1. Geometriai vektorok

Ebben a bevezető fejezetben a középiskolában tanult vektorfogalmat ("irányított szakasz") ismételjük át, és kiegészítjük néhány új fogalommal. A későbbiekben a vektornak általánosabb értelmezést adunk. Ezért a később definiálandó vektorfogalomtól való megkülönböztetésül az irányított szakaszokat **geometriai vektor**oknak fogjuk nevezni.

1.1. A geometriai vektor fogalma

A középiskolai definíció szerint a vektor nem más, mint irányított szakasz, a síkon vagy a térben. (A sík és a tér alapfogalmak a geometriában, ezeket nem definiáljuk. A szakasz egy egyenes két pontja közé eső része (azaz ponthalmaz).) Az irányítás azt jelenti, hogy a szakasz két szélső pontját megkülönböztetjük: az egyiket kijelöljük kezdőpontnak, a másik a végpont. (Ábrázoláskor a végpontba nyilat rajzolunk.) A vektorokat általában kisbetűvel és aláhúzással fogjuk jelölni, pl. $\underline{a}$.

A geometriai vektoroknak van hosszuk és irányuk. Értelmezzük ezeket a fogalmakat a matematika nyelvén!

A síkon egy $\underline{a}$ vektor hosszát a következőképpen értelmezzük. Felveszünk egy egyenest, amely átmegy a vektor kezdőpontján. Ezután felveszünk még egy egyenest az előbbire merőlegesen, szintén a kezdőponton keresztül. Az irányított szakasznak képezzük a vetületét mindkét egyenesre (jelölje ezeket a vetületeket x és y), és a vetületek hosszából a keletkező derékszögű háromszögben Pitagorasz-tétellel kapott számot nevezzük a továbbiakban a vektor hosszának. Az $\underline{a}$ vektor hosszát jelölje $|\underline{a}|$. Tehát

$$|\underline{a}| = \sqrt{x^2 + y^2}.$$

(A merőlegesség és a vetület fogalmát a geometriából ismertnek tételezzük fel.)

A geometriai vektor irányát is tudjuk számszerűen jellemezni: pl. annak a ϕ forgásszögnek a megadásával amelyet az irányított szakasz az egyik rögzített egyenes valamelyik féltengelyével bezár. (0 és 360^o közötti szög.)

Megjegyezzük, hogy térvektorok hosszát szintén értelmezhetjük a Pitagorasz-tétellel egy harmadik, az előzőekre merőleges egyenes felvételével ($|\underline{a}| = \sqrt{x^2 + y^2 + z^2}$), az irányt pedig két szög segítségével tudjuk megadni.

A továbbiakban azokat a geometriai vektorokat, amelyeknek megegyezik a hosszuk és az irányuk, nem tekintjük különbözőnek. Ebből következik, hogy ha a síkon/a térben rögzítünk egy pontot (nevezzük ezt origónak és jelöljük O-val), akkor minden egyes geometriai vektor azonosítható egy olyan vektorral, amelynek az origó a kiindulópontja. Ezért a továbbiakban elegendő csak az origóból kiinduló vektorokkal foglalkozni.

1.1 Megjegyzés. *A sík összes vektorát (a síkvektorokat) a továbbiakban jelölje $\mathcal{E}_2$, a tér összes vektorát (a térvektorokat) pedig $\mathcal{E}_3$. Hangsúlyozzuk, hogy a térben végtelen sok sík van, de "a sík" és "egy sík a térben" két különböző dolog. A sík nem része a térnek, és így a mi értelmezésünkben az $\mathcal{E}_2$ halmaz sem részhalmaza az $\mathcal{E}_3$ halmaznak.*

Mindkét vektorhalmazt, a sík összes vektorát és a tér összes vektorát is kiegészítjük egy-egy különleges elemmel: egy olyan "szakasszal", amelynek nemcsak a kezdő-, hanem a végpontja is az origó, ezt nevezzük nullvektornak. Jelölése: $\underline{0}$. (A sík nullvektora és a tér nullvektora nem azonos, de jelölésben nem teszünk közöttük különbséget, mert a szövegkörnyezetből mindig világos, hogy melyik halmaznak a nullvektoráról van szó.)

Fontos nevezetes vektorok az ún. descartes-i egységvektorok. Ezek értelmezéséhez nevezzük x- és y-tengelynek (térvektorok esetén x-, y- és z-tengelynek) az origón átmenő, egymásra merőlegesen felvett segédegyeneseket. Az egyenesek origóval elválasztott darabjait nevezzük ki pozitív és negatív féltengelynek úgy, hogy a pozitív féltengelyek jobbsodrású rendszert alkossanak. Az ezek irányába eső egységvektorokat jelölje $\underline{i}$, $\underline{j}$ (ill. $\underline{k}$). (Bár a jelölés ugyanaz, ismételten hansúlyozzuk, hogy a sík $\underline{i}$ egységvektora nem ugyanaz, mint a tér $\underline{i}$ egységvektora, hiszen két különböző alaphalmazból vett vektorokról van szó.)

1.2. Műveletek geometriai vektorokkal

A geometriai vektorok körében sokfajta műveletet értelmezünk. Ezek közül a legalapvetőbbek az összeadás és a skalárral (valós számmal) való szorzás (a középiskolában tanult módon). Emlékeztetőül: két vektort a paralelogramma-szabállyal adunk össze, az $\underline{a}$ vektor λ számszorosán pedig azt a vektort értjük (jelölése $\lambda\underline{a}$), amelynek a nagysága $|\lambda\underline{a}| = |\lambda| \cdot |\underline{a}|$, iránya pedig $\underline{a}$ irányával megegyező, ha $\lambda > 0$, és azzal ellentétes, ha $\lambda < 0$. $\lambda = 0$ esetén $\lambda \cdot \underline{a} = \underline{0}$, továbbá a nullvektort bármely számmal szorozva a nullvektort kapjuk. Vegyük észre, hogy mindkét művelet felfogható egy-egy leképezésként. Az összeadásnál két vektor párjához egy harmadik vektort rendelünk: az összegüket. Ez tehát egy $\mathcal{E}_n \times \mathcal{E}_n \to \mathcal{E}_n$ ($n = 2$ vagy 3) típusú leképezés. A skalárral való szorzásnál egy vektor és egy valós szám párjához rendelünk egy vektort, a skalárral való szorzás tehát nem más, mint egy $\mathcal{E}_n \times \mathbb{R} \to \mathcal{E}_n$ típusú leképezés.

A két műveletre igaz a következő hét tulajdonság:

1. $\underline{a} + \underline{b} = \underline{b} + \underline{a} \qquad \forall \underline{a}, \underline{b} \in \mathcal{E}_n (n = 2, 3)$ (az összeadás kommutatív)
2. $(\underline{a} + \underline{b}) + \underline{c} = \underline{a} + (\underline{b} + \underline{c}) \qquad \forall \underline{a}, \underline{b}, \underline{c} \in \mathcal{E}_n (n = 2, 3)$ (az összeadás asszociatív)
3. A nullvektor egy olyan vektor, amely azzal a tulajdonsággal rendelkezik, hogy bármely $\underline{a} \in \mathcal{E}_n (n = 2, 3)$ vektorhoz adva $\underline{a}$-t kapjuk, azaz $\underline{a} + \underline{0} = \underline{a}$.

4. Minden $\underline{a} \in \mathcal{E}_n(n = 2,3)$ vektorhoz létezik olyan vektor, amelyet $\underline{a}$-hoz adva a nullvektort kapjuk: ez az $\underline{a}$-val ellentétes irányú, vele megegyező hosszúságú vektor. Ezt a vektort jelölje $-\underline{a}$, és elnevezése: ellentett vektor.

5. Vektorok összegét tagonként szorozhatjuk:
$\lambda \cdot (\underline{a} + \underline{b}) = \lambda \cdot \underline{a} + \lambda \cdot \underline{b} \qquad \forall \underline{a}, \underline{b} \in \mathcal{E}_n(n = 2,3)$ és $\forall \lambda \in \mathbb{R}$ (vektordisztributivitás)

6. Két valós szám összegével bármely vektort tagonként megszorozhatunk :
$(\lambda + \mu) \cdot \underline{a} = \lambda \cdot \underline{a} + \mu \cdot \underline{a} \qquad \forall \underline{a} \in \mathcal{E}_n(n = 2,3)$ és $\forall \lambda, \mu \in \mathbb{R}$ (skalárdisztributivitás)

7. Két valós számmal tetszőleges sorrendben szorozhatunk:
$\lambda \cdot (\mu \cdot \underline{a}) = \mu \cdot (\lambda \cdot \underline{a}) = (\lambda\mu) \cdot \underline{a} \qquad \forall \underline{a} \in \mathcal{E}_n(n = 2,3)$ és $\forall \lambda, \mu \in \mathbb{R}$ (skalárasszociativitás)

Megjegyzés: A fentiek segítségével értelmezhető a kivonás: $\underline{a} - \underline{b} := \underline{a} + (-\underline{b})$ és a skalárral való osztás (ha $\lambda \neq 0$): $\frac{\underline{a}}{\lambda} := \frac{1}{\lambda} \cdot \underline{a}$

Feladat. Legyen $\underline{a} \neq \underline{0}$. Milyen hosszú az $\frac{\underline{a}}{|\underline{a}|}$ vektor?
Megoldás: $\left|\frac{\underline{a}}{|\underline{a}|}\right| = |\frac{1}{|\underline{a}|} \cdot \underline{a}| = |\frac{1}{|\underline{a}|}| \cdot |\underline{a}| = \frac{1}{|\underline{a}|} \cdot |\underline{a}| = 1$

1.3. Lineáris kombináció, lineáris összefüggőség

1.2 Definíció. *Az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_p$ geometriai vektorok* **lineáris kombinációi***ján valamely α_i, $i = 1, 2, \ldots p$ valós számok mellett az $\alpha_1 \cdot \underline{a}_1 + \alpha_2 \cdot \underline{a}_2 + \ldots + \alpha_p \cdot \underline{a}_p$ vektort értjük.*

Megjegyzés: egy darab $\underline{a}$ vektor lineáris kombinációi az $\alpha \cdot \underline{a}$ szorzatok.

Pl. egy $\underline{a}$ és egy $\underline{b}$ vektornak lineáris kombinációi: $2\underline{a} + 0\underline{b}$, $\frac{3}{4}\underline{a} + 10\underline{b}$, $\pi\underline{a} - 2\underline{b}$ stb.
Hány lineáris kombinációt lehet gyártani? Végtelen sokat.
Mi lesz a lineáris kombináció, ha mindkét vektort 0-val szorozzuk? $0\underline{a} + 0\underline{b} = \underline{0}$ (ún. triviális lineáris kombináció)

Fontos kérdés: Adott $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_p$ vektorokból hogyan kaphatjuk meg lineáris kombinációval a nullvektort? Nyilvánvalóan, ha mindegyik $\alpha_i = 0$, akkor megkapjuk a nullvektort. De vajon van más lehetőség is?

1.3 Definíció. *Az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_p$ vektorokat* **lineárisan függetlene***knek nevezzük, ha lineáris kombinációjuk csak akkor lehet egyenlő a nullvektorral, ha minden vektort nullával szorzunk. (Röviden: $\alpha_1\underline{a}_1 + \alpha_2\underline{a}_2 + \ldots + \alpha_p\underline{a}_p = \underline{0} \Leftrightarrow \alpha_i = 0 \ \forall i = 1, 2, \ldots, p$.)*

Ennek ellentéte a lineáris összefüggőség:

1.4 Definíció. *Az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_p$ vektorokat* **lineárisan összefüggő***knek nevezzük, ha lineáris kombinációjuk úgy is egyenlő lehet a nullvektorral, ha valamelyik vektort nemnulla számmal szorozzuk.*

A továbbiakban sokat foglalkozunk vektorokból álló halmazokkal (ún. vektorhalmaz vagy vektorrendszer). Értelemszerűen, ha azt a kifejezést használjuk, hogy az $\{\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_p\}$ vektorhalmaz lineárisan összefüggő/független, akkor azon azt kell érteni, hogy az $\underline{a}_1$, $\underline{a}_2, \ldots, \underline{a}_p$ vektorok lineárisan összefüggőek/függetlenek.

Feladat. Lineárisan összefüggő-e az $\{\underline{a}, \underline{0}\}$ vektorhalmaz, ahol $\underline{a} \in \mathcal{E}_2, \underline{a} \neq \underline{0}$ tetszőleges?
Megoldás: Milyen együtthatókkal állhat fenn: $\alpha_1\underline{a} + \alpha_2\underline{0} = \underline{0}$? Ha $\alpha_1 = 0$, akkor α_2 tetszőlegesen megválasztható $\Rightarrow$ a vektorhalmaz lineárisan összefüggő.
(Hasonlóan meggondolható, hogy minden olyan vektorrendszer lineárisan összefüggő, amelyben benne van a nullvektor.)

Annak eldöntéséhez, hogy egy vektorhalmaz lineárisan összefüggő vagy független-e, sok esetben fel tudjuk használni a következő tételt.

1.5 Állítás. *Az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_p$ vektorok pontosan akkor lineárisan összefüggőek, ha* **legalább egyikük** *előállítható a többi vektor lineáris kombinációjaként.*

Biz.: A bizonyítás két részből áll.
1. ($\Rightarrow$) Tegyük fel, hogy $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_p$ lineárisan összefüggőek. Ekkor léteznek olyan $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{R}$ számok, amelyekkel

$$\alpha_1\underline{a}_1 + \alpha_2\underline{a}_2 + \ldots + \alpha_p\underline{a}_p = \underline{0},$$

és legalább az egyik együttható nem nulla, legyen α_i egy ilyen együttható. Ekkor ezen α_i-vel eloszthatjuk az egyenlőség mindkét oldalát. Átrendezve:

$$\underline{a}_i = -\frac{\alpha_1}{\alpha_i}\underline{a}_1 - \frac{\alpha_2}{\alpha_i}\underline{a}_2 - \ldots - \frac{\alpha_{i-1}}{\alpha_i}\underline{a}_{i-1} - \frac{\alpha_{i+1}}{\alpha_i}\underline{a}_{i+1} - \ldots - \frac{\alpha_p}{\alpha_i}\underline{a}_p,$$

vagyis $\underline{a}_i$ kifejezhető a többi vektor lineáris kombinációjaként.
2. ($\Leftarrow$) Tegyük fel, hogy valamely $\underline{a}_i$ vektor kifejezhető a többi vektor lineáris kombinációjaként, azaz

$$\underline{a}_i = \alpha_1\underline{a}_1 + \alpha_2\underline{a}_2 + \ldots + \alpha_{i-1}\underline{a}_{i-1} + \alpha_{i+1}\underline{a}_{i+1} + \ldots + \alpha_p\underline{a}_p.$$

Ekkor

$$1 \cdot \underline{a}_i - \alpha_1 \underline{a}_1 - \alpha_2 \underline{a}_2 - \ldots - \alpha_{i-1} \underline{a}_{i-1} - \alpha_{i+1} \underline{a}_{i+1} - \ldots - \alpha_p \underline{a}_p = \underline{0}.$$

A bal oldalon egy nemtriviális lineáris kombináció áll (a legelső együttható 1), így a vektorok lineárisan összefüggőek.

1.6 Következmény. *Az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_p$ vektorok pontosan akkor lineárisan függetlenek, ha egyikük sem állítható elő a többi vektor lineáris kombinációjaként.*

Mindezek alapján beláthatók a következő állítások.

1. Két nemnulla vektor pontosan akkor lineárisan összefüggő, ha kollineárisak (azaz egy egyenesbe esők).
2. Három nemnulla vektor pontosan akkor lineárisan összefüggő, ha komplanárisak (azaz egy síkba esők).
3. Négy vagy több geometriai vektor már mindig lineárisan összefüggő.

Fontos következmények:

- 2. $\Rightarrow$ A síkon három vektor mindig lineárisan összefüggő (hiszen egy síkban vannak).
- 1. $\Rightarrow$ A síkon az $\underline{i}, \underline{j}$ descartes-i egységvektorok lineárisan függetlenek.
- 2. $\Rightarrow$ A térben az $\underline{i}, \underline{j}, \underline{k}$ descartes-i egységvektorok lineárisan függetlenek (páronként merőlegesek egymásra, így nem esnek egy síkba).

1.4. Geometriai vektorterek

Vegyük észre, hogy $\mathcal{E}_2$-re igaz a következő két tulajdonság:
1. Láttuk, hogy a síkvektorok összeadása $\mathcal{E}_2 \times \mathcal{E}_2 \to \mathcal{E}_2$ típusú művelet. Következésképpen, ha $\underline{a} \in \mathcal{E}_2$ és $\underline{b} \in \mathcal{E}_2$, akkor $\underline{a} + \underline{b} \in \mathcal{E}_2$, azaz bármely két síkvektor összege is síkvektor, tehát szintén az $\mathcal{E}_2$ halmazban van. Más szóval az $\mathcal{E}_2$ halmaz **zárt az összeadásra**.
(Egy példa arra, amikor egy halmaz nem zárt valamilyen műveletre: a pozitív számok halmaza nem zárt a kivonásra, hiszen pl. $2 - 10 = -8$, ami már nem pozitív szám.)
2. A síkvektorok skalárral való szorzása $\mathcal{E}_2 \times \mathbb{R} \to \mathcal{E}_2$ típusú művelet. Így, ha $\underline{a} \in \mathcal{E}_2$, akkor minden λ valós számra $\lambda \cdot \underline{a} \in \mathcal{E}_2$, azaz bármely síkvektor tetszőleges számszorosa is síkvektor, tehát szintén az $\mathcal{E}_2$ halmazban van. Az $\mathcal{E}_2$ halmaz tehát **zárt a skalárral való szorzásra**.
Ugyanezek igazak $\mathcal{E}_3$-ra is, azaz bármely két térvektor összege és bármely térvektor tetszőleges számszorosa is magában a halmazban, azaz $\mathcal{E}_3$-ban van.

1.7 Definíció. *(Geometriai vektortér)*
Ha geometriai vektorok egy V halmazára igaz, hogy
1. minden $\underline{a}, \underline{b} \in V$ esetén $\underline{a} + \underline{b} \in V$, és
2. minden $\underline{a} \in V$ és $\lambda \in \mathbb{R}$ esetén $\lambda \underline{a} \in V$,
akkor a V halmazt **geometriai vektortér***nek nevezzük.*

Geometriai vektorterek-e a következő halmazok?

a.) $H_1 := \{\underline{i},\ \underline{j}\} \subset \mathcal{E}_2$
Nem, ugyanis pl. $\underline{i} + \underline{j}$ nincs benne H_1-ben.

b.) $H_2 := \{\underline{v} \in \mathcal{E}_2 : \underline{v} \parallel \underline{u},\ \underline{u} \in \mathcal{E}_2$ rögzített$\}$, azaz azon síkvektorok halmaza, amelyek kollineárisak egy rögzített $\underline{u}$ síkvektorral.
Legyen $\underline{v}_1$ és $\underline{v}_2$ két tetszőleges vektor a H_2-ben. Ekkor $\underline{v}_1 \parallel \underline{u}$ és $\underline{v}_2 \parallel \underline{u}$. Egy egyenesbe eső vektorok összege is ugyanabban az egyenesben van. Így

$$\underline{v}_1 + \underline{v}_2 \parallel \underline{u}, \text{ azaz } \underline{v}_1 + \underline{v}_2 \in H_2.$$

Legyen továbbá $\lambda \in \mathbb{R}$ tetszőleges. Ekkor $\lambda \cdot \underline{v} \parallel \underline{v}$ és $\underline{v} \parallel \underline{u}$ (hiszen $\underline{v} \in H_2$), következésképpen $\lambda \cdot \underline{v} \parallel \underline{u}$, azaz $\lambda \cdot \underline{v} \in H_2$. Tehát a H_2 halmazra már igaz 1. és 2., tehát geometriai vektortér.

1.8 Megjegyzés. *Ha egy V geometriai vektortér valamely $\tilde{V}$ részhalmazára teljesül, hogy maga is geometriai vektortér, akkor $\tilde{V}$-t a V alterének nevezzük. Így a fenti H_2 geometriai vektortér altere az $\mathcal{E}_2$-nek.*

1.9 Definíció. *Azt mondjuk, hogy egy geometriai vektortér dimenziója k, ha található benne k darab vektorból álló lineárisan független rendszer, $k+1$ darabból álló azonban már nem.*

Pl. $\dim\mathcal{E}_2 = 2$, ugyanis a síkon két nem egy egyenesbe eső vektor lineárisan független rendszert alkot, három egy síkba eső vektor azonban már lineárisan összefüggő. Hasonlóan belátható: $\dim H_2 = 1, \dim\mathcal{E}_3 = 3$. Tehát $\dim\mathcal{E}_n = n (n = 2, 3)$.

Kérdés: Tudunk-e egy vektortérben olyan elemeket találni, amelyek lineáris kombinációi kiadják a teljes vektorteret? Az ilyen elemek halmazának nevet is adunk:

1.10 Definíció. *Az $U \subset V$ halmazt a V vektortér* **generálórendszer***ének nevezzük, ha V minden eleme előállítható U elemeinek lineáris kombinációjával.*

Feladatok. A megadott V vektortereknek generálórendszerei-e az alábbi U halmazok?
1.) $V = H_2$
a.) $U = \{\underline{a}\}$, ahol $\underline{a} \in H_2$, $\underline{a} \neq \underline{0}$ tetsz. igen
b.) $U = \{\underline{a}, \underline{b}\}$, ahol $\underline{a}, \underline{b} \in H_2$ tetsz. igen
2.) $V = \mathcal{E}_2$
a.) $U = \{\underline{a}\}$, ahol $\underline{a} \in \mathcal{E}_2$, $\underline{a} \neq \underline{0}$ tetsz. nem
b.) $U = \{\underline{a}, \underline{b}\}$, ahol $\underline{a}, \underline{b} \in \mathcal{E}_2$, $\underline{a} \nparallel \underline{b}$ igen
c.) $U = \{\underline{a}, \underline{b}\}$, ahol $\underline{a}, \underline{b} \in \mathcal{E}_2$, $\underline{a} \parallel \underline{b}$ nem
d.) $U = \{\underline{a}, \underline{b}, \underline{c}\}$, ahol $\underline{a}, \underline{b}, \underline{c} \in \mathcal{E}_2$, és $\underline{a}, \underline{b}, \underline{c}$ páronként hegyesszöget zár be igen

Látszik: sokféle generálórendszer megadható. De ezek különbözhetnek abban, hogy belőlük hogyan állíthatók elő a vektortér elemei! Hasonlítsuk össze az 1./a.) és b.) példát. Hogyan állítható elő pl. a $\underline{c} = 2\underline{a}$ vektor az a.) és a b.) generálórendszer segítségével?
a.) $\underline{c} = 2\underline{a}$ az egyetlen előállítás
b.) Legyen $\underline{b} = -\underline{a}$. Ekkor végtelen sok lehetőség van:

$$\underline{c} = 2\underline{a} + 0\underline{b} = 3\underline{a} + 1\underline{b} = 4\underline{a} + 2\underline{b} \text{ stb.}$$

Előnyös, ha az előállítás egyértelmű. Vajon a generálórendszer melyik tulajdonsága biztosítja, hogy a vektorokat egyértelműen lehessen a generálórendszerből előállítani? Mint látni fogjuk, ez a lineáris függetlenség. Ezért bevezetjük a következő definíciót.

1.11 Definíció. *A lineárisan független generálórendszert bázisnak nevezzük.*

Egy bázisnak nem lehet akárhány eleme. Erről szól a következő állítás.

1.12 Állítás. *Ha V geometriai vektortér, és $B \subset V$ bázis V-ben, akkor B elemszáma* $\mathrm{dim}V$*-vel egyenlő.*

Biz.: Legyen V dimenziója n. Belátjuk, hogy B-nek nem lehet sem n-nél kevesebb, sem n-nél több eleme.
1. Tegyük fel, hogy B elemszáma k, és $k < n$. Jelölje az elemeit $\underline{b}_1, \ldots \underline{b}_k$. Mivel $\mathrm{dim}V = n$, így V-ben található n darab lineárisan független vektor, jelölje ezeket $\underline{a}_1, \ldots \underline{a}_n$. B bázis, következésképpen az $\underline{a}_1$ vektor előállítható B elemeinek lineáris kombinációjaként, azaz

$$\underline{a}_1 = \alpha_1 \underline{b}_1 + \ldots + \alpha_k \underline{b}_k.$$

A jobb oldalon biztos, hogy valamelyik együttható nem nulla, különben a jobb oldalon a nullvektor állna, a bal oldali vektor viszont nem nullvektor (hiszen lineárisan független vektorhalmaz eleme). Legyen ez a nemnulla együttható pl. az α_1. Ekkor α_1-gyel oszthatunk, és így $\underline{b}_1$ kifejezhető $\underline{a}_1, \underline{b}_2, \ldots, \underline{b}_k$ lineáris kombinációjaként. $\Rightarrow \{\underline{a}_1, \underline{b}_2, \ldots, \underline{b}_k\}$ generálórendszere V-nek. Akkor viszont pl. $\underline{a}_2$ kifejezhető $\underline{a}_1, \underline{b}_2, \ldots, \underline{b}_k$ lineáris kombinációjaként:

$$\underline{a}_2 = \beta_1 \underline{a}_1 + \beta_2 \underline{b}_2 + \ldots + \beta_k \underline{b}_k.$$

A jobb oldalon valamelyik $\underline{b}_i$ együtthatója biztos nem nulla, különben $\underline{a}_2$ és $\underline{a}_1$ lineárisan összefüggő volna. Legyen ez pl. a $\underline{b}_2$. $\Rightarrow \{\underline{a}_1, \underline{a}_2, \underline{b}_3 \ldots, \underline{b}_k\}$ generálórendszer. Folytatva az eljárást, arra jutunk, hogy $\{\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k\}$ generálórendszer V-ben. Ez azonban nem lehet, mert pl. az $\underline{a}_n$ vektor nem állítható elő ezen vektorok lineáris kombinációjaként, hiszen lineárisan független tőlük.
2. Tegyük fel, hogy B-nek $k > n$ darab eleme van.
$\dim V = n$ miatt V-ben nem található $n+1$ darab lineárisan független vektor, tehát $k > n$ darab vektor nem lehet lineárisan független, így B nem lehet bázis. Ezzel az állítást beláttuk.

Az állítás alapján $\mathcal{E}_2$-ben két vektorból, $\mathcal{E}_3$-ban három vektorból, és a H_2 vektortérben egy vektorból állnak a bázisok.

Igaz a következő állítás is.

1.13 Állítás. *Ha V geometriai vektortér, $B \subset V$ lineárisan független és B elemszáma $\dim V$-vel egyenlő, akkor B bázis V-ben.*

Biz.: A feltevés szerint B lineárisan független, tehát már csak azt kell megmutatni, hogy generálórendszer is. Ha V dimenziója $\dim V$, akkor V-ben $\dim V + 1$ darab lineárisan független vektor már nem található, vagyis B-hez tetszőleges $\underline{v} \in V$ elemet hozzávéve lineárisan összefüggő vektorrendszert kapunk. A hozzávett elem előállítható B elemeiből. (Ennek bizonyítása hasonló az 1. gyakorló példasor 8. példájának bizonyításához.) Tehát V tetszőleges eleme előállítható a B elemeiből, vagyis B generálórendszere V-nek.

Ebből következik, hogy a síkon bármely két lineárisan független vektor, a térben bármely három lineárisan független vektor, a H_2 vektortérben pedig bármely nemnulla vektor bázist alkot. Speciálisan, a síkvektorok között bázist alkotnak az $\underline{i}$ és $\underline{j}$, a térvektorok között az $\underline{i}, \underline{j}$ és $\underline{k}$ descartes-i egységvektorok (ún. descartes-i bázis a síkon/térben).

1.14 Állítás. *Ha B bázis a V geometriai vektortérben, akkor B elemeiből minden $\underline{v} \in V$ vektor egyértelműen (azaz csak egyféleképpen) állítható elő lineáris kombinációval.*

Bizonyítás helyett egy példán szemléltetjük, hogy mindez miért éppen a lineáris függetlenségen múlik.
Legyen V az $\mathcal{E}_2$ vektortér. Tekintsük azt az $\underline{a}$ vektort, amelynek az x és y-tengelyre eső vetülete 3 ill. 2 hosszúságú, és az első síknegyedben van. Az előzőek értelmében az $\mathcal{E}_2$-ben az $\{\underline{i}, \underline{j}\}$ halmaz bázis, hiszen két lineárisan független vektor a síkon bázist alkot. (Ez a leggyakrabban használt, ún. descartes-i bázis az $\mathcal{E}_2$-ben.) Világos, hogy $\underline{a}$ előáll $3\underline{i} + 2\underline{j}$ alakban. Gondoljuk meg, hogy miért nem lehetnek itt mások az együtthatók! Tegyük fel, hogy az $\underline{a}$ az $\alpha_1\underline{i} + \alpha_2\underline{j}$ alakban is felírható, ahol $\alpha_1 \neq 3$ és $\alpha_2 \neq 2$. Ekkor egyszerre fennállna a következő két egyenlőség:

$$\underline{a} = 3\underline{i} + 2\underline{j},$$
$$\underline{a} = \alpha_1\underline{i} + \alpha_2\underline{j}.$$

Vonjuk ki az első egyenletből a másodikat:

$$\underline{0} = (3 - \alpha_1)\underline{i} + (2 - \alpha_2)\underline{j}.$$

A jobb oldalon az $\underline{i}$ és a $\underline{j}$ vektorok lineáris kombinációja van. Mivel $\underline{i}$ és $\underline{j}$ lineárisan független rendszert alkot, ezért lineáris kombinációjuk csak akkor adhatja a nullvektort, ha mindkét vektor együtthatója nulla, azaz

$$3 - \alpha_1 = 0 \quad \text{és} \quad 2 - \alpha_2 = 0.$$

Ez viszont azt jelenti, hogy $\alpha_1 = 3$ és $\alpha_2 = 2$, azaz ellentmondásba kerültünk azzal a feltevéssel, hogy $\alpha_1 \neq 3$ és $\alpha_2 \neq 2$. Így tehát nincs más előállítás. A fentiekből látható, hogy ez a tény a báziselemek lineáris függetlenségének a következménye.

Miért jó az egyértelmű előállítás? Mivel minden egyes $\underline{v} \in \mathcal{E}_2$ vektort csak egyféleképpen lehet az $\alpha_1\underline{b}_1 + \alpha_2\underline{b}_2$ alakban felírni, ezért a $\underline{v}$ vektort egyértelműen azonosítja az α_1 és α_2 szám. Mivel nem mindegy, hogy ezek közül melyik a $\underline{b}_1$, és melyik a $\underline{b}_2$ együtthatója, ezért a két számot rendezett számpárként ($\mathbb{R}^2$-beli elem) adjuk meg: mindig előre írjuk az első báziselem ($\underline{b}_1$) és hátra a második báziselem ($\underline{b}_2$) együtthatóját. Így pl. az $\underline{a} = 3\underline{i} + 2\underline{j}$ vektort egyértelműen megadhatjuk a $(3, 2)$ számpárral. Az α_1 és α_2 számot a vektornak az adott bázisra vonatkozó **koordinátá**inak, az $\alpha_1\underline{b}_1$ és $\alpha_2\underline{b}_2$ vektorokat pedig a vektor **komponens**einek (összetevőinek) nevezzük. A descartes-i bázisban a síkvektor koordinátái az x-és y-tengelyre eső előjeles vetületek, ami igen egyszerűvé teszi ezen bázis használatát.
A koordináták és a komponensek természetesen függnek a bázistól. Tekintsük pl. az $\underline{a} = 3\underline{i} + 2\underline{j}$ vektort az $\{\underline{i}, \underline{b}\}$ bázisban, ahol legyen most $\underline{b} = -\underline{j}$. Ekkor $\underline{a} = 3\underline{i} - 2\underline{b}$, vagyis

az $\underline{a}$ vektorhoz ebben a bázisban a $(3, -2)$ számpárt rendelhetjük hozzá.

A térvektorokat ugyanilyen módon számhármasokkal ($\mathbb{R}^3$-beli elemekkel) azonosíthatjuk.

1.5. A polárkoordináta-rendszer

Síkvektorokat gyakran kényelmes ún. polárkoordinátákkal megadni: az r hosszúsággal és a ϕ irányszöggel ($\phi \in [0, 2\pi)$, ahol az intervallum balról zárt, jobbról nyílt!):

$$\underline{a} \mapsto (r, \phi)_p$$

(Vigyázat: itt nem bázisvektorokhoz tartozó együtthatókról van szó!)

Feladat: a.) Adjuk meg az $\underline{a} = \underline{i} + \underline{j}$ vektor polárkoordinátás alakját!

$$r = \sqrt{1+1} = \sqrt{2}, \quad \phi = \frac{\pi}{4}\ (45^o) \Rightarrow\ \underline{a} \text{ polárkoordinátás alakja: } (\sqrt{2}, \frac{\pi}{4})_p$$

b.) Ha egy $\underline{a}$ vektor polárkoordinátás alakja $(4, \frac{\pi}{3})_p$, akkor mik a Descartes-koordinátái?

$$\begin{array}{l} x = r\cos\phi = 4\cos\frac{\pi}{3} = 2 \\ y = r\sin\phi = 4\sin\frac{\pi}{3} = 2\sqrt{3} \end{array} \quad \Rightarrow\ \underline{a} \text{ Descartes-koordinátás alakja: } (2, 2\sqrt{3})$$

2. Az $\mathbb{R}^n$ vektortér

Ebben a fejezetben kibővítjük a vektortér fogalmát.

Feladat: Adjuk meg síkbeli vektorok

a.) összegének,

b.) skalárszorosának

Descartes-koordinátáit!

a.) Legyen $\underline{a} = \alpha_1\underline{i} + \alpha_2\underline{j}$ és $\underline{b} = \beta_1\underline{i} + \beta_2\underline{j}$ két $\mathcal{E}_2$-beli vektor. Ekkor

$$\underline{a} + \underline{b} = (\alpha_1\underline{i} + \alpha_2\underline{j}) + (\beta_1\underline{i} + \beta_2\underline{j}) = (\alpha_1 + \beta_1)\underline{i} + (\alpha_2 + \beta_2)\underline{j},$$

amiből $\underline{a} + \underline{b}$ koordinátái: $(\alpha_1 + \beta_1, \alpha_2 + \beta_2)$. Azaz összeadásnál a megfelelő koordináták összeadódnak.

b.) Legyen $\underline{a} = \alpha_1\underline{i} + \alpha_2\underline{j}$ és $\lambda \in \mathbb{R}$ tetszőleges. Ekkor

$$\lambda \cdot \underline{a} = \lambda \cdot (\alpha_1\underline{i} + \alpha_2\underline{j}) = (\lambda \cdot \alpha_1)\underline{i} + (\lambda \cdot \alpha_2)\underline{j}.$$

Ebből a $\lambda \cdot \underline{a}$ vektor Descartes-koordinátái: $(\lambda \cdot \alpha_1, \lambda \cdot \alpha_2)$. Vagyis az $\underline{a}$ vektor mindkét Descartes-koordinátája λ-val szorzódik. Könnyen meggondolható, hogy ez nemcsak a Descartes-féle, hanem tetszőleges bázisvektorok esetén is igaz. (A levezetés ugyanez.)

Tekintsük most $\mathbb{R}^2$-t önmagában. Ez a rendezett valós számpárok halmaza. Vezessünk be ezen a halmazon két műveletet:

1. Összeadáson értsük a következőt: $(a, b) \oplus (c, d) := (a + c, b + d)$;

2. Skalárral való szorzáson pedig a következőt: $\lambda \odot (a, b) := (\lambda a, \lambda b)$.

Belátható, hogy ez a két művelet ugyanazzal a hét tulajdonsággal rendelkezik, mint a geometriai vektorok összeadása és skalárral való szorzása! Mutassuk meg pl., hogy az összeadás kommutatív:

$$(a, b) \oplus (c, d) = (c, d) \oplus (a, b)!$$

Ez igaz, mert $(a, b) \oplus (c, d) = (a + c, b + d) = (c + a, d + b) = (c, d) \oplus (a, b)$. (A második egyenlőség azért áll fenn, mert a valós számok összeadása kommutatív.)
A nullvektor tulajdonságával rendelkezik a $(0, 0)$ számpár: $(a, b) \oplus (0, 0) = (a, b)$
(A maradék öt tulajdonságot lássuk be önállóan.)
Igaz továbbá, hogy az összeadás és a skalárral való szorzás eredménye is mindig számpár, azaz benne van $\mathbb{R}^2$-ben. Így jogos $\mathbb{R}^2$-t (a fenti két művelettel ellátva!) szintén vektortérnek nevezni, elemeit, a valós számpárokat pedig vektoroknak. Vegyük észre, hogy most már $\mathbb{R}^2$-ben is van értelme a következő fogalmaknak: lineáris kombináció, lineáris összefüggőség/függetlenség, generálórendszer, bázis, dimenzió.

Feladat: Adjunk meg bázist $\mathbb{R}^2$-ben!
Például

$$B := \{(1, 0), (0, 1)\}$$

Ez egyrészt generálórendszere $\mathbb{R}^2$-nek, mivel bármely $(a, b) \in \mathbb{R}^2$ számpár felírható ezen két számpár lineáris kombinációjaként, mégpedig az a és b együtthatókkal:

$$(a, b) = a \odot (1, 0) \oplus b \odot (0, 1).$$

Belátandó még, hogy B lineárisan független rendszer: igaz-e, hogy

$$\alpha \odot (1, 0) \oplus \beta \odot (0, 1) = (0, 0) \Leftrightarrow \alpha, \beta = 0?$$

A bal oldalon elvégezve a műveleteket:

$$(\alpha, 0) \oplus (0, \beta) = (0, 0),$$

$$(\alpha, \beta) = (0, 0) \Leftrightarrow \alpha, \beta = 0.$$

(A továbbiakban a karikákat elhagyjuk a műveletek jelölésénél.)

Az $\mathbb{R}^2$ vektortér szoros kapcsolatban van a síkvektorokkal: láttuk, hogy adott bázist választva kölcsönösen egyértelműen megfeleltethetők egymásnak a két halmaz elemei. Ez a bijekció (kölcsönösen egyértelmű megfeleltetés) ráadásul művelettartó is, ami a következőt jelenti. Ha az $\underline{a}$ és $\underline{b}$ vektoroknak az (a_1, a_2) és a (b_1, b_2) számpár felel meg, akkor az $\underline{a} + \underline{b}$ összegnek éppen az a számpár felel meg, amelyet úgy kapunk, hogy az $\mathbb{R}^2$-n definiált összeadási szabállyal összeadjuk az (a_1, a_2) és a (b_1, b_2) számpárt. Hasonlóan, a $\lambda \cdot \underline{a}$ vektornak az a számpár felel meg, amelyet úgy kapunk, hogy az $\mathbb{R}^2$-n definiált szorzási szabályt követve megszorozzuk az (a_1, a_2) számpárt λ-val. Ez a művelettartó bijekció biztosítja, hogy a síkvektorok helyett számolhatunk a nekik megfeleltetett számpárokkal. Egy számpár és a megfelelő síkvektor közé egyenlőségjelet teszünk.

Feladat: Legyen $\underline{a} = \underline{i} + 2\underline{j}$ és $\underline{b} = -2\underline{i} - \underline{j}$. Adjuk meg a $3(\underline{a} + \underline{b})$ vektor Descartes-koordináta-rendszerbeli alakját.

1. $3(\underline{a} + \underline{b}) = 3(\underline{i} + 2\underline{j} - 2\underline{i} - \underline{j}) = 3(-\underline{i} + \underline{j}) = -3\underline{i} + 3\underline{j} = (-3, 3)$
2. $3(\underline{a} + \underline{b}) = 3 \odot ((1, 2) \oplus (-2, -1)) = 3 \odot (-1, 1) = (-3, 3).$

Hasonlóan bevezethetők az $\mathbb{R}^3, \mathbb{R}^4, \ldots \mathbb{R}^n$ vektorterek. (A térvektorok vektortere $\mathbb{R}^3$-mal azonosítható, a többit azonban nem tudjuk geometriai vektortérrel azonosítani.) Miközben a geometriai vektorterek között nem tudtunk 3-nál többdimenziósat találni, az $\mathbb{R}^n$ vektortér n-dimenziós, így most már tudunk mondani tetszőleges dimenziójú vektorteret.

2.1 Megjegyzés. *A vektortér egészen általános definíciójával a következő félévben fogunk részletesen foglalkozni. Itt csak megemlítjük, hogy általánosan vektortérnek nevezünk minden olyan halmazt, amelyen értelmezve van ugyanolyan tulajdonságú két művelet, mint a geometriai vektorok összeadása és skalárral való szorzása. Az egzakt definíció a következőképpen szól.*

2.2 Definíció. *Legyen X egy tetszőleges halmaz, amelyen értelmezve van egy $\oplus : X \times X \to X$ és egy $\odot : X \times \mathbb{R} \to X$ művelet a következő tulajdonságokkal:*

1. $x \oplus y = y \oplus x \quad \forall x, y \in X$

2. $(x \oplus y) \oplus z = x \oplus (y \oplus z) \quad \forall x, y, z \in X$

3. Létezik nullelem (nullvektor), azaz olyan $0_X \in X$ *elem, amelyre* $x \oplus 0_X = x \quad \forall x \in X$

4. Minden $x \in X$ *elemhez létezik negatív elem, azaz olyan elem (jelölje -x), amelyre* $x \oplus (-x) = 0_X$*, ahol* 0_X *a 3. tulajdonság szerini nullvektor.*

5. $\lambda \odot (x \oplus y) = (\lambda \odot x) \oplus (\lambda \odot y) \quad \forall x, y \in X$ *és* $\forall \lambda \in \mathbb{R}$

6. $(\lambda + \mu) \odot x = (\lambda \odot x) \oplus (\mu \odot x) \quad \forall x \in X$ *és* $\forall \lambda, \mu \in \mathbb{R}$

7. $\lambda \odot (\mu \odot x) = (\lambda \cdot \mu) \odot x = \mu \odot (\lambda \odot x) \quad \forall x \in X$ *és* $\forall \lambda, \mu \in \mathbb{R}$

Ekkor az $(X, \oplus, \odot)$ *rendezett hármast vektortérnek, az* X *elemeit pedig vektoroknak nevezzük.*

Az eddigiekben csak azon speciális vektorterekkel foglalkoztunk, amelyekben az X *halmaz valamilyen geometriai vektortér ill.* $\mathbb{R}^n$ *volt. Léteznek azonban egyéb vektorterek is, pl. olyanok, ahol az* X *halmaz függvényekből áll. Ilyen vektorterekkel szintén a következő félévben találkozunk. Megjegyezzük, hogy a legtöbb vektortérben értelmezhető az elemek hosszúsága, az irány azonban kizárólag a geometriai vektorok sajátossága.*

3. Geometriai vektorok skaláris, vektoriális és vegyes szorzata

Térjünk vissza a geometriai vektorokra! Ezek körében további műveleteket értelmezünk, amelyek bevezetését a fizikai alkalmazások is indokolják.

3.1. Skaláris szorzat

3.1 Definíