n a_{nk_n} b_{k_n 1} & \cdots & \sum_{k_n=1}^n a_{nk_n} b_{k_n n} \end{array} \right| = \\
 &= \sum_{k_n=1}^n \cdots \sum_{k_1=1}^n \left| \begin{array}{ccc} a_{1k_1} b_{k_1 1} & \cdots & a_{1k_1} b_{k_1 n} \\ \vdots & & \vdots \\ a_{nk_n} b_{k_n 1} & \cdots & a_{nk_n} b_{k_n n} \end{array} \right| = \\
 &= \sum_{k_n=1}^n \cdots \sum_{k_1=1}^n a_{1k_1} \cdots a_{nk_n} \left| \begin{array}{ccc} b_{k_1 1} & \cdots & b_{k_1 n} \\ \vdots & & \vdots \\ b_{k_n 1} & \cdots & b_{k_n n} \end{array} \right| = \sum_{k_n=1}^n \cdots \sum_{k_1=1}^n a_{1k_1} \cdots a_{nk_n} \left| \begin{array}{c} k_1[\mathbf{B}] \\ \vdots \\ k_n[\mathbf{B}] \end{array} \right|
 \end{aligned}$$

Vegyük észre, hogy az összegben k_1, \dots, k_n között lehetnek egyenlőek is, ezek pedig biztos nullát adnak, hisz ekkor a jobb oldali determinánsnak lesz két egyező sora, mely így 0. Ekkor mi történik? Ha a közös tagokat kiejtjük, már permutáció segítségével írhatjuk fel a fenti összeget. Vigyük be továbbá a (-1) -hatványokat a szorzatba kétszer, ezzel az értéken nem változtatunk hisz $1 \cdot 1 = 1$ és $(-1) \cdot (-1) = 1$.

$$= \sum_{\substack{k_1, \dots, k_n \\ (1, \dots, n)}} (-1)^{I(k_1, \dots, k_n)} a_{1k_1} \cdots a_{nk_n} (-1)^{I(k_1, \dots, k_n)} \left| \begin{array}{c} k_1[\mathbf{B}] \\ \vdots \\ k_n[\mathbf{B}] \end{array} \right|$$

Láttuk már korábban, hogy

$$(-1)^{I(k_1, \dots, k_n)} \left| \begin{array}{c} k_1[\mathbf{B}] \\ \vdots \\ k_n[\mathbf{B}] \end{array} \right| = |\mathbf{B}|$$

tehát tovább írva

$$\begin{aligned}
 &= \sum_{\substack{k_1, \dots, k_n \\ (1, \dots, n)}} (-1)^{I(k_1, \dots, k_n)} a_{1k_1} \cdots a_{nk_n} |\mathbf{B}| = \\
 &= \left(\sum_{\substack{k_1, \dots, k_n \\ (1, \dots, n)}} (-1)^{I(k_1, \dots, k_n)} a_{1k_1} \cdots a_{nk_n} \right) |\mathbf{B}| = |\mathbf{A}||\mathbf{B}|,
 \end{aligned}$$

és ezzel végre készen vagyunk. □

6. fejezet

Sajátérték és sajátvektor

6.1. Diagonalizálhatóság

Mind bizonyos problémák megértésében, mind a gyakorlati alkalmazásokban fontos lehet, hogy egy adott mátrix oszlopvektorait milyen bázisban írjuk fel. Nagyon egyszerű például azoknak a mátrixoknak (és később az általuk képzett lineáris leképezéseknek) a kezelése, amelyek valamely bázisban diagonálisak.

37. Definíció. Azt mondjuk $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ esetén \mathbf{A} hasonló \mathbb{R} felett \mathbf{B} -hez (jelölésben $\mathbf{A} \sim_{\mathbb{R}} \mathbf{B}$), ha létezik olyan invertálható $\mathbf{D} \in \mathbb{R}^{n \times n}$, melyre $\mathbf{B} = \mathbf{D}^{-1} \mathbf{A} \mathbf{D}$.

Könnyen megmutatható, hogy a hasonlóság szimmetrikus reláció, tehát ha $\mathbf{A} \sim_{\mathbb{R}} \mathbf{B}$, akkor $\mathbf{B} \sim_{\mathbb{R}} \mathbf{A}$ is teljesül, így beszélhetünk egyszerűen mátrixok hasonlóságáról. Különösen kedvező az a helyzet, amikor egy mátrix valamely diagonális mátrixhoz hasonló.

38. Definíció. Azt mondjuk $\mathbf{A} \in \mathbb{R}^{n \times n}$ mátrix diagonalizálható, ha \mathbb{R} felett hasonló valamely diagonális mátrixhoz.

A legtöbb tudományágban mely mátrixokkal dolgozik, általában a diagonális mátrixok előnyt élveznek alkalmazhatóság szempontjából, de már csak az is előnyös tud lenni, ha egy mátrix „majdnem” diagonális, például tridiagonális, vagy a diagonális elemek közelében nagyon kicsi értékek szerepelnek. Tekintsünk egy nagyon primitív kis alkalmazást, például adott „nem szép” \mathbf{A} mátrixot kell valamely k -adik hatványra emelnünk. Amennyiben \mathbf{A} diagonalizálható, úgy létezik egy olyan invertálható \mathbf{D} mátrix, mellyel

$$\mathbf{B} = \mathbf{D}^{-1} \mathbf{A} \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Ezt a mátrixot „négyzetre emelve”:

$$\mathbf{D}^{-1}\mathbf{A}\mathbf{D}\mathbf{D}^{-1}\mathbf{A}\mathbf{D} = \mathbf{D}^{-1}\mathbf{A}\mathbf{A}\mathbf{D} = \mathbf{D}^{-1}\mathbf{A}^2\mathbf{D}$$

$$\text{Hasonlóan } (\mathbf{D}^{-1}\mathbf{A}\mathbf{D})^k = \mathbf{D}^{-1}\mathbf{A}^k\mathbf{D} \implies \mathbf{A}^k = \mathbf{D}\mathbf{B}^k\mathbf{D}^{-1} =$$

$$= \mathbf{D} \left(\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \right)^k \mathbf{D}^{-1} = \mathbf{D} \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix} \mathbf{D}^{-1},$$

így tehát a hatvány könnyedén meghatározható, azonban „szerencsénknek kellett legyen”, hiszen a mátrix diagonalizálható volt. Mit is jelent ez? Tekintsük ismét a fenti diagonális \mathbf{B} mátrixot, és vegyük \mathbb{R}^n triviális $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázisát. Könnyen láthatóan ekkor $\mathbf{B}\mathbf{e}_1 = \lambda_1\mathbf{e}_1$, és így tovább, $\mathbf{B}\mathbf{e}_n = \lambda_n\mathbf{e}_n$. Ez utóbbi egyenlet teljesülését úgy fogjuk nevezni, hogy \mathbf{B} -nek \mathbf{e}_j sajátvektora, λ_j pedig a hozzá tartozó sajátérték ($j = 1, \dots, n$).

39. Definíció. Legyen $\mathbf{A} \in \mathbb{R}^{n \times n}$. Azt mondjuk, a $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$ jobb oldali sajátvektora \mathbf{A} -nak, ha létezik olyan $\lambda_0 \in \mathbb{R}$, melyre $\mathbf{A}\mathbf{x} = \lambda_0\mathbf{x}$. Ilyenkor a λ_0 számot az \mathbf{x} jobb oldali sajátvektorhoz tartozó jobb oldali sajátértéknek nevezzük. Az $\mathbf{0} \neq \mathbf{y} \in \mathbb{R}^{1 \times n}$ bal oldali sajátvektora \mathbf{A} -nak, ha létezik olyan $\mu_0 \in \mathbb{R}$, hogy $\mathbf{y}\mathbf{A} = \mu_0\mathbf{y}$. Ekkor μ_0 az \mathbf{y} bal oldali sajátvektorhoz tartozó bal oldali sajátérték.

Mi csak jobb oldali sajátvektorokkal foglalkozunk, hiszen $\mathbf{y}\mathbf{A} = \mu_0\mathbf{y} \iff \mathbf{A}^\top \mathbf{y}^\top = \mu_0\mathbf{y}^\top$ összefüggés végett a bal oldali eset is visszavezethető a jobb oldalra, tehát sajátvektor és sajátérték alatt ezentúl mindig a jobb oldaliakra gondolunk majd.

Térjünk vissza a korábbi „szép” \mathbf{B} mátrixra. Láttuk, hogy $\mathbf{B}\mathbf{e}_j = \lambda_j\mathbf{e}_j$ minden szóbaeső j -re. Láttuk, hogy $\mathbf{B} = \mathbf{D}^{-1}\mathbf{A}\mathbf{D}$, innen átszorzással $\mathbf{A} = \mathbf{D}\mathbf{B}\mathbf{D}^{-1}$. Vizsgálódjunk kicsit:

$$\mathbf{B}\mathbf{e}_j = \lambda_j\mathbf{e}_j \implies \mathbf{D}^{-1}\mathbf{A}\mathbf{D}\mathbf{e}_j = \lambda_j\mathbf{e}_j \implies \mathbf{A}\mathbf{D}\mathbf{e}_j = \mathbf{D}\lambda_j\mathbf{e}_j \implies \mathbf{A}(\mathbf{D}\mathbf{e}_j) = \lambda_j(\mathbf{D}\mathbf{e}_j).$$

Írjuk fel \mathbf{D} -t $\mathbf{D} = [\mathbf{d}_1, \dots, \mathbf{d}_n]$ alakban. Nyilván $\mathbf{D} = \mathbf{D}\mathbf{I}_n = \mathbf{D}[\mathbf{e}_1, \dots, \mathbf{e}_n] = [\mathbf{D}\mathbf{e}_1, \dots, \mathbf{D}\mathbf{e}_n]$. Emiatt minden j -re $\mathbf{d}_j = \mathbf{D}\mathbf{e}_j$. A korábbiak miatt így $\mathbf{A}\mathbf{d}_j = \lambda_j\mathbf{d}_j$. Tudjuk továbbá hogy \mathbf{D} invertálható, azaz $\varrho(\mathbf{D}) = n$ így az oszlopai lineárisan független rendszert alkotnak \mathbb{R}^n -ben, s mivel ez n db vektort jelent így ez azt is jelenti, hogy $\mathbf{d}_1, \dots, \mathbf{d}_n$ bázis \mathbb{R}^n -ben, s ezek a bázisvektorok \mathbf{A} -nak sajátvektorai. Ezzel megadtuk a most következő tétel \implies irányát, melynek másik irányá a fentiek visszafelé olvasásából kapható.

42. Tétel. Legyen $\mathbf{A} \in \mathbb{R}^{n \times n}$. Ekkor \mathbf{A} diagonalizálható \iff létezik \mathbb{R}^n -ben \mathbf{A} sajátvektoraiból álló bázis (röviden: sajátbázis).

Most adott sajátérték mellett tekintsük az összes hozzá tartozó sajátvektort. A nullvektort hozzávéve \mathbb{R}^n egy alterét kapjuk:

40. Definíció. Legyen $\mathbf{A} \in \mathbb{R}^{n \times n}$, valamint $\lambda_0 \in \mathbb{R}$ egy sajátértéke \mathbf{A} -nak. A λ_0 -hoz tartozó sajátaltér:

$$W_{\lambda_0} := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \lambda_0\mathbf{x}\}$$

A létezésről még nem ejtettünk szót, de azt most feltéve mi igazolja azt, hogy a fenti halmaz valóban altér? Vegyünk két tetszőleges $\mathbf{x}, \mathbf{y} \in W_{\lambda_0}$ -t. Ekkor $\mathbf{A}\mathbf{x} = \lambda_0\mathbf{x}$ és $\mathbf{A}\mathbf{y} = \lambda_0\mathbf{y}$, és emiatt

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y} = \lambda_0\mathbf{x} + \lambda_0\mathbf{y} = \lambda_0(\mathbf{x} + \mathbf{y}),$$

így az összegvektor is eleme lesz a fenti halmaznak. Hasonlóan tetszőleges $\lambda \in \mathbb{R}$ -re

$$\mathbf{A}(\lambda\mathbf{x}) = \lambda\mathbf{A}\mathbf{x} = \lambda\lambda_0\mathbf{x} = \lambda_0(\lambda\mathbf{x}),$$

tehát a fenti halmaz valóban altér. Nézzünk egy példát. Legyen

$$\mathbf{A} := \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$$

Keressük a sajátvektorokat

$$\mathbf{0} \neq \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$$

alakban:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \iff \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \iff \begin{cases} 3x_1 = \lambda x_1 \\ x_1 + 2x_2 = \lambda x_2 \end{cases} \iff \begin{cases} (\lambda - 3)x_1 = 0 \\ x_1 = (\lambda - 2)x_2 \end{cases}$$

Elsőnek tegyük fel, hogy $x_1 \neq 0$. Ekkor $\lambda - 3 = 0$, tehát $\lambda = 3$, a második egyenlet miatt pedig $x_1 = x_2$. Második eset, ha $x_1 = 0$. Ekkor $x_2 = 0$ nem lehet, mert $\mathbf{x} = \mathbf{0}$ esetet kizártuk. Ekkor $0 = (\lambda - 2)x_2$, tehát $2x_2 = \lambda x_2$, vagyis $\lambda = 2$. Így tehát az \mathbf{A} sajátértékei $\lambda_1 = 3$, $\lambda_2 = 2$. A λ_1 -hez tartozó sajátvektorokban a komponensek megegyeztek, tehát

$$W_3 = \left\{ \begin{bmatrix} x \\ x \end{bmatrix} : x \in \mathbb{R} \right\}$$

a hozzá tartozó sajátaltér, míg λ_2 -höz:

$$W_2 = \left\{ \begin{bmatrix} 0 \\ x \end{bmatrix} : x \in \mathbb{R} \right\}$$

Létezik sajátbázis, pl.

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

Láttuk, hogy ebből készíthetjük \mathbf{D} oszlopait, tehát

$$\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

ez a mátrix független oszlopai révén invertálható, és

$$\mathbf{D}^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \implies \mathbf{D}^{-1}\mathbf{A}\mathbf{D} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

Látjuk, hogy a kapott diagonális mátrix elemei pont a sajátértékek, ahogy azt korábban láttuk. Sajnos mivel esetszétválasztást kellett végrehajtanunk, nagyobb mátrixok esetén ez a módszer gyakorlatilag kivitelezhetetlen. Tovább vizsgálódunk, hogy jobb módszerhez jussunk.

6.2. Karakterisztikus polinom

Végezzünk el néhány ekvivalens átalakítást, és nézzük meg mire jutunk. Tudjuk, hogy $\lambda_0 \in \mathbb{R}$ esetén λ_0 akkor és csak akkor sajátértéke $\mathbf{A} \in \mathbb{R}^{n \times n}$ -nek, ha létezik $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{A}\mathbf{x} = \lambda_0\mathbf{x} \iff \mathbf{A}\mathbf{x} - \lambda_0\mathbf{x} = \mathbf{0} \iff (\mathbf{A} - \mathbf{I}_n\lambda_0)\mathbf{x} = \mathbf{0} \iff |\mathbf{A} - \mathbf{I}_n\lambda_0| = 0.$$

Ez a determináns a következő alakú:

$$|\mathbf{A} - \mathbf{I}_n\lambda_0| = \begin{vmatrix} a_{11} - \lambda_0 & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} - \lambda_0 & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} - \lambda_0 & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} - \lambda_0 \end{vmatrix}$$

A fenti determináns értékének tehát nullának kell lennie ahhoz, hogy λ_0 sajátérték legyen. Ennek nem könnyű a felírása, az viszont látszik, hogy λ_0 legfeljebb n -edik hatványon (illetve

annál kisebbeken) fel fog tűnni a kifejtésben. Most λ_0 -ról feltettük hogy egy ismert sajátérték, hogy keressünk ekvivalens feltételt. Ha λ_0 helyett λ jelölést használunk mint határozatlan értéket, akkor a fenti kifejezés egy $f : \mathbb{R} \rightarrow \mathbb{R}$ függvényt határoz meg:

$$f(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} - \lambda & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} - \lambda \end{vmatrix} \in \mathbb{R}$$

Belátható, hogy ez a függvény egy n -edfokú polinom, melyet karakterisztikus polinomnak fogunk nevezni.

41. Definíció. Az $\mathbf{A} \in \mathbb{R}^{n \times n}$ mátrix karakterisztikus polinomja $k_{\mathbf{A}}(\lambda) = |\mathbf{A} - \mathbf{I}_n \lambda|$.

A fentiek értelmében a sajátértékek éppen ennek a függvénynek a (valós) zérushelyei. Írjuk fel ezt a polinomot a korábbi példánkra:

$$k_{\mathbf{A}}(\lambda) = \begin{vmatrix} 3 - \lambda & 0 \\ 1 & 2 - \lambda \end{vmatrix} = (3 - \lambda)(2 - \lambda).$$

Szépen látszódik, hogy ennek a másodfokú polinomnak a gyökei a 3 és a 2.

Megmutatjuk, hogy hasonló mátrixok karakterisztikus polinomjai megegyeznek (így sajátértékeik is):

43. Tétel. Legyenek $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$, $\mathbf{A} \sim_{\mathbb{R}} \mathbf{B}$. Ekkor $k_{\mathbf{A}}(\lambda) = k_{\mathbf{B}}(\lambda)$.

Bizonyítás. Egyszerűen technikai:

$$\begin{aligned} k_{\mathbf{B}}(\lambda) &= |\mathbf{B} - \mathbf{I}_n \lambda| = |\mathbf{D}^{-1} \mathbf{A} \mathbf{D} - \mathbf{I}_n \lambda| = |\mathbf{D}^{-1} (\mathbf{A} - \mathbf{I}_n \lambda) \mathbf{D}| = \\ &= |\mathbf{D}^{-1}| |\mathbf{A} - \mathbf{I}_n \lambda| |\mathbf{D}| = |\mathbf{D}^{-1}| |\mathbf{D}| |\mathbf{A} - \mathbf{I}_n \lambda| = |\mathbf{D}^{-1} \mathbf{D}| |\mathbf{A} - \mathbf{I}_n \lambda| = \\ &= |\mathbf{I}_n| |\mathbf{A} - \mathbf{I}_n \lambda| = |\mathbf{A} - \mathbf{I}_n \lambda| = k_{\mathbf{A}}(\lambda), \end{aligned}$$

és készen vagyunk. □

Nézzünk most egy nem kívánatos helyzetet. Legyen

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

Ekkor $k_{\mathbf{A}}(\lambda) = (-\lambda)^2 + 1 = \lambda^2 + 1$, melynek nincs valós gyöke. Az algebra alaptétele alapján minden legalább elsőfokú polinomnak létezik komplex gyöke, így tehát sajnos a sajátértékek ezesetben komplex számok, és $\mathbb{R}^{n \times n}$ -ből „át kell mennünk” $\mathbb{C}^{n \times n}$ -be. A fenti polinomnak i és $-i$ a gyökei, így ezek lesznek a sajátértékei.

7. fejezet

Euklideszi terek

7.1. Bevezetés

Emlékezzünk vissza a geometriai vektoroknál bevezetett skaláris szorzatra. Szeretnénk ezt a fogalmat általánosítani először \mathbb{R}^n , majd \mathbb{C}^n -re, lehetőleg úgy, hogy a két esetet együtt tudjuk kezelni. Emlékszünk, hogy a koordináták segítségével megkaptuk a skaláris szorzatot úgy, hogy komponensenként összeszoroztunk, majd a kapott szorzatokat összeadtuk. Ennek $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ -re egy természetes általánosítása $\mathbf{y}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{y}$, melyet $\langle \mathbf{x}, \mathbf{y} \rangle$ módon fogunk jelölni. Egy vektor hossza a vektornak az önmagával vett skaláris szorzat négyzetgyöke volt, így a $\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ a hossz fogalmának egy természetes általánosítása. Komplex számoknál a transzponálás általánosítása az adjungálás volt, de most felmerül a kérdés, hogy a két vektor közül melyiket adjungáljuk? Kénytelenek leszünk a kommutativitásról részben lemondani, a skaláris szorzást $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ esetére $\mathbf{y}^* \mathbf{x}$ módon fogjuk definiálni. Ezek itt a geometriai vektornál megismert skaláris szorzat általánosításai, ám most a skaláris szorzatot tulajdonságaival definiáljuk, így a fenti képlet csak egy konkrét skaláris szorzatot ad a számos közül.

42. Definíció. Legyen $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, valamint V egy vektortér \mathbb{K} felett. Azt mondjuk, V valós vagy komplex euklideszi tér, ha adott benne egy skaláris szorzatnak nevezett

$$\langle \mathbf{x}, \mathbf{y} \rangle : V \times V \rightarrow \mathbb{K}$$

függvény, melyre a következők teljesülnek minden $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\lambda \in \mathbb{K}$ esetén:

1. $\langle \mathbf{y}, \mathbf{x} \rangle = \overline{\langle \mathbf{x}, \mathbf{y} \rangle}$;
2. $\langle \lambda \mathbf{x}, \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle$;
3. $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$;

$$4. \ 0 \leq \langle \mathbf{x}, \mathbf{x} \rangle \in \mathbb{R};$$

$$5. \ \langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}.$$

A fenti tulajdonságok alapján néhány következmény észrevehető. Vigyázni kell például ha a második vektorból akarunk skalárt kiemelni, hiszen

$$\langle \mathbf{x}, \lambda \mathbf{y} \rangle = \overline{\langle \lambda \mathbf{y}, \mathbf{x} \rangle} = \overline{\lambda \langle \mathbf{y}, \mathbf{x} \rangle} = \bar{\lambda} \cdot \overline{\langle \mathbf{y}, \mathbf{x} \rangle} = \bar{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle.$$

Könnyen megmutatható továbbá a másik disztributivitás, tehát $\langle \mathbf{z}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{z}, \mathbf{x} \rangle + \langle \mathbf{z}, \mathbf{y} \rangle$. Skaláris szorzatból sokféle definiálható. Legyen $\mathbf{B} \in \mathbb{C}^{n \times n}$ egy teljes rangú mátrix és legyen $\mathbf{A} := \mathbf{B}^* \mathbf{B}$. A vektortér legyen $V := \mathbb{C}^n$. Ekkor megmutatható, hogy $\mathbf{x}, \mathbf{y} \in V$ esetén

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{y}^* \mathbf{A} \mathbf{x}$$

szintén skaláris szorzatot definiál. Egy másik példa: $V := \mathbb{C}^{n \times n}$. Ekkor

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{Tr}(\mathbf{B}^* \mathbf{A})$$

szintén skaláris szorzat, ahol Tr a mátrix nyomát jelenti, azaz a főátlóban lévő elemek összegét. Amikor egy euklideszi térről beszélünk, mindig egy rögzített skaláris szorzatot fogunk értelmezni rajta.

7.2. Euklideszi norma

43. Definíció. Legyen V valós vagy komplex euklideszi tér, $\mathbf{x} \in V$ esetén az \mathbf{x} (euklideszi) normája: $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

A fenn megbeszéltek szerint ez a hossz (abszolút érték) egy természetes általánosítása, ezt azonban pontosabban is megmutatjuk, nevezetesen tehát hogy a norma rendelkezik az abszolút értéknél megismert tulajdonságokkal. Legyenek $\mathbf{x}, \mathbf{y} \in V$, $\lambda \in \mathbb{K}$. A skaláris szorzat tulajdonságaiból adódik, hogy $\|\mathbf{x}\| \geq 0$ és $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$. Könnyen megmutatható továbbá, hogy $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$. Fontos kérdés a háromszög-egyenlőtlenség teljesülése:

$$\|\mathbf{x} + \mathbf{y}\| \stackrel{?}{\leq} \|\mathbf{x}\| + \|\mathbf{y}\|$$

Fogalmazzuk át:

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| &\iff \|\mathbf{x} + \mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 \iff \\ &\iff \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle \leq \langle \mathbf{x}, \mathbf{x} \rangle + 2\|\mathbf{x}\|\|\mathbf{y}\| + \langle \mathbf{y}, \mathbf{y} \rangle \iff \end{aligned}$$

$$\begin{aligned}
&\Longleftrightarrow \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \leq \langle \mathbf{x}, \mathbf{x} \rangle + 2\|\mathbf{x}\|\|\mathbf{y}\| + \langle \mathbf{y}, \mathbf{y} \rangle \Longleftrightarrow \\
&\Longleftrightarrow \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle \leq 2\|\mathbf{x}\|\|\mathbf{y}\| \Longleftrightarrow \langle \mathbf{x}, \mathbf{y} \rangle + \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \leq 2\|\mathbf{x}\|\|\mathbf{y}\| \Longleftrightarrow \\
&\Longleftrightarrow 2\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle \leq 2\|\mathbf{x}\|\|\mathbf{y}\| \Longleftrightarrow \operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{x}\|\|\mathbf{y}\|.
\end{aligned}$$

Bizonyítunk a háromszög-egyenlőtlenségnél egy erősebb tételt, mely majd bizonyítást ad a háromszög-egyenlőtlenségre is.

44. Tétel (Cauchy-egyenlőtlenség). *Legyen V valós vagy komplex euklideszi tér. Ekkor bármely $\mathbf{x}, \mathbf{y} \in V$ -re $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$, egyenlőség pedig pontosan akkor teljesül, ha \mathbf{x} és \mathbf{y} lineárisan összefüggő.*

Bizonyítás. Tekintsük az egyenlőséget. Ha a két vektor összefüggő, akkor egyik valamely másikkal számszorosa, pl. $\mathbf{x} = \lambda\mathbf{y}$. Ekkor $\|\mathbf{x}\| = |\lambda|\|\mathbf{y}\|$, ezt felhasználva

$$|\langle \mathbf{x}, \mathbf{y} \rangle| = |\langle \lambda\mathbf{y}, \mathbf{y} \rangle| = |\lambda\langle \mathbf{y}, \mathbf{y} \rangle| = |\lambda|\|\mathbf{y}\|^2 = \|\mathbf{x}\|\|\mathbf{y}\|.$$

Most feltesszük a vektorok függetlenségét, ekkor biztosan egyik vektor sem nullvektor, továbbá tetszőleges $\lambda \in \mathbb{K}$ -ra $\lambda\mathbf{x} + \mathbf{y} \neq \mathbf{0}$. Ekkor

$$\begin{aligned}
0 &< \langle \lambda\mathbf{x} + \mathbf{y}, \lambda\mathbf{x} + \mathbf{y} \rangle = \langle \lambda\mathbf{x}, \lambda\mathbf{x} \rangle + \langle \lambda\mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \lambda\mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle = \\
&= \|\lambda\mathbf{x}\|^2 + 2\operatorname{Re}\langle \lambda\mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 = |\lambda|^2\|\mathbf{x}\|^2 + 2\operatorname{Re}(\lambda\langle \mathbf{x}, \mathbf{y} \rangle) + \|\mathbf{y}\|^2
\end{aligned}$$

Legyen $\lambda = \alpha\overline{\langle \mathbf{x}, \mathbf{y} \rangle}$ alakú, ahol $\alpha \in \mathbb{R}$. Így folytatva

$$\begin{aligned}
0 &< |\alpha\overline{\langle \mathbf{x}, \mathbf{y} \rangle}|^2\|\mathbf{x}\|^2 + 2\operatorname{Re}(\alpha\overline{\langle \mathbf{x}, \mathbf{y} \rangle}\langle \mathbf{x}, \mathbf{y} \rangle) + \|\mathbf{y}\|^2 = \\
&= \alpha^2|\overline{\langle \mathbf{x}, \mathbf{y} \rangle}|^2\|\mathbf{x}\|^2 + 2\alpha|\langle \mathbf{x}, \mathbf{y} \rangle|^2 + \|\mathbf{y}\|^2.
\end{aligned}$$

Szorozzuk meg az egyenlőtlenséget $0 < \|\mathbf{x}\|^2$ -vel, azt kapjuk, hogy

$$\begin{aligned}
0 &< \alpha^2|\overline{\langle \mathbf{x}, \mathbf{y} \rangle}|^2\|\mathbf{x}\|^4 + 2\alpha|\langle \mathbf{x}, \mathbf{y} \rangle|^2\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\|\mathbf{x}\|^2 = \\
&= |\overline{\langle \mathbf{x}, \mathbf{y} \rangle}|^2\left(\alpha^2\|\mathbf{x}\|^4 + 2\alpha\|\mathbf{x}\|^2\right) + \|\mathbf{x}\|^2\|\mathbf{y}\|^2 \Longleftrightarrow \\
\Longleftrightarrow 0 &< |\overline{\langle \mathbf{x}, \mathbf{y} \rangle}|^2\left(\alpha^2\|\mathbf{x}\|^4 + 2\alpha\|\mathbf{x}\|^2 + 1\right) - |\overline{\langle \mathbf{x}, \mathbf{y} \rangle}|^2 + \|\mathbf{x}\|^2\|\mathbf{y}\|^2 \Longleftrightarrow \\
&|\overline{\langle \mathbf{x}, \mathbf{y} \rangle}|^2 < |\overline{\langle \mathbf{x}, \mathbf{y} \rangle}|^2\left((\alpha\|\mathbf{x}\|^2 + 1)^2\right) + \|\mathbf{x}\|^2\|\mathbf{y}\|^2,
\end{aligned}$$

innen pedig $\alpha = -1/\|\mathbf{x}\|^2$ választással kapjuk a bizonyítandó állítást. □

Ezek után definiálhatunk távolságot euklideszi térben: $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

7.3. Ortonormált bázisok

Ismertünk a geometriai vektorok terében olyan fogalmat, mint két vektor hajlásszöge. Ezt a $[0, \pi]$ -beli értéket koszinuszával definiáltuk: $\cos \gamma(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} / (\|\mathbf{a}\| \|\mathbf{b}\|)$. Ezt az eddigi ismereteinkkel általánosíthatjuk euklideszi terekre:

$$\cos \gamma(\mathbf{a}, \mathbf{b}) = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \|\mathbf{b}\|}.$$

Maga a szög azonban annyira nem fontos, inkább azok lesznek az érdekes esetek, amikor a vektorok merőlegesek egymásra (más szóval ortogonálisak). Ez pontosan akkor teljesül, ha $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, ez így módon már tetszőleges valós vagy komplex euklideszi térre definiálható. Vegyük észre, hogy a geometriai vektoroknál megismert $\mathbf{i}, \mathbf{j}, \mathbf{k}$ bázisvektorok páronként merőlegesek voltak, illetve másik fontos tulajdonságuk, hogy hosszuk 1 volt. Világos, hogy bármely vektor egységnyi hosszúra transzformálható, ha szorozzuk a hosszának reciprokával, hiszen (most még csak geometriai vektorokra megfogalmazva):

$$\left| \frac{\mathbf{a}}{\|\mathbf{a}\|} \right| = \frac{\|\mathbf{a}\|}{\|\mathbf{a}\|} = 1.$$

Az abszolút értéket a norma fogalmával általánosítottuk euklideszi terekre. Így már értelmes a következő definíció:

44. Definíció. Legyen V n -dimenziós (valós vagy komplex) euklideszi tér, $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V -ben. Az $\mathbf{e}_1, \dots, \mathbf{e}_n$ ortogonális bázis V -ben, ha bázis, és elemei páronként merőlegesek. Az $\mathbf{e}_1, \dots, \mathbf{e}_n$ ortonormált bázis V -ben, ha ortogonális bázis, és minden vektor normája 1.

A következő tétel bizonyításában az ún. Gram-Schmidt-féle ortogonalizációs eljárást fogjuk alkalmazni.

45. Tétel. $n > 0$ -ra tetszőleges n -dimenziós euklideszi térben létezik ortonormált bázis.

Bizonyítás. Elegendő ortogonális bázist találnunk, hiszen ekkor (a fenn elmondottak végett) ha $\mathbf{e}_1, \dots, \mathbf{e}_n$ egy ortogonális bázis, akkor $\mathbf{e}_1/\|\mathbf{e}_1\|, \dots, \mathbf{e}_n/\|\mathbf{e}_n\|$ ortonormált bázis. Legyen $\mathbf{b}_1, \dots, \mathbf{b}_n$ bázis V -ben, és legyen $\mathbf{e}_1 := \mathbf{b}_1$. Ha \mathbf{b}_2 merőleges \mathbf{b}_1 -re, akkor $\mathbf{e}_2 := \mathbf{b}_2$ és folytathatjuk, ha nem, akkor ortogonalizálnunk kell a következőképp: tegyük fel, hogy $\mathbf{e}_1, \dots, \mathbf{e}_i$ már ortogonálisak, illetve $\text{Span}(\mathbf{e}_1, \dots, \mathbf{e}_i) = \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_i)$. Keressük \mathbf{e}_{i+1} -et a következő alakban (erre a bizonyítás után mutatunk \mathbb{R}^2 -beli példát):

$$\mathbf{e}_{i+1} = \mathbf{b}_{i+1} + \lambda_1 \mathbf{e}_1 + \dots + \lambda_i \mathbf{e}_i$$

Azt kell biztosítani, hogy $\text{Span}(\mathbf{e}_1, \dots, \mathbf{e}_{i+1}) = \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_{i+1})$ teljesüljön úgy, hogy \mathbf{e}_{i+1} merőleges az összes megelőző \mathbf{e} -re. Az egyenlőség kétoldali tartalmazással belátható, hiszen könnyen láthatóan $\text{Span}(\mathbf{e}_1, \dots, \mathbf{e}_i) \subseteq \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_{i+1})$, továbbá $\mathbf{b}_{i+1} \in \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_{i+1})$, ezért a fenn definiált \mathbf{e}_{i+1} is benne van, tehát a \subseteq irányt beláttuk, hasonlóan látható be a másik irány is. Most a merőlegességet próbáljuk meg biztosítani. Legyen $1 \leq j \leq i$. Ekkor \mathbf{e}_{i+1} -nek merőlegesnek kell lennie \mathbf{e}_j -re (j választása tetszőleges, ezért biztosítja az összes merőlegességet). Ez azt jelenti, hogy $\langle \mathbf{e}_{i+1}, \mathbf{e}_j \rangle = 0$ kell legyen.

$$\begin{aligned} \langle \mathbf{e}_{i+1}, \mathbf{e}_j \rangle &= \langle \mathbf{b}_{i+1} + \lambda_1 \mathbf{e}_1 + \dots + \lambda_i \mathbf{e}_i, \mathbf{e}_j \rangle = \\ &= \langle \mathbf{b}_{i+1}, \mathbf{e}_j \rangle + \langle \lambda_1 \mathbf{e}_1, \mathbf{e}_j \rangle + \dots + \langle \lambda_i \mathbf{e}_i, \mathbf{e}_j \rangle = \\ &= \langle \mathbf{b}_{i+1}, \mathbf{e}_j \rangle + \lambda_1 \underbrace{\langle \mathbf{e}_1, \mathbf{e}_j \rangle}_0 + \dots + \lambda_{j-1} \underbrace{\langle \mathbf{e}_{j-1}, \mathbf{e}_j \rangle}_0 + \lambda_j \langle \mathbf{e}_j, \mathbf{e}_j \rangle + \lambda_{j+1} \underbrace{\langle \mathbf{e}_{j+1}, \mathbf{e}_j \rangle}_0 + \dots + \lambda_i \underbrace{\langle \mathbf{e}_i, \mathbf{e}_j \rangle}_0 = \\ &= \langle \mathbf{b}_{i+1}, \mathbf{e}_j \rangle + \lambda_j \langle \mathbf{e}_j, \mathbf{e}_j \rangle \end{aligned}$$

Mivel $\mathbf{b}_1, \dots, \mathbf{b}_i$ független, ezért $\text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_i) = \text{Span}(\mathbf{e}_1, \dots, \mathbf{e}_i)$ dimenziója i , így tehát \mathbf{b}_j nem lehet nullvektor, oszthatunk tehát:

$$0 = \langle \mathbf{b}_{i+1}, \mathbf{e}_j \rangle + \lambda_j \langle \mathbf{e}_j, \mathbf{e}_j \rangle \implies \lambda_j = -\frac{\langle \mathbf{b}_{i+1}, \mathbf{e}_j \rangle}{\langle \mathbf{e}_j, \mathbf{e}_j \rangle}$$

ezzel a választással tehát a kívánt merőlegességhez jutunk. □

Ígértünk \mathbb{R}^2 -beli példát. Legyen az eredeti bázis \mathbf{e}_1, \mathbf{b} , ahol pl.:

$$\mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Ez nem merőleges \mathbf{e}_1 -re, így az \mathbf{e}_2 új bázisvektort úgy keressük, hogy \mathbf{b} -hez hozzávesszük \mathbf{e}_1 valamely alkalmas λ -szorosát: $\mathbf{e}_2 = \mathbf{b} + \lambda \mathbf{e}_1$. Tudjuk, hogy $\langle \mathbf{e}_1, \mathbf{e}_1 \rangle = 1$, illetve $\langle \mathbf{b}, \mathbf{e}_1 \rangle = 2$, így $\lambda = -2$. $\mathbf{b} - 2\mathbf{e}_1$ pont az általunk ismert második kanonikus egységvektort adja ki, tehát itt éppenséggel a merőlegesség mellett normált vektort is kaptunk, ez persze nem feltétlen teljesül, így adott esetben még normálnunk kellhet.

46. Tétel (Valós szimmetrikus mátrixok spektráltétele). *Legyen $\mathbf{A} \in \mathbb{R}^{n \times n}$. Ekkor \mathbf{A} szimmetrikus \iff létezik saját ortonormált bázisa \mathbb{R}^n -ben, továbbá minden sajátértéke valós.*

7.4. Kvadratikus alak, definit

Kvadratikus alatt egy speciális leképezést fogunk érteni valós n -dimenziós vektorok teréből a valós számok terébe (megjegyzés: a determináns is értelmezhető leképezésként, melyet a négyzetes mátrixokból képeztünk le a valós számokra, tehát $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$).

45. Definíció. Legyen $\mathbf{A} \in \mathbb{R}^{n \times n}$ szimmetrikus mátrix. Az \mathbf{A} mátrixhoz tartozó Q kvadratikus alak (vagy kvadratikus forma) olyan $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ leképezés, melyre $Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x}$.

Legyen például

$$\mathbf{A} := \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Ekkor pl.:

$$Q\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} = 26.$$

Vegyük észre, hogy a valós vektortéren értelmezett $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^\top \mathbf{x}$ skaláris szorzattal a kvadratikus alak a következővel ekvivalens: $Q(\mathbf{x}) = \langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle$. Csupán a koordináták segítségével is definiálhatjuk a kvadratikus alakot, hisz

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ esetén } Q(\mathbf{x}) = \sum_{j=1}^n x_j \sum_{k=1}^n a_{jk} x_k$$

Írjuk fel $n = 2$ -re, látjuk is honnan az elnevezés (kvadratikus jelentése: négyzetes). Ne feledjük, hogy a mátrix szimmetriája végett $a_{21} = a_{12}$:

$$Q(\mathbf{x}) = x_1(a_{11}x_1 + a_{12}x_2) + x_2(a_{21}x_1 + a_{22}x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2.$$

Tudjuk a spektráltételből, hogy a szimmetria miatt létezik saját ortonormált bázis, tehát olyan $\mathbf{u}_1, \dots, \mathbf{u}_n$ ortonormált bázis, melyre $\mathbf{A} \mathbf{u}_k = \lambda_k \mathbf{u}_k$, ahol $k = 1, \dots, n$ és $\lambda_k \in \mathbb{R}$. Írjuk fel a tetszőleges \mathbf{x} -et ebben a bázisban: $\mathbf{x} = \sum_{k=1}^n \alpha_k \mathbf{u}_k$. Ekkor (kiírva a lépéseket kellő részletességgel)

$$\begin{aligned} \mathbf{A} \mathbf{x} &= \sum_{k=1}^n \alpha_k \mathbf{A} \mathbf{u}_k = \sum_{k=1}^n \lambda_k \alpha_k \mathbf{u}_k \implies \\ \implies Q(\mathbf{x}) &= \langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle \stackrel{\text{valós szimmetria}}{=} \langle \mathbf{x}, \mathbf{A} \mathbf{x} \rangle = \left\langle \sum_{j=1}^n \alpha_j \mathbf{u}_j, \sum_{k=1}^n \lambda_k \alpha_k \mathbf{u}_k \right\rangle = \\ &= \langle \alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n, \lambda_1 \alpha_1 \mathbf{u}_1 + \dots + \lambda_n \alpha_n \mathbf{u}_n \rangle = \end{aligned}$$

$$= \sum_{j=1}^n \sum_{k=1}^n \langle \alpha_j \mathbf{u}_j, \lambda_k \alpha_k \mathbf{u}_k \rangle = \sum_{j=1}^n \sum_{k=1}^n \alpha_j \lambda_k \alpha_k \langle \mathbf{u}_j, \mathbf{u}_k \rangle \stackrel{(*)}{=} \sum_{k=1}^n \alpha_k \lambda_k \alpha_k = \sum_{k=1}^n \lambda_k \alpha_k^2$$

A csillagozott lépést még indokolnunk kell. $\mathbf{u}_1, \dots, \mathbf{u}_n$ ortogonális rendszer, tehát $j \neq k$ -ra $\langle \mathbf{u}_j, \mathbf{u}_k \rangle = 0$, tehát csak azok a tagok maradnak meg, ahol $j = k$, ezért hagyható el az egyik összegzés. Továbbá mivel a vektor hossza $\|\mathbf{u}_k\| = 1$, ami definíció szerint $\|\mathbf{u}_k\| = \sqrt{\langle \mathbf{u}_k, \mathbf{u}_k \rangle}$, így négyzetre emelve kapjuk, hogy egységvektor önmagával vett skaláris szorzata 1. Kaptunk egy négyzetes kifejezést, melyekben a sajátértékek is feltűnnek, ennek speciális eseteivel definiáljuk a definitséget. $\mathbf{x} = \mathbf{0}$ nem túl érdekes, hiszen $Q(\mathbf{0}) = 0$, így csak a többi esettel foglalkozunk.

46. Definíció. Legyen $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Ekkor ha minden $\lambda_k > 0$, vagy ha $Q(\mathbf{x}) > 0$ minden nemnulla \mathbf{x} -re, akkor Q -t pozitív definitnek nevezzük. Hasonlóan, ha minden $\lambda_k < 0$ vagy $Q(\mathbf{x}) < 0$ minden nemnulla \mathbf{x} -re, akkor Q negatív definit. Ha egyenlőséget is megengedünk a fentiekben, akkor pozitív, ill. negatív szemidefinitről beszélünk. Minden más esetben Q indefinit.

A pozitív, ill. negatív definitség eldönthető a sajátértékek nélkül is, ezt bizonyítás nélkül közöljük:

47. Tétel. Legyen

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n \times n}$$

szimmetrikus mátrix. A mátrix karakterisztikus sorozatának nevezzük a bal felső sarokdeterminánsokból képzett sorozatot, tehát

$$\Delta_0 := 1, \quad \Delta_1 := a_{11}, \quad \Delta_2 := \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \Delta_3 := \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \dots \quad \Delta_n := |\mathbf{A}|.$$

Az \mathbf{A} -hoz tartozó Q akkor és csak akkor pozitív definit, ha a fenti sorozat minden tagja pozitív. Q akkor és csak akkor negatív definit, ha $\forall j \in \{0, 1, \dots, [n/2]\}$ -re $\Delta_{2j} > 0$, illetve $\forall j \in \{0, 1, \dots, [(n-1)/2]\}$ -re $\Delta_{2j+1} < 0$.

Az első esetben azt is mondjuk, hogy a sorozat jeltartó, a második esetben pedig jelváltó.

7.5. Saját ortonormált bázis keresése

A gyakorlatban valós szimmetrikus mátrixhoz saját ortonormált bázist kereshetünk sajátaltérként, majd az ezekből kapott bázisvektorokat összerakhatjuk. Ezt a következő tétel biztosítja, mely kimondja, hogy a különböző sajátaltérbeli vektorok ortogonálisak.

48. Tétel. Legyen $\mathbf{A} = \mathbf{A}^\top \in \mathbb{R}^{n \times n}$ és $\lambda, \mu \in \mathbb{R}$ az \mathbf{A} különböző sajátértékei, továbbá $\mathbf{x} \in W_\lambda$, $\mathbf{y} \in W_\mu$. Ekkor $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Bizonyítás. Elsőnek megmutatjuk, hogy $\lambda \langle \mathbf{x}, \mathbf{y} \rangle = \mu \langle \mathbf{x}, \mathbf{y} \rangle$:

$$\lambda \langle \mathbf{x}, \mathbf{y} \rangle = \langle \lambda \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{A} \mathbf{x}, \mathbf{y} \rangle = (\mathbf{A} \mathbf{x})^\top \mathbf{y} = \mathbf{x}^\top \mathbf{A}^\top \mathbf{y} = \mathbf{x}^\top \mathbf{A} \mathbf{y} = \langle \mathbf{x}, \mathbf{A} \mathbf{y} \rangle = \langle \mathbf{x}, \mu \mathbf{y} \rangle = \mu \langle \mathbf{x}, \mathbf{y} \rangle.$$

Ezért $(\lambda - \mu) \langle \mathbf{x}, \mathbf{y} \rangle = 0$, de $\lambda - \mu \neq 0$, és innen kapjuk az állítást. \square

Nézzünk egy példát, legyen

$$\mathbf{A} := \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

Elsőnek írjuk fel a karakterisztikus polinomot:

$$p_{\mathbf{A}}(\lambda) = \begin{vmatrix} 2 - \lambda & -1 \\ -1 & 2 - \lambda \end{vmatrix} = (2 - \lambda)^2 - 1$$

Ennek gyökei $\lambda_1 = 1$ és $\lambda_2 = 3$, így a sajátértékek megvannak. Írjuk fel a sajátaltérket, ehhez szükség van a sajátvektorok meghatározására:

$$\mathbf{A} \mathbf{x} = \lambda_1 \mathbf{x} \implies \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \implies 2x_1 - x_2 = x_1 \text{ és } 2x_2 - x_1 = x_2.$$

Azt kapjuk, hogy $2x_1 = 2x_2 = x_1 + x_2$, vagyis $x_1 = x_2$. Ez az altér tehát (már egyből bázissal megadva)

$$W_1 = \text{Span} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$$

Hasonlóan:

$$\mathbf{A} \mathbf{x} = \lambda_2 \mathbf{x} \implies \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 \\ 3x_2 \end{bmatrix} \implies 2x_1 - x_2 = 3x_1 \text{ és } 2x_2 - x_1 = 3x_2.$$

Innen $x_1 = -x_2$, tehát

$$W_3 = \text{Span} \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)$$

A kapott bázisvektorok jól láthatóan merőlegesek, így ezek ortogonális bázist alkotnak \mathbb{R}^2 -ben. Szükség van még a normalizálásra. Könnyen látszik, hogy mindkét vektor normája $\sqrt{2}$, így végül az ortonormált bázis:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

A kapott bázis a triviális bázisnak 45 fokos elforgatása. Ez véletlen? A következő fejezetben kiderül. ☺

8. fejezet

Lineáris transzformációk

8.1. Vektortérhomomorfizmus

Vezessük be a következő leképezést. Legyen $\mathbf{A} \in \mathbb{R}^{m \times n}$ és $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\varphi(\mathbf{x}) := \mathbf{Ax}$. Meglátjuk, hogy ez nem csak egy szimpla leképezés, rendelkezik pár igen kedvező tulajdonsággal. Vizsgáljuk meg ugyanis, hogyan viselkedik a leképezés két vektor összege esetén:

$$\varphi(\mathbf{x} + \mathbf{y}) = \mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{Ax} + \mathbf{Ay} = \varphi(\mathbf{x}) + \varphi(\mathbf{y}).$$

Azt látjuk tehát, hogy a vektorok összegének képe a vektorok képeinek összege. Vizsgáljuk még meg a skalárszorítás kapcsolatát a leképezéssel:

$$\varphi(\lambda \mathbf{x}) = \mathbf{A}(\lambda \mathbf{x}) = \lambda \mathbf{Ax} = \lambda \varphi(\mathbf{x})$$

A skalárszoros képe tehát a kép skalárszorosa, más szóval a skalár kiemelhető volt a leképezésből. Innen már levonhatjuk a következtetést, hogy a fenn definiált leképezés mit csinál tetszőleges lineáris kombinációval:

$$\varphi(\lambda_1 \mathbf{a}_1 + \cdots + \lambda_k \mathbf{a}_k) = \lambda_1 \varphi(\mathbf{a}_1) + \cdots + \lambda_k \varphi(\mathbf{a}_k).$$

Minden olyan függvényt, mely rendelkezik a fenti tulajdonságokkal, vektortérhomomorfizmusnak fogunk nevezni.

47. Definíció. Legyen $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$. φ vektortérhomomorfizmus (vagy (homogén) lineáris leképezés), ha bármely $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ -re $\varphi(\mathbf{u} + \mathbf{v}) = \varphi(\mathbf{u}) + \varphi(\mathbf{v})$, illetve minden $\lambda \in \mathbb{R}$ és $\mathbf{u} \in \mathbb{R}^n$ -re $\varphi(\lambda \mathbf{u}) = \lambda \varphi(\mathbf{u})$. Ha egy lineáris leképezés bijektív, akkor vektortér-izomorfizmusnak nevezzük.

Bijektív leképezés egyszerre jelent injektív leképezést (kölsönösen egyértelmű) és szürjektív leképezést (a képhalmaz minden eleme hozzá van rendelve valamely elemhez). A lineáris leképezéseket művelettartó leképezéseknek is mondják, pontosan a tulajdonságai miatt. Megmutatjuk, hogy bázishoz egyértelműen tudunk csak vektorokat rendelni.

49. Tétel (Egyértelmű kiterjesztési tétel). *Legyen $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis \mathbb{R}^n -ben, $\mathbf{w}_1, \dots, \mathbf{w}_n$ tetszőleges vektorok \mathbb{R}^m -ben. Ekkor pontosan egy $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineáris leképezés létezik, melyre $\varphi(\mathbf{e}_i) = \mathbf{w}_i$ minden i -re.*

Bizonyítás. Elsőnek az egyértelműséget mutatjuk meg. Legyen φ egy jó leképezés, és $\mathbf{x} := x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n \in \mathbb{R}^n$. Ekkor

$$\varphi(\mathbf{x}) = \varphi(x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n) = x_1\varphi(\mathbf{e}_1) + \dots + x_n\varphi(\mathbf{e}_n) = x_1\mathbf{w}_1 + \dots + x_n\mathbf{w}_n.$$

A fentiből a létezés is adódik definíció végett. □

A bevezető példából látszik, hogy lineáris leképezéseket mátrix segítségével is megadhatunk:

48. Definíció. *Ha $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis \mathbb{R}^n -ben, $\mathbf{f}_1, \dots, \mathbf{f}_m$ bázis \mathbb{R}^m -ben, $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, akkor a φ mátrixa az $\mathbf{e}; \mathbf{f}$ bázispárban:*

$$[\varphi]^{\mathbf{e}; \mathbf{f}} := [[\varphi(\mathbf{e}_1)]_{\mathbf{f}}, \dots, [\varphi(\mathbf{e}_n)]_{\mathbf{f}}] \in \mathbb{R}^{m \times n}.$$

Ha például a leképezés $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ és

$$\varphi(\mathbf{x}) = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} \mathbf{x}$$

és mindkét vektortérben a triviális bázissal koordinálunk, akkor

$$[\varphi(\mathbf{e}_1)]_{\mathbf{e}_1, \mathbf{e}_2} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad [\varphi(\mathbf{e}_2)]_{\mathbf{e}_1, \mathbf{e}_2} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad [\varphi(\mathbf{e}_3)]_{\mathbf{e}_1, \mathbf{e}_2} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

ez pedig egyértelmű, mint láttuk. Továbbá tetszőleges $\mathbf{x} \in \mathbb{R}^n$ esetén $[\varphi(\mathbf{x})]_{\mathbf{f}} = [\varphi]^{\mathbf{e}; \mathbf{f}}[\mathbf{x}]_{\mathbf{e}}$.

49. Definíció. $\mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$ jelölje az összes \mathbb{R}^n -ből \mathbb{R}^m -be képező vektortérhomomorfizmusok halmazát.

Lineáris leképezések összegét és számmal vett szorzatát teljesen ugyanúgy definiáljuk, mint a többi függvény esetén: legyenek $\varphi, \psi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$ és legyen $\varphi + \psi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ úgy, hogy minden $\mathbf{u} \in \mathbb{R}^n$ -re $(\varphi + \psi)(\mathbf{u}) := \varphi(\mathbf{u}) + \psi(\mathbf{u})$. Hasonlóan $\lambda \in \mathbb{R}$ mellett $\lambda\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ úgy, hogy $(\lambda\varphi)(\mathbf{u}) := \lambda\varphi(\mathbf{u})$. A kompozícióval persze a sorrend végett vigyázni kell:

50. Definíció. *Legyen $V_1 := \mathbb{R}^n$, $V_2 := \mathbb{R}^m$, $V_3 := \mathbb{R}^s$, $\varphi \in \mathcal{H}om(V_1, V_2)$, $\psi \in \mathcal{H}om(V_2, V_3)$. Legyen $\psi\varphi : V_1 \rightarrow V_3$ úgy, hogy $\mathbf{u} \in V_1$ -re $(\psi\varphi)(\mathbf{u}) := \psi(\varphi(\mathbf{u}))$. Szokás még a $\psi \circ \varphi$ jelölés is.*

8.2. Képtér és magtér

51. Definíció. Legyen $\varphi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$. Ekkor
 φ képtere: $\mathcal{I}m\varphi := \{\mathbf{v} \in \mathbb{R}^m : \exists \mathbf{u} \in \mathbb{R}^n : \varphi(\mathbf{u}) = \mathbf{v}\},$
 φ magtere: $\mathcal{K}er\varphi := \{\mathbf{x} \in \mathbb{R}^n : \varphi(\mathbf{x}) = \mathbf{0}\}.$

A képtér tehát az összes képvektor halmaza, míg a magtérben azok a vektorok vannak, melyeket a leképezés a nullvektorba visz (persze \mathbb{R}^m -beli nullvektorba). Fontos észrevétel, hogy a nullvektor mindig benne van a magtérben. Az már egy érdekesebb következmény, hogy a fenti két halmaz mindig altér:

50. Tétel. $\varphi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$ esetén $\mathcal{I}m\varphi \leq \mathbb{R}^m$ és $\mathcal{K}er\varphi \leq \mathbb{R}^n$.

Bizonyítás. Legyenek $\mathbf{u}', \mathbf{v}' \in \mathcal{I}m\varphi$. Ekkor $\exists \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, melyekre $\varphi(\mathbf{u}) = \mathbf{u}'$, $\varphi(\mathbf{v}) = \mathbf{v}'$. Mivel

$$\varphi(\mathbf{u} + \mathbf{v}) = \mathbf{u}' + \mathbf{v}',$$

ezért a vektorok összege is a képtérben lesz. Hasonlóan mutatható meg a skalárszoros is. A magtér esetén ha két magtérbeli vektort veszünk, azok képe a nullvektor, és mivel vektorok összegének képe a képek összege, így az összegvektor is benne lesz a magtérben. Hasonlóan mutatható meg a skalárszoros itt is. \square

Vegyük észre, hogy a leképezés szürjektivitása azzal ekvivalens, hogy $\mathcal{I}m\varphi = \mathbb{R}^m$. Megmutatjuk, hogy lineáris leképezések injektivitását eldönteni nevetségesen könnyű feladat. Mivel

$$\varphi(\mathbf{x}) = \varphi(\mathbf{y}) \iff \varphi(\mathbf{x} - \mathbf{y}) = \mathbf{0} \iff \mathbf{x} - \mathbf{y} \in \mathcal{K}er\varphi,$$

az injektivitás pedig a fenti állítások valamelyikéből következtet arra, hogy $\mathbf{x} = \mathbf{y}$, ezért ez pontosan akkor teljesül, ha $\mathcal{K}er\varphi = \{\mathbf{0}\}$.

52. Definíció. Legyen $\varphi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$. φ rangja alatt $r(\varphi) := \dim \mathcal{I}m\varphi$ -t értjük, defektusa alatt pedig $d(\varphi) := \dim \mathcal{K}er\varphi$ -t.

51. Tétel (Dimenzióösszefüggés). $\varphi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$ esetén $\dim \mathcal{I}m\varphi + \dim \mathcal{K}er\varphi = \dim \mathbb{R}^n$.

Bizonyítás. Később. \square

Nézzünk egy példát. Legyen $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$,

$$\varphi \left(\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \right) := \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$$

Elsőnek győződjünk meg róla, hogy a fenti leképezés valóban lineáris-e.

$$\varphi \left(\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \right) = \begin{bmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \alpha_3 + \beta_3 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \varphi \left(\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \right) + \varphi \left(\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \right)$$

Hasonlóan mutatható meg a másik feltétel, tehát ez valóban lineáris leképezés. Határozzuk meg a magterét! Olyan \mathbb{R}^3 -beli vektorokat keresünk tehát, melyek az \mathbb{R}^2 -beli nullvektorba képeznek:

$$\varphi \left(\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \right) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \alpha_1 = 0 \text{ és } \alpha_2 = 0.$$

Tehát $\text{Ker} \varphi = \text{Span}(\mathbf{e}_3)$. Innen már az is látszik, hogy a leképezés nem lehet injektív. Határozzuk meg a képteret! Itt most könnyen látszik, hogy minden \mathbb{R}^2 -beli vektorhoz található (ráadásul végtelen sok) \mathbb{R}^3 -beli vektor, melynek ő a képe, így a képtér \mathbb{R}^2 , a szürjektivitás tehát teljesül. Szépen látszik a dimenzióösszefüggés is. Írjuk fel a leképezés mátrixát az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3; \mathbf{e}_1, \mathbf{e}_2$ bázispárban, tehát mindkét tér triviális bázisában:

$$[\varphi]^{\mathbf{e};\mathbf{e}} = [[\varphi(\mathbf{e}_1)]_{\mathbf{e}}, [\varphi(\mathbf{e}_2)]_{\mathbf{e}}, [\varphi(\mathbf{e}_3)]_{\mathbf{e}}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

53. Definíció. A $\text{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ -beli lineáris leképezéseket lineáris transzformációknak nevezzük.

8.3. Diagonalizálhatóság, sajátérték, sajátvektor

A címben szereplő már ismerősen csengő fogalmak értelmezhetők lineáris transzformációkra is, hiszen ezek mint láttuk, reprezentálhatók mátrixszal. Nem kevés olyan szakirodalom van, melyek a transzformációk tárgyalását a mátrixok előtt teszi meg. Mi már ismerjük azt a fogalmat hogy diagonalizálhatóság, nézzük meg hogy transzformációk esetén hogyan válhat ez hasznunkra. Mindenek előtt definiáljuk a sajátérték és sajátvektor fogalmát lineáris transzformációkra.

54. Definíció. Legyen $\varphi \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^n)$. A $\mathbf{0} \neq \mathbf{u} \in \mathbb{R}^n$ sajátvektora φ -nek, ha létezik olyan $\lambda_0 \in \mathbb{R}$, amivel $\varphi(\mathbf{u}) = \lambda_0 \mathbf{u}$. Ilyenkor λ_0 az \mathbf{u} sajátvektorhoz tartozó sajátértéke φ -nek.

Vizsgálódjunk kicsit. Legyen $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis \mathbb{R}^n -ben, a $\varphi \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ leképezés mátrixa pedig $\mathbf{A} := [\varphi]^{\mathbf{e}} (= [\varphi]^{\mathbf{e};\mathbf{e}})$, $\lambda_0 \in \mathbb{R}$. Ekkor λ_0 pontosan akkor sajátértéke φ -nek, ha $\exists \mathbf{0} \neq \mathbf{u} \in \mathbb{R}^n$, melyre

$$\varphi(\mathbf{u}) = \lambda_0 \mathbf{u} \iff [\varphi(\mathbf{u})]_{\mathbf{e}} = \lambda_0 [\mathbf{u}]_{\mathbf{e}} \iff [\varphi]^{\mathbf{e}} [\mathbf{u}]_{\mathbf{e}} = \lambda_0 [\mathbf{u}]_{\mathbf{e}} \iff \mathbf{A} \mathbf{x} = \lambda_0 \mathbf{x},$$

tehát pontosan akkor, ha \mathbf{x} jobb oldali sajátvektora \mathbf{A} -nak, ami azzal ekvivalens, hogy

$$|\mathbf{A} - \mathbf{I}_n \lambda_0| = 0.$$

55. Definíció. $\varphi \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ karakterisztikus polinomja $k_\varphi(\lambda) = k_{[\varphi]^e}(\lambda)$.

Belátunk egy fontos tételt, melyet a mátrixok sajátvektorainak vizsgálatánál is megtehetünk volna. Megmutatjuk, hogy a különböző sajátértékekhez tartozó sajátvektorok lineárisan függetlenek.

52. Tétel. Legyen $\varphi \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ és $\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{R}^n$ sajátvektora φ -nek, továbbá $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ a megfelelő sajátértékek, melyek páronként különbözőek. Ekkor $\mathbf{u}_1, \dots, \mathbf{u}_k$ lineárisan független sajátvektorrendszer.

Bizonyítás. Teljes indukcióval bizonyítunk. $k = 1$ esete világos, hiszen egy sajátvektor nem lehet nullvektor. Tegyük fel, hogy állításunk igaz $k - 1$ -re, tehát $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$ lineárisan független. Tegyük fel indirekt úton, hogy $\mathbf{u}_1, \dots, \mathbf{u}_k$ összefüggő. Ekkor jól tudjuk, hogy \mathbf{u}_k lineárisan függ a többi vektortól, tehát előáll $\mathbf{u}_k = \sum_{j=1}^{k-1} \alpha_j \mathbf{u}_j$ alakban. Innen

$$\varphi(\mathbf{u}_k) = \sum_{j=1}^{k-1} \alpha_j \underbrace{\varphi(\mathbf{u}_j)}_{\lambda_j \mathbf{u}_j} = \sum_{j=1}^{k-1} \alpha_j \lambda_j \mathbf{u}_j = \lambda_k \mathbf{u}_k.$$

Másrészt $\varphi(\mathbf{u}_k) = \lambda_k \mathbf{u}_k$, így

$$\varphi(\mathbf{u}_k) = \sum_{j=1}^{k-1} \alpha_j \lambda_k \mathbf{u}_j = \lambda_k \mathbf{u}_k$$

szintén teljesül. Kétféleképp állítottuk elő $\lambda_k \mathbf{u}_k$ -t, tehát

$$\mathbf{0} = \sum_{j=1}^{k-1} \alpha_j (\lambda_j - \lambda_k) \mathbf{u}_j.$$

Tudjuk, hogy $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$ független, tehát minden $j = 1, \dots, k - 1$ -re $\alpha_j (\lambda_j - \lambda_k) = 0$, viszont $\lambda_j \neq \lambda_k$ (mert ezek páronként különbözőek), így $\alpha_j = 0$. Ez viszont azt jelenti, hogy $\mathbf{u}_k = \mathbf{0}$, ami ellentmond annak, hogy sajátvektor. \square

Egyszerű következménye a fentieknek, hogy ha éppen n darab páronként különböző sajátérték van φ -nek, akkor szükségképpen létezik φ -nek sajátbázisa \mathbb{R}^n -ben, hisz kapunk n db független vektorból álló rendszert. Vigyázzunk, hiszen ez csak elégséges feltétel sajátbázis létezésére. Előfordulhat, hogy létezik sajátbázis, ám nincs n darab különböző sajátérték.

8.4. Általánosítás tetszőleges vektorterekre

Bár korábban megfogalmaztuk mit értünk vektortér alatt, eddig szinte végig \mathbb{R}^n -ben dolgoztunk, ami persze vektortér, de csak egy speciális vektortér. Ha egy vektortér véges és n dimenziós, akkor szoros kapcsolat létezik az adott tér és \mathbb{R}^n között, ám ennek vizsgálatához mindenképpen szükséges általánosítanunk az eddigi ismereteink nagy részét. Mostantól jelöljön V egy \mathbb{R} feletti vektorteret.

56. Definíció. Az \mathbb{R} feletti V vektortér dimenziója:

$$\dim V = \begin{cases} 0 & \text{ha } V = \{\mathbf{0}\} \\ \text{egy tetszőleges bázis elemszáma} & \text{ha } V \neq \{\mathbf{0}\} \text{ és van} \\ & V\text{-ben véges generátorrendszer} \\ \infty & \text{különben.} \end{cases}$$

Tekintsük a $\mathbb{R} \rightarrow \mathbb{R}$ függvények halmazát. Megmutatható, hogy ez vektortér. Ebben a vektortérben a polinomfüggvények alteret alkotnak, melynek azonban nincs véges generátorrendszere, így dimenziója végtelen.

57. Definíció. Legyenek V_1, V_2 vektorterek \mathbb{R} felett, illetve $\varphi : V_1 \rightarrow V_2$. φ vektortérhomomorfizmus (vagy (homogén) lineáris leképezés), ha bármely $\mathbf{u}, \mathbf{v} \in V_1$ -re $\varphi(\mathbf{u} + \mathbf{v}) = \varphi(\mathbf{u}) + \varphi(\mathbf{v})$, illetve minden $\lambda \in \mathbb{R}$ és $\mathbf{u} \in V_1$ -re $\varphi(\lambda \mathbf{u}) = \lambda \varphi(\mathbf{u})$. Ha egy lineáris leképezés bijektív, akkor vektortér-izomorfizmusnak nevezzük.

53. Tétel (Egyértelmű kiterjesztési tétel). Legyenek V_1, V_2 vektorterek \mathbb{R} felett, $\dim V_1 = n > 0$. Legyen továbbá $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V_1 -ben, $\mathbf{w}_1, \dots, \mathbf{w}_n$ tetszőleges vektorok V_2 -ben. Ekkor pontosan egy $\varphi : V_1 \rightarrow V_2$ lineáris leképezés létezik, melyre $\varphi(\mathbf{e}_i) = \mathbf{w}_i$ minden i -re.

58. Definíció. Ha $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V_1 -ben, $\mathbf{f}_1, \dots, \mathbf{f}_m$ bázis V_2 -ben, $\varphi : V_1 \rightarrow V_2$, akkor a φ mátrixa az $\mathbf{e}; \mathbf{f}$ bázispárban:

$$[\varphi]^{\mathbf{e}; \mathbf{f}} := [[\varphi(\mathbf{e}_1)]_{\mathbf{f}}, \dots, [\varphi(\mathbf{e}_n)]_{\mathbf{f}}] \in \mathbb{R}^{m \times n}.$$

59. Definíció. A $\mathcal{H}om(V, V)$ -beli lineáris leképezéseket lineáris transzformációknak nevezzük.

60. Definíció. Legyenek V_1, V_2 vektorterek \mathbb{R} felett, $\varphi, \psi \in \mathcal{H}om(V_1, V_2)$, és legyen $\varphi + \psi : V_1 \rightarrow V_2$ úgy, hogy minden $\mathbf{u} \in V_1$ -re $(\varphi + \psi)(\mathbf{u}) := \varphi(\mathbf{u}) + \psi(\mathbf{u})$. Hasonlóan $\lambda \in \mathbb{R}$ mellett $\lambda \varphi : V_1 \rightarrow V_2$ úgy, hogy $(\lambda \varphi)(\mathbf{u}) := \lambda \varphi(\mathbf{u})$.

61. Definíció. Legyenek V_1, V_2 vektorterek \mathbb{R} felett, $\varphi \in \mathcal{H}om(V_1, V_2)$. Ekkor

φ képtere: $\mathcal{I}m\varphi := \{\mathbf{v} \in V_2 : \exists \mathbf{u} \in V_1 : \varphi(\mathbf{u}) = \mathbf{v}\},$

φ magtere: $\mathcal{K}er\varphi := \{\mathbf{x} \in V_1 : \varphi(\mathbf{x}) = \mathbf{0}\}.$

54. Tétel. $\varphi \in \mathcal{H}om(V_1, V_2)$ esetén $\mathcal{I}m\varphi \leq V_2$ és $\mathcal{K}er\varphi \leq V_1$.

55. Tétel. V_1, V_2 vektorterek között pontosan akkor létezik vektortér izomorfizmus, ha a dimenziójuk megegyezik. Formálisan:

$$V_1 \simeq V_2 \iff \dim V_1 = \dim V_2.$$

Bizonyítás. Nézzük a \Rightarrow irányt. Tegyük fel tehát, hogy $\varphi : V_1 \rightarrow V_2$ vektortér-izomorfizmus (tehát bijektív homomorfizmus). Legyen $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V_1 -ben. Meg fogjuk mutatni, hogy ekkor $\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_n)$ bázis V_2 -ben, ahonnan már az állítás ezen iránya következik. Ehhez meg kell mutatnunk, hogy a rendszer generátorrendszer, és hogy lineárisan független. Kezdjük az előbbivel. Legyen $\mathbf{v}' \in V_2$. Mivel φ szürjektív, létezik $\mathbf{u} \in V_1$, melyre $\varphi(\mathbf{u}) = \mathbf{v}'$. A bázisban felírva

$$\varphi(u_1\mathbf{e}_1 + \dots + u_n\mathbf{e}_n) = u_1\varphi(\mathbf{e}_1) + \dots + u_n\varphi(\mathbf{e}_n),$$

tehát a tetszőleges V_2 -beli vektort előállítottuk a kívánt rendszerrel. Tekintsük most a függetlenséget. Tegyük fel, hogy

$$\alpha_1\varphi(\mathbf{e}_1) + \dots + \alpha_n\varphi(\mathbf{e}_n) = \mathbf{0} \implies \varphi(\alpha_1\mathbf{e}_1 + \dots + \alpha_n\mathbf{e}_n) = \mathbf{0} = \varphi(\mathbf{0}).$$

Mivel φ injektív, ezért $\alpha_1\mathbf{e}_1 + \dots + \alpha_n\mathbf{e}_n = \mathbf{0}$, ahonnan $\alpha_1 = \dots = \alpha_n = 0$. Tekintsük most a \Leftarrow irányt. Legyen $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V_1 -ben, $\mathbf{f}_1, \dots, \mathbf{f}_n$ bázis V_2 -ben. Tudjuk, hogy ekkor egyértelműen létezik $\varphi \in \mathcal{H}om(V_1, V_2)$, melyre $\varphi(\mathbf{e}_i) = \mathbf{f}_i$ minden i -re. Azt állítjuk, hogy ez a φ bijektív. Ehhez a szürjektivitást és injektivitást kell belátni. Kezdjük az előbbivel. Legyen $\mathbf{v}' \in V_2$, és írjuk fel a V_2 bázisában:

$$\mathbf{v}' = \beta_1\mathbf{f}_1 + \dots + \beta_n\mathbf{f}_n = \beta_1\varphi(\mathbf{e}_1) + \dots + \beta_n\varphi(\mathbf{e}_n) = \varphi(\beta_1\mathbf{e}_1 + \dots + \beta_n\mathbf{e}_n).$$

Az injektivitás alatt mint már tudjuk azt kell megmutatni, hogy $\mathcal{K}er\varphi = \{\mathbf{0}\}$. Tegyük fel, hogy $\mathbf{x} \in \mathcal{K}er\varphi$, vagyis $\varphi(\mathbf{x}) = \mathbf{0}$. Bázisban felírva:

$$\mathbf{0} = \varphi(x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n) = x_1\varphi(\mathbf{e}_1) + \dots + x_n\varphi(\mathbf{e}_n) = x_1\mathbf{f}_1 + \dots + x_n\mathbf{f}_n,$$

ez utóbbi független, így $x_1 = \dots = x_n = 0$, tehát $\mathcal{K}er\varphi = \{\mathbf{0}\}$ lehet csak, és készen vagyunk. \square

56. Tétel (Dimenzióösszefüggés). *Legyenek V_1, V_2 vektorterek \mathbb{R} felett, $\varphi \in \mathcal{H}om(V_1, V_2)$ esetén $\dim \mathcal{I}m\varphi + \dim \mathcal{K}er\varphi = \dim V_1$.*

Bizonyítás. Legyen $\dim V_1 = n$. Ha $n = 0$, akkor $\dim \mathcal{I}m\varphi = \dim \mathcal{K}er\varphi = 0$ és rendben vagyunk. Legyen tehát $n > 0$ és legyen k a defektus, azaz $k = \dim \mathcal{K}er\varphi$. Ha $k = n$, akkor $\mathcal{K}er\varphi = V_1$, a képtér tehát szükségképpen $\mathcal{I}m\varphi = \{\mathbf{0}\}$ (hiszen ekkor „minden út a nullába vezet”). Legyen tehát $0 < k < n$ és $\mathbf{a}_1, \dots, \mathbf{a}_k$ bázis $\mathcal{K}er\varphi$ -ben. Ez nyilván független rendszer V_1 -ben, melyről tudjuk, hogy kiegészíthető V_1 bázisává. Legyenek ezek a hiányzó vektorok $\mathbf{b}_1, \dots, \mathbf{b}_{n-k}$, tehát most

$$\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{b}_1, \dots, \mathbf{b}_{n-k}$$

bázis V_1 -ben. Megmutatjuk, hogy ekkor $\varphi(\mathbf{b}_1), \dots, \varphi(\mathbf{b}_{n-k})$ bázis $\mathcal{I}m\varphi$ -ben, ami igazolja állításunkat. Elsőnek azt mutatjuk meg, hogy ez a rendszer generátorrendszer $\mathcal{I}m\varphi$ -ben. Tudjuk, hogy $\mathcal{I}m\varphi$ -ben $\varphi(\mathbf{u})$ alakú elemek vannak, ahol $\mathbf{u} \in V_1$. A V_1 bázisában felírva legyen

$$\begin{aligned} \mathbf{u} &= \alpha_1 \mathbf{a}_1 + \dots + \alpha_k \mathbf{a}_k + \beta_1 \mathbf{b}_1 + \dots + \beta_{n-k} \mathbf{b}_{n-k}, \text{ így} \\ \varphi(\mathbf{u}) &= \varphi(\alpha_1 \mathbf{a}_1 + \dots + \alpha_k \mathbf{a}_k + \beta_1 \mathbf{b}_1 + \dots + \beta_{n-k} \mathbf{b}_{n-k}) = \\ &= \alpha_1 \varphi(\mathbf{a}_1) + \dots + \alpha_k \varphi(\mathbf{a}_k) + \beta_1 \varphi(\mathbf{b}_1) + \dots + \beta_{n-k} \varphi(\mathbf{b}_{n-k}). \end{aligned}$$

Mivel $\mathbf{a}_1, \dots, \mathbf{a}_k$ $\mathcal{K}er\varphi$ -beli (bázis)vektorok, ezért a képük biztosan a nullvektor, tehát

$$\varphi(\mathbf{u}) = \beta_1 \varphi(\mathbf{b}_1) + \dots + \beta_{n-k} \varphi(\mathbf{b}_{n-k})$$

tehát $\varphi(\mathbf{u}) \in \mathcal{I}m\varphi$ -t előállítottuk a rendszerrel, ez a rész tehát kész. Még meg kell mutatnunk a függetlenséget. Ehhez szokásos módon tegyük fel, hogy

$$\begin{aligned} \gamma_1 \varphi(\mathbf{b}_1) + \dots + \gamma_{n-k} \varphi(\mathbf{b}_{n-k}) &= \mathbf{0} \implies \\ \implies \varphi(\gamma_1 \mathbf{b}_1 + \dots + \gamma_{n-k} \mathbf{b}_{n-k}) &= \mathbf{0} \end{aligned}$$

tehát $\gamma_1 \mathbf{b}_1 + \dots + \gamma_{n-k} \mathbf{b}_{n-k} \in \mathcal{K}er\varphi$. $\ker \varphi$ -ben $\mathbf{a}_1, \dots, \mathbf{a}_k$ bázis, tehát felírható ez a vektor ebben a bázisban:

$$\begin{aligned} \gamma_1 \mathbf{b}_1 + \dots + \gamma_{n-k} \mathbf{b}_{n-k} &= \delta_1 \mathbf{a}_1 + \dots + \delta_k \mathbf{a}_k \implies \\ \implies \gamma_1 \mathbf{b}_1 + \dots + \gamma_{n-k} \mathbf{b}_{n-k} - \delta_1 \mathbf{a}_1 - \dots - \delta_k \mathbf{a}_k &= \mathbf{0} \end{aligned}$$

Mivel $\mathbf{a}_1, \dots, \mathbf{a}_k, \mathbf{b}_1, \dots, \mathbf{b}_{n-k}$ bázis V_1 -ben, így minden együttható nulla, szükségképpen tehát minden $\gamma_i = 0$, és készen vagyunk. Ha $k = 0$, akkor ugyanez a procedura csak \mathbf{a} -k nincsenek, és $\mathbf{b}_1, \dots, \mathbf{b}_n$ lesz egy bázis V_1 -ben. \square

57. Tétel (Szorzástétel). Legyenek V_1, V_2, V_3 vektorterek \mathbb{R} felett, dimenziójuk rendre n, m, s (pozitív egészek), $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V_1 -ben, $\mathbf{f}_1, \dots, \mathbf{f}_m$ bázis V_2 -ben, $\mathbf{g}_1, \dots, \mathbf{g}_s$ bázis V_3 -ban, továbbá $\varphi \in \mathcal{H}om(V_1, V_2)$, $\psi \in \mathcal{H}om(V_2, V_3)$. Ekkor

$$[\psi\varphi]^{\mathbf{e};\mathbf{g}} = [\psi]^{\mathbf{f};\mathbf{g}}[\varphi]^{\mathbf{e};\mathbf{f}}$$

Bizonyítás. Tudjuk, hogy definíció szerint $[\varphi]^{\mathbf{e};\mathbf{f}} = [[\varphi(\mathbf{e}_1)]_{\mathbf{f}}, \dots, [\varphi(\mathbf{e}_n)]_{\mathbf{f}}] \in \mathbb{R}^{m \times n}$. Tehát minden szóbjövő i -re

$$\varphi(\mathbf{e}_i) = \sum_{j=1}^m [\varphi]_i^{\mathbf{e};\mathbf{f}} \mathbf{f}_j$$

hiszen minden oszlop az \mathbf{f} bázisban van felírva. Hasonlóan $[\psi]^{\mathbf{f};\mathbf{g}} = [[\psi(\mathbf{f}_1)]_{\mathbf{g}}, \dots, [\psi(\mathbf{f}_m)]_{\mathbf{g}}]$ ami pedig $\mathbb{R}^{s \times m}$ -beli, ezért minden j -re

$$\psi(\mathbf{f}_j) = \sum_{t=1}^s [\psi]_j^{\mathbf{f};\mathbf{g}} \mathbf{g}_t.$$

Végül pedig $[\psi\varphi]^{\mathbf{e};\mathbf{g}} = [[(\psi\varphi)(\mathbf{e}_1)]_{\mathbf{g}}, \dots, [(\psi\varphi)(\mathbf{e}_n)]_{\mathbf{g}}] \in \mathbb{R}^{s \times n}$, így minden i -re

$$(\psi\varphi)(\mathbf{e}_i) = \sum_{t=1}^s [\psi\varphi]_i^{\mathbf{e};\mathbf{g}} \mathbf{g}_t.$$

Kezdjük átalakítani:

$$\begin{aligned} (\psi\varphi)(\mathbf{e}_i) &= \psi(\varphi(\mathbf{e}_i)) = \psi\left(\sum_{j=1}^m [\varphi]_i^{\mathbf{e};\mathbf{f}} \mathbf{f}_j\right) \stackrel{\text{lin.}}{=} \sum_{j=1}^m [\varphi]_i^{\mathbf{e};\mathbf{f}} \psi(\mathbf{f}_j) = \\ &= \sum_{j=1}^m [\varphi]_i^{\mathbf{e};\mathbf{f}} \sum_{t=1}^s [\psi]_j^{\mathbf{f};\mathbf{g}} \mathbf{g}_t = \sum_{t=1}^s \left(\sum_{j=1}^m [\varphi]_i^{\mathbf{e};\mathbf{f}} [\psi]_j^{\mathbf{f};\mathbf{g}}\right) \mathbf{g}_t = \sum_{t=1}^s \left(\sum_{j=1}^m [\psi]_j^{\mathbf{f};\mathbf{g}} [\varphi]_i^{\mathbf{e};\mathbf{f}}\right) \mathbf{g}_t, \end{aligned}$$

Egy \mathbf{g} bázisban felírt vektort kaptunk, a felírás egyértelmű, az együttható pedig éppen a szorzatmátrix t -edik sorának i -edik eleme. \square

Most megnézzük, hogyan változik egy lineáris transzformáció mátrixa új bázisra való áttéréskor. Ismerős lesz az eredmény:

58. Tétel (Új bázisra való áttérés). Legyen V vektortér \mathbb{R} felett, $\dim V = n > 0$, valamint $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V -ben, és $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ egy másik bázis V -ben. Ekkor egyértelműen létezik olyan $\tau \in \mathcal{H}om(V, V)$, melyre $\tau(\mathbf{e}_i) = \mathbf{e}'_i$ minden i -re. Legyen $\mathbf{D} := [\tau]^e$. Ekkor \mathbf{D} invertálható és bármely $\varphi \in \mathcal{H}om(V, V)$ esetén

$$[\varphi]^{\mathbf{e}'} = \mathbf{D}^{-1}[\varphi]^e \mathbf{D}$$

Bizonyítás. Az egyértelmű kiterjesztési tételt a bázisokban oda-vissza alkalmazva tudjuk, hogy egyértelműen léteznek olyan $\tau, \gamma \in \mathcal{H}om(V, V)$ transzformációk, melyekre minden szó-bajjövő i -re $\tau(\mathbf{e}_i) = \mathbf{e}'_i$ és $\gamma(\mathbf{e}'_i) = \mathbf{e}_i$, tehát

$$\gamma(\tau(\mathbf{e}_i)) = \mathbf{e}_i$$

tehát $\gamma\tau = \varepsilon$ az identitás (minden vektort helyben hagy). Ha most minden bázis helyébe \mathbf{e} -t írunk és alkalmazzuk a szorzástételt:

$$[\gamma]^\mathbf{e}[\tau]^\mathbf{e} = [\gamma\tau]^\mathbf{e} = [\varepsilon]^\mathbf{e} = \mathbf{I}_n$$

tehát $\mathbf{D} := [\tau]^\mathbf{e}$ invertálható. Legyen $\varphi \in \mathcal{H}om(V, V)$ tetszőleges. Ekkor mint tudjuk minden i -re

$$\begin{aligned} \varphi(\mathbf{e}_i) &= \sum_{j=1}^m {}_j[\varphi]^\mathbf{e} \mathbf{e}_j \implies \varphi(\mathbf{e}'_i) = \sum_{j=1}^m {}_j[\varphi]^\mathbf{e}' \mathbf{e}'_j \implies \varphi(\tau(\mathbf{e}_i)) = \sum_{j=1}^m {}_j[\varphi]^\mathbf{e}' \tau(\mathbf{e}_j) \implies \\ &\implies \gamma(\varphi(\tau(\mathbf{e}_i))) = \gamma\left(\sum_{j=1}^m {}_j[\varphi]^\mathbf{e}' \tau(\mathbf{e}_j)\right) = \sum_{j=1}^m {}_j[\varphi]^\mathbf{e}' \gamma(\tau(\mathbf{e}_j)) = \sum_{j=1}^m {}_j[\varphi]^\mathbf{e}' \mathbf{e}_j \end{aligned}$$

Másképp

$$\begin{aligned} [\varphi]^\mathbf{e}' &= [[\varphi(\mathbf{e}'_1)]_{\mathbf{e}'}, \dots, [\varphi(\mathbf{e}'_n)]_{\mathbf{e}'}] = [[\varphi(\tau(\mathbf{e}_1))]_{\mathbf{e}'}, \dots, [\varphi(\tau(\mathbf{e}_n))]_{\mathbf{e}'}] = \\ &= [[\gamma(\varphi(\tau(\mathbf{e}_1)))]_{\mathbf{e}}, \dots, [\gamma(\varphi(\tau(\mathbf{e}_n)))]_{\mathbf{e}}] \end{aligned}$$

Ezért

$$[\varphi]^\mathbf{e}' = [\gamma\varphi\tau]^\mathbf{e} = [\gamma]^\mathbf{e}[\varphi]^\mathbf{e}[\tau]^\mathbf{e} = \mathbf{D}^{-1}[\varphi]^\mathbf{e}\mathbf{D},$$

és készen vagyunk. □

Egyszerű következmény, hogy új bázisra való áttéréskor az eredeti mátrixhoz hasonló mátrixot kapunk \mathbb{R} felett. Nézzünk egy példát, legyen $\mathbf{e}_1, \mathbf{e}_2$ bázis \mathbb{R}^2 -ben, $\varphi \in \mathcal{H}om(\mathbb{R}^2, \mathbb{R}^2)$ egy lineáris transzformáció, valamint

$$[\varphi]^\mathbf{e} := \begin{bmatrix} 2 & -1 \\ 5 & -3 \end{bmatrix}$$

Írjuk fel a $[\varphi]^\mathbf{e}'$ mátrixot, ha $\mathbf{e}'_1 = -3\mathbf{e}_1 + 7\mathbf{e}_2$, $\mathbf{e}'_2 = \mathbf{e}_1 - 2\mathbf{e}_2$. Elsőnek egy gyors bázistranszformációval ellenőrizzük, hogy \mathbf{e}' valóban bázis-e:

$$\begin{array}{c} \mathbf{e}_1 \\ \mathbf{e}_2 \end{array} \left\| \begin{array}{cc} \mathbf{e}'_1 & \mathbf{e}'_2 \\ -3 & \boxed{1} \\ 7 & -2 \end{array} \right\| \begin{array}{c} \mathbf{e}'_2 \\ \mathbf{e}_2 \end{array} \left\| \begin{array}{cc} \mathbf{e}'_1 & \mathbf{e}'_2 \\ -3 & 1 \\ \boxed{1} & 0 \end{array} \right\| \begin{array}{c} \mathbf{e}'_1 \\ \mathbf{e}'_2 \end{array} \left\| \begin{array}{cc} \mathbf{e}'_1 & \mathbf{e}'_2 \\ 0 & 1 \\ 1 & 0 \end{array} \right\|$$

Tehát \mathbf{e}' valóban bázis. A \mathbf{D} mátrix i -ik oszlopa $[\mathbf{e}'_i]_{\mathbf{e}}$, tehát

$$\mathbf{D} = \begin{bmatrix} -3 & 1 \\ 7 & -2 \end{bmatrix}$$

Invertáljuk:

$$\begin{array}{c} \mathbf{e}_1 \\ \mathbf{e}_2 \end{array} \left| \begin{array}{cc|cc} \mathbf{d}_1 & \mathbf{d}_2 & \mathbf{e}_1 & \mathbf{e}_2 \\ \hline -3 & 1 & 1 & 0 \\ 7 & -2 & 0 & 1 \end{array} \right\| \begin{array}{c} \mathbf{d}_1 \\ \mathbf{e}_2 \end{array} \left| \begin{array}{cc|cc} \mathbf{d}_2 & & \mathbf{e}_1 & \mathbf{e}_2 \\ \hline -1/3 & & -1/3 & 0 \\ 1/3 & & 7/3 & 1 \end{array} \right\| \begin{array}{c} \mathbf{d}_1 \\ \mathbf{d}_2 \end{array} \left| \begin{array}{cc|cc} \mathbf{e}_1 & \mathbf{e}_2 & & \\ \hline 2 & 1 & & \\ 7 & 3 & & \end{array} \right\| \Rightarrow \mathbf{D}^{-1} = \begin{bmatrix} 2 & 1 \\ 7 & 3 \end{bmatrix}$$

Innen

$$[\varphi]^{\mathbf{e}'} = \begin{bmatrix} 2 & 1 \\ 7 & 3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 5 & -3 \end{bmatrix} \begin{bmatrix} -3 & 1 \\ 7 & -2 \end{bmatrix} = \begin{bmatrix} -62 & 19 \\ -199 & 61 \end{bmatrix}$$

Ellenőrizzünk, de hogyan? Vegyünk egy tetszőleges vektort, és írjuk fel a képét az eredeti bázisban:

$$\mathbf{x} := \begin{bmatrix} 3 \\ 2 \end{bmatrix} \Rightarrow \varphi(\mathbf{x}) = \begin{bmatrix} 2 & -1 \\ 5 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}$$

Az elemi bázistranszformáció felhasználásával írjuk fel $[\mathbf{x}]_{\mathbf{e}'}$ -t és $[\varphi(\mathbf{x})]_{\mathbf{e}'}$ -t:

$$\begin{array}{c} \mathbf{e}_1 \\ \mathbf{e}_2 \end{array} \left| \begin{array}{cc|cc} \mathbf{e}'_1 & \mathbf{e}'_2 & \mathbf{x} & \varphi(\mathbf{x}) \\ \hline -3 & 1 & 3 & 4 \\ 7 & -2 & 2 & 9 \end{array} \right\| \begin{array}{c} \mathbf{e}'_1 \\ \mathbf{e}'_2 \end{array} \left| \begin{array}{cc|cc} \mathbf{e}'_1 & & \mathbf{x} & \varphi(\mathbf{x}) \\ \hline -3 & & 3 & 4 \\ 1 & & 8 & 17 \end{array} \right\| \begin{array}{c} \mathbf{e}'_2 \\ \mathbf{e}'_1 \end{array} \left| \begin{array}{cc|cc} \mathbf{x} & \varphi(\mathbf{x}) & & \\ \hline 27 & 55 & & \\ 8 & 17 & & \end{array} \right\| \Rightarrow$$

$$\Rightarrow [\mathbf{x}]_{\mathbf{e}'} = \begin{bmatrix} 8 \\ 27 \end{bmatrix}, [\varphi(\mathbf{x})]_{\mathbf{e}'} = \begin{bmatrix} 17 \\ 55 \end{bmatrix}$$

Még azt kell megmutatni, hogy $[\varphi]^{\mathbf{e}'}[\mathbf{x}]_{\mathbf{e}'} = [\varphi(\mathbf{x})]_{\mathbf{e}'}$:

$$\begin{bmatrix} -62 & 19 \\ -199 & 61 \end{bmatrix} \begin{bmatrix} 8 \\ 27 \end{bmatrix} = \begin{bmatrix} 17 \\ 55 \end{bmatrix}$$

tehát rendben vagyunk.

9. fejezet

Bilineáris függvények

Ebben a fejezetben a skaláris szorzat fogalmát fogjuk még tovább általánosítani.

62. Definíció. Legyen $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, továbbá V vektortér \mathbb{K} felett. Az

$$\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{K}$$

függvényt (valós vagy komplex) bilineáris függvénynek, bilineáris alaknak vagy bilineáris formának hívjuk, ha teljesülnek a következők minden $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\lambda \in \mathbb{K}$ -ra:

1. $\mathcal{A}(\lambda \mathbf{x}, \mathbf{y}) = \lambda \mathcal{A}(\mathbf{x}, \mathbf{y})$,
2. $\mathcal{A}(\mathbf{x}, \lambda \mathbf{y}) = \bar{\lambda} \mathcal{A}(\mathbf{x}, \mathbf{y})$,
3. $\mathcal{A}(\mathbf{x} + \mathbf{y}, \mathbf{z}) = \mathcal{A}(\mathbf{x}, \mathbf{z}) + \mathcal{A}(\mathbf{y}, \mathbf{z})$,
4. $\mathcal{A}(\mathbf{z}, \mathbf{x} + \mathbf{y}) = \mathcal{A}(\mathbf{z}, \mathbf{x}) + \mathcal{A}(\mathbf{z}, \mathbf{y})$.

$\mathcal{A}(\mathbf{x}, \mathbf{y})$ -t rögzített \mathbf{y} mellett szokták az első változójában lineáris leképezésnek nevezni. Másik változó rögzítésnél már másodfajú lineáris leképezést szokás említeni a konjugálás végett.

Legyen V egy vektortér \mathbb{K} felett, és $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V -ben, valamint $\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{K}$ bilineáris függvény. Írjuk fel ezzel a bilineáris függvénnyel két tetszőleges lineáris kombinációt az \mathbf{e} bázisban (a 3. és 4. tulajdonságot alkalmazzuk ismételve):

$$\mathcal{A}\left(\sum_{j=1}^n x_j \mathbf{e}_j, \sum_{k=1}^n y_k \mathbf{e}_k\right) = \sum_{j=1}^n \sum_{k=1}^n x_j \bar{y}_k \mathcal{A}(\mathbf{e}_j, \mathbf{e}_k).$$

Ez a felírás fogja motiválni (különösen a kvadratikus alaknál látott forma miatt) a bilineáris függvény adott bázisbeli mátrixának bevezetését.

63. Definíció. Legyen $\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{K}$ a V -n értelmezett bilineáris függvény, továbbá $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V -ben. Ekkor $[\mathcal{A}]^e \in \mathbb{K}^{n \times n}$, és minden szóba jövő j, k -ra ${}_j[\mathcal{A}]_k^e = \mathcal{A}(\mathbf{e}_j, \mathbf{e}_k)$

A mátrixunk tehát a következőképp fest:

$$[\mathcal{A}]^e = \begin{bmatrix} \mathcal{A}(\mathbf{e}_1, \mathbf{e}_1) & \mathcal{A}(\mathbf{e}_1, \mathbf{e}_2) & \cdots & \mathcal{A}(\mathbf{e}_1, \mathbf{e}_n) \\ \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1) & \mathcal{A}(\mathbf{e}_2, \mathbf{e}_2) & \cdots & \mathcal{A}(\mathbf{e}_2, \mathbf{e}_n) \\ \vdots & \vdots & & \vdots \\ \mathcal{A}(\mathbf{e}_n, \mathbf{e}_1) & \mathcal{A}(\mathbf{e}_n, \mathbf{e}_2) & \cdots & \mathcal{A}(\mathbf{e}_n, \mathbf{e}_n) \end{bmatrix}$$

Így a korábban felírt lineáris kombinációk bilineáris alakja

$$\mathcal{A}\left(\sum_{j=1}^n x_j \mathbf{e}_j, \sum_{k=1}^n y_k \mathbf{e}_k\right) = \sum_{j=1}^n \sum_{k=1}^n x_j \overline{y_k} [{}_j[\mathcal{A}]_k^e,$$

egy speciálisabb esetben:

$$\mathcal{A}\left(\sum_{j=1}^n x_j \mathbf{e}_j, \sum_{k=1}^n x_k \mathbf{e}_k\right) = \sum_{j=1}^n \sum_{k=1}^n x_j \overline{x_k} [{}_j[\mathcal{A}]_k^e,$$

ugyanaz valós esetben ($\mathbb{K} = \mathbb{R}$):

$$\mathcal{A}\left(\sum_{j=1}^n x_j \mathbf{e}_j, \sum_{k=1}^n x_k \mathbf{e}_k\right) = \sum_{j=1}^n \sum_{k=1}^n x_j x_k [{}_j[\mathcal{A}]_k^e.$$

Speciális kvadratikus függvényhez jutunk, ha a skaláris szorzatnál megismert szimmetria fogalmakat is megköveteljük:

64. Definíció. Legyen $\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{K}$ a V -n értelmezett bilineáris függvény. Azt mondjuk, hogy \mathcal{A} Hermite-féle bilineáris függvény (valós esetben szimmetrikus bilineáris függvény is használatos), ha $\mathbf{x}, \mathbf{y} \in V$ esetén teljesül, hogy $\mathcal{A}(\mathbf{y}, \mathbf{x}) = \overline{\mathcal{A}(\mathbf{x}, \mathbf{y})}$.

Világos, hogy ha ez teljesül, akkor a kvadratikus függvényhez rendelt $[\mathcal{A}]^e$ mátrix komplex esetben önadjungált, valós esetben pedig szimmetrikus (önadjungált azt jelenti, hogy $\mathbf{A}^* = \mathbf{A}$, mely a valós szimmetria általánosítása). Így értelmezhető a kvadratikus alak.

65. Definíció. Legyen $\mathbb{K} = \mathbb{R}$ és $\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{R}$ a V -n értelmezett szimmetrikus bilineáris függvény. Ekkor az \mathcal{A} -hoz tartozó \mathcal{Q} kvadratikus alak $\mathcal{Q} : V \rightarrow \mathbb{R}$, $\mathcal{Q}(\mathbf{x}) = \mathcal{A}(\mathbf{x}, \mathbf{x})$.

66. Definíció. Legyen $\mathbb{K} = \mathbb{C}$ és $\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{C}$ a V -n értelmezett bilineáris függvény. Ekkor az \mathcal{A} -hoz tartozó \mathcal{Q} kvadratikus alak $\mathcal{Q} : V \rightarrow \mathbb{C}$, $\mathcal{Q}(\mathbf{x}) = \mathcal{A}(\mathbf{x}, \mathbf{x})$.

Érdekes lehet, hogy \mathbb{C} esetén nem tettük fel, hogy legyen a bilineáris függvény Hermite-féle. Erről majd később ejtünk szót, most vizsgáljuk meg a valós esetet és azt, hogy ott mire jó a szimmetria. Megmutatjuk, hogy szimmetrikus esetben \mathcal{A} minden értéke kifejezhető a hozzá tartozó kvadratikus alak segítségével.

59. Tétel. Legyen $\mathbb{K} = \mathbb{R}$, $\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{R}$ pedig a V -n értelmezett szimmetrikus bilineáris függvény. Ekkor \mathcal{A} minden értéke kifejezhető az \mathcal{A} -hoz tartozó \mathcal{Q} kvadratikus alak alkalmas értékei segítségével.

Bizonyítás. A szimmetria miatt

$$\mathcal{A}(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = \mathcal{A}(\mathbf{x}, \mathbf{x}) + \mathcal{A}(\mathbf{x}, \mathbf{y}) + \mathcal{A}(\mathbf{y}, \mathbf{x}) + \mathcal{A}(\mathbf{y}, \mathbf{y}) = \mathcal{A}(\mathbf{x}, \mathbf{x}) + 2\mathcal{A}(\mathbf{x}, \mathbf{y}) + \mathcal{A}(\mathbf{y}, \mathbf{y})$$

tehát

$$\mathcal{Q}(\mathbf{x} + \mathbf{y}) = \mathcal{Q}(\mathbf{x}) + 2\mathcal{A}(\mathbf{x}, \mathbf{y}) + \mathcal{Q}(\mathbf{y}) \implies \mathcal{A}(\mathbf{x}, \mathbf{y}) = \frac{\mathcal{Q}(\mathbf{x} + \mathbf{y}) - \mathcal{Q}(\mathbf{x}) - \mathcal{Q}(\mathbf{y})}{2},$$

és készen vagyunk. □

Nézzünk egy példát az eddigiekre mielőtt tovább mennénk. Tekintsük \mathbb{R}^2 -t a szokásos $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{y}$ skaláris szorzattal, továbbá legyen $\varphi \in \text{Hom}(\mathbb{R}^2, \mathbb{R}^2)$ tetszőleges, $\mathbf{e}_1, \mathbf{e}_2$ pedig ortonormált bázis \mathbb{R}^2 -ben (ilyen a triviális bázis is). Elsőnek belátunk egy állítást:

1. Állítás. A fenti feltételekkel $\mathcal{A}(\mathbf{x}, \mathbf{y}) = \langle \varphi(\mathbf{x}), \mathbf{y} \rangle$ bilineáris függvény.

Bizonyítás. A négy tulajdonságot kell megmutatnunk.

1. $\mathcal{A}(\lambda \mathbf{x}, \mathbf{y}) = \langle \lambda \varphi(\mathbf{x}), \mathbf{y} \rangle = \lambda \langle \varphi(\mathbf{x}), \mathbf{y} \rangle = \lambda \mathcal{A}(\mathbf{x}, \mathbf{y})$,
2. $\mathcal{A}(\mathbf{x}, \lambda \mathbf{y}) = \langle \varphi(\mathbf{x}), \lambda \mathbf{y} \rangle \stackrel{\text{valós szimm.}}{=} \langle \lambda \mathbf{y}, \varphi(\mathbf{x}) \rangle = \lambda \langle \mathbf{y}, \varphi(\mathbf{x}) \rangle = \lambda \langle \varphi(\mathbf{x}), \mathbf{y} \rangle = \lambda \mathcal{A}(\mathbf{x}, \mathbf{y})$,
3. $\mathcal{A}(\mathbf{x} + \mathbf{y}, \mathbf{z}) = \langle \varphi(\mathbf{x} + \mathbf{y}), \mathbf{z} \rangle = \langle \varphi(\mathbf{x}) + \varphi(\mathbf{y}), \mathbf{z} \rangle = \langle \varphi(\mathbf{x}), \mathbf{z} \rangle + \langle \varphi(\mathbf{y}), \mathbf{z} \rangle = \mathcal{A}(\mathbf{x}, \mathbf{z}) + \mathcal{A}(\mathbf{y}, \mathbf{z})$,

az utolsó állítás pedig a valós szimmetria felhasználásával teljesen hasonlóan mutatható meg. □

Tekintsük most azt a $\varphi \in \text{Hom}(\mathbb{R}^2, \mathbb{R}^2)$ lineáris leképezést, melyre $\varphi(\mathbf{e}_1) = 2\mathbf{e}_1 + \mathbf{e}_2$ és $\varphi(\mathbf{e}_2) = \mathbf{e}_1 + 2\mathbf{e}_2$. Legyenek $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ és $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2$. Ekkor

$$\begin{aligned} \varphi(\mathbf{x}) &= \varphi(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = x_1\varphi(\mathbf{e}_1) + x_2\varphi(\mathbf{e}_2) = \\ &= 2x_1\mathbf{e}_1 + x_1\mathbf{e}_2 + x_2\mathbf{e}_1 + 2x_2\mathbf{e}_2 = (2x_1 + x_2)\mathbf{e}_1 + (x_1 + 2x_2)\mathbf{e}_2 \end{aligned}$$

Hasonlóan kapjuk, hogy $\varphi(\mathbf{y}) = (2y_1 + y_2)\mathbf{e}_1 + (y_1 + 2y_2)\mathbf{e}_2$. Teljesül a szimmetria?

$$\begin{aligned}\mathcal{A}(\mathbf{y}, \mathbf{x}) &= \langle \varphi(\mathbf{y}), \mathbf{x} \rangle = \langle (2y_1 + y_2)\mathbf{e}_1 + (y_1 + 2y_2)\mathbf{e}_2, x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \rangle = \\ &= (2y_1 + y_2)x_1\langle \mathbf{e}_1, \mathbf{e}_1 \rangle + (2y_1 + y_2)x_2\langle \mathbf{e}_1, \mathbf{e}_2 \rangle + (y_1 + 2y_2)x_1\langle \mathbf{e}_2, \mathbf{e}_1 \rangle + (y_1 + 2y_2)x_2\langle \mathbf{e}_2, \mathbf{e}_2 \rangle = \\ &= (2y_1 + y_2)x_1 + (y_1 + 2y_2)x_2\end{aligned}$$

Hasonlóan $\mathcal{A}(\mathbf{x}, \mathbf{y}) = (2x_1 + x_2)y_1 + (x_1 + 2x_2)y_2$, melyek egyenlőek, így ez tehát szimmetrikus bilineáris függvény. Írjuk fel a mátrixát. Tudjuk, hogy $\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2) = \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1)$, így ezt csak egyszer kell kiszámolni:

$$\begin{aligned}\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2) &= \langle 2\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_2 \rangle = 2\langle \mathbf{e}_1, \mathbf{e}_2 \rangle + \langle \mathbf{e}_2, \mathbf{e}_2 \rangle = 1, \\ [\mathcal{A}]^{\mathbf{e}} &= \begin{bmatrix} \mathcal{A}(\mathbf{e}_1, \mathbf{e}_1) & \mathcal{A}(\mathbf{e}_1, \mathbf{e}_2) \\ \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1) & \mathcal{A}(\mathbf{e}_2, \mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} \langle \varphi(\mathbf{e}_1), \mathbf{e}_1 \rangle & 1 \\ 1 & \langle \varphi(\mathbf{e}_2), \mathbf{e}_2 \rangle \end{bmatrix} = \\ &= \begin{bmatrix} \langle 2\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 \rangle & 1 \\ 1 & \langle \mathbf{e}_1 + 2\mathbf{e}_2, \mathbf{e}_2 \rangle \end{bmatrix} = \begin{bmatrix} 2\langle \mathbf{e}_1, \mathbf{e}_1 \rangle + \langle \mathbf{e}_2, \mathbf{e}_1 \rangle & 1 \\ 1 & \langle \mathbf{e}_1, \mathbf{e}_2 \rangle + 2\langle \mathbf{e}_2, \mathbf{e}_2 \rangle \end{bmatrix} = \\ &= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\end{aligned}$$

Vizsgáljuk meg a kvadratikus alakot, leginkább előjel szempontjából:

$$\mathcal{Q}(\mathbf{x}) = \mathcal{A}(\mathbf{x}, \mathbf{x}) = (2x_1 + x_2)x_1 + (x_1 + 2x_2)x_2 = 2(x_1^2 + x_1x_2 + x_2^2),$$

ez pedig mindig pozitív ($\mathbf{x} \neq \mathbf{0}$), tehát itt is hasonlóan vizsgálhatók a definitségi kérdések. Most a korábbi tételünk komplex verzióját mondjuk ki és bizonyítjuk, látni fogjuk, hogy itt nem lesz szükségünk Hermite-féle bilineáris függvényre.

60. Tétel. *Legyen $\mathbb{K} = \mathbb{C}$, $\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{C}$ pedig a V -n értelmezett bilineáris függvény. Ekkor \mathcal{A} minden értéke kifejezhető az \mathcal{A} -hoz tartozó \mathcal{Q} kvadratikus alak alkalmas értékei segítségével.*

Bizonyítás. Írjuk fel a következőket:

$$\mathcal{A}(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = \mathcal{A}(\mathbf{x}, \mathbf{x}) + \mathcal{A}(\mathbf{x}, \mathbf{y}) + \mathcal{A}(\mathbf{y}, \mathbf{x}) + \mathcal{A}(\mathbf{y}, \mathbf{y}),$$

$$\mathcal{A}(\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y}) = \mathcal{A}(\mathbf{x}, \mathbf{x}) - \mathcal{A}(\mathbf{x}, \mathbf{y}) - \mathcal{A}(\mathbf{y}, \mathbf{x}) + \mathcal{A}(\mathbf{y}, \mathbf{y}),$$

$$\mathcal{A}(\mathbf{x} + i\mathbf{y}, \mathbf{x} + i\mathbf{y}) = \mathcal{A}(\mathbf{x}, \mathbf{x}) - i\mathcal{A}(\mathbf{x}, \mathbf{y}) + i\mathcal{A}(\mathbf{y}, \mathbf{x}) + \mathcal{A}(\mathbf{y}, \mathbf{y}),$$

$$\mathcal{A}(\mathbf{x} - i\mathbf{y}, \mathbf{x} - i\mathbf{y}) = \mathcal{A}(\mathbf{x}, \mathbf{x}) + i\mathcal{A}(\mathbf{x}, \mathbf{y}) - i\mathcal{A}(\mathbf{y}, \mathbf{x}) + \mathcal{A}(\mathbf{y}, \mathbf{y}).$$

Innen

$$\begin{aligned}\mathcal{A}(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) &= \mathcal{A}(\mathbf{x}, \mathbf{x}) + \mathcal{A}(\mathbf{x}, \mathbf{y}) + \mathcal{A}(\mathbf{y}, \mathbf{x}) + \mathcal{A}(\mathbf{y}, \mathbf{y}), \\ -\mathcal{A}(\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y}) &= -\mathcal{A}(\mathbf{x}, \mathbf{x}) + \mathcal{A}(\mathbf{x}, \mathbf{y}) + \mathcal{A}(\mathbf{y}, \mathbf{x}) - \mathcal{A}(\mathbf{y}, \mathbf{y}), \\ i\mathcal{A}(\mathbf{x} + i\mathbf{y}, \mathbf{x} + i\mathbf{y}) &= i\mathcal{A}(\mathbf{x}, \mathbf{x}) + \mathcal{A}(\mathbf{x}, \mathbf{y}) - \mathcal{A}(\mathbf{y}, \mathbf{x}) + i\mathcal{A}(\mathbf{y}, \mathbf{y}), \\ -i\mathcal{A}(\mathbf{x} - i\mathbf{y}, \mathbf{x} - i\mathbf{y}) &= -i\mathcal{A}(\mathbf{x}, \mathbf{x}) + \mathcal{A}(\mathbf{x}, \mathbf{y}) - \mathcal{A}(\mathbf{y}, \mathbf{x}) - i\mathcal{A}(\mathbf{y}, \mathbf{y}).\end{aligned}$$

Tehát összevonva

$$\mathcal{A}(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) - \mathcal{A}(\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y}) + i\mathcal{A}(\mathbf{x} + i\mathbf{y}, \mathbf{x} + i\mathbf{y}) - i\mathcal{A}(\mathbf{x} - i\mathbf{y}, \mathbf{x} - i\mathbf{y}) = 4\mathcal{A}(\mathbf{x}, \mathbf{y})$$

vagyis

$$\mathcal{A}(\mathbf{x}, \mathbf{y}) = \frac{\mathcal{Q}(\mathbf{x} + \mathbf{y}) - \mathcal{Q}(\mathbf{x} - \mathbf{y}) + i\mathcal{Q}(\mathbf{x} + i\mathbf{y}) - i\mathcal{Q}(\mathbf{x} - i\mathbf{y})}{4},$$

és készen vagyunk. □

Egy nagyon szép összefüggést mutatunk a komplex Hermite-féle bilineáris függvényekkel kapcsolatban:

61. Tétel. Legyen $\mathbb{K} = \mathbb{C}$. Ekkor bármely $\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{C}$ bilineáris függvény esetén

\mathcal{A} Hermite-féle \iff az \mathcal{A} -hoz tartozó \mathcal{Q} kvadratikus alak minden értéke valós.

Bizonyítás. Nézzük a \Rightarrow irányt. Ha \mathcal{A} Hermite-féle, akkor

$$\mathcal{Q}(\mathbf{x}) = \mathcal{A}(\mathbf{x}, \mathbf{x}) = \overline{\mathcal{A}(\mathbf{x}, \mathbf{x})} \in \mathbb{R}$$

ez így tehát rendben van, nézzük a másik irányt. Láttuk, hogy

$$\mathcal{A}(\mathbf{x}, \mathbf{y}) = \frac{\mathcal{Q}(\mathbf{x} + \mathbf{y}) - \mathcal{Q}(\mathbf{x} - \mathbf{y}) + i\mathcal{Q}(\mathbf{x} + i\mathbf{y}) - i\mathcal{Q}(\mathbf{x} - i\mathbf{y})}{4}$$

Innen a változók cseréjével

$$\mathcal{A}(\mathbf{y}, \mathbf{x}) = \frac{\mathcal{Q}(\mathbf{y} + \mathbf{x}) - \mathcal{Q}(\mathbf{y} - \mathbf{x}) + i\mathcal{Q}(\mathbf{y} + i\mathbf{x}) - i\mathcal{Q}(\mathbf{y} - i\mathbf{x})}{4}$$

Mivel $\mathcal{Q}(\lambda\mathbf{x}) = \mathcal{A}(\lambda\mathbf{x}, \lambda\mathbf{x}) = \lambda\bar{\lambda}\mathcal{A}(\mathbf{x}, \mathbf{x}) = |\lambda|^2\mathcal{Q}(\mathbf{x})$, tehát $\mathcal{Q}(\mathbf{y} - \mathbf{x}) = \mathcal{Q}((-1) \cdot (\mathbf{x} - \mathbf{y})) = \mathcal{Q}(\mathbf{x} - \mathbf{y})$, és ezt a többi tagra alkalmazva kapjuk, hogy

$$\mathcal{A}(\mathbf{y}, \mathbf{x}) = \frac{\mathcal{Q}(\mathbf{x} + \mathbf{y}) - \mathcal{Q}(\mathbf{x} - \mathbf{y}) + i\mathcal{Q}(\mathbf{x} - i\mathbf{y}) - i\mathcal{Q}(\mathbf{x} + i\mathbf{y})}{4}$$

Most feltettük, hogy \mathcal{Q} értékei valósak, így

$$\overline{\mathcal{A}(\mathbf{x}, \mathbf{y})} = \frac{\mathcal{Q}(\mathbf{x} + \mathbf{y}) - \mathcal{Q}(\mathbf{x} - \mathbf{y}) - i\mathcal{Q}(\mathbf{x} + i\mathbf{y}) + i\mathcal{Q}(\mathbf{x} - i\mathbf{y})}{4} = \mathcal{A}(\mathbf{y}, \mathbf{x}),$$

és készen vagyunk. □

Végül pedig bilineáris függvényeknél is megvizsgáljuk az új bázisra áttérést, ahogy lineáris leképezéseknél is tettük.

62. Tétel (Új bázisra való áttérés bilineáris függvényeknél). *Legyen V vektortér \mathbb{R} felett, $\dim V = n > 0$, $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázis V -ben és $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ szintén bázis V -ben. Ekkor mint már tudjuk, $\exists! \tau \in \mathcal{H}om(V, V) : \tau(\mathbf{e}_i) = \mathbf{e}'_i$ ($i = 1, \dots, n$). Legyen $\mathbf{D} = [\tau]^e$. Ekkor \mathbf{D} invertálható (már ezt is tudjuk), és tetszőleges $\mathcal{A}(\mathbf{x}, \mathbf{y}) : V \times V \rightarrow \mathbb{R}$ bilineáris függvény esetén*

$$[\mathcal{A}]^{e'} = \mathbf{D}^\top [\mathcal{A}]^e \mathbf{D}.$$

Bizonyítás. Pusztán technikai:

$$\begin{aligned} {}_j[\mathcal{A}]_k^{e'} &= \mathcal{A}(\mathbf{e}'_j, \mathbf{e}'_k) = \mathcal{A}(\tau(\mathbf{e}_j), \tau(\mathbf{e}_k)) = \mathcal{A}\left(\sum_{t=1}^n {}_t[\mathbf{D}]_j \mathbf{e}_t, \sum_{s=1}^n {}_s[\mathbf{D}]_k \mathbf{e}_s\right) = \\ &= \sum_{t=1}^n \sum_{s=1}^n {}_t[\mathbf{D}]_j {}_s[\mathbf{D}]_k \mathcal{A}(\mathbf{e}_t, \mathbf{e}_s) = \sum_{t=1}^n \sum_{s=1}^n {}_j[\mathbf{D}^\top]_{ts} {}_k[\mathbf{D}]_{st} [\mathcal{A}]_s^e = \\ &= \sum_{t=1}^n \sum_{s=1}^n {}_j[\mathbf{D}^\top]_{ts} [\mathcal{A}]_s^e {}_k[\mathbf{D}]_t = {}_j[\mathbf{D}^\top [\mathcal{A}]^e \mathbf{D}]_k, \end{aligned}$$

és ezzel készen is vagyunk. □

Említettük korábban az ún. ortogonális mátrixokat, melyeknél $\mathbf{D}^{-1} = \mathbf{D}^\top$. Ha ez teljesül az áttérési mátrixra, akkor ugyanezen feltételekkel megadott φ transzformáció és bilineáris függvény mátrixa ugyanúgy transzformálódik.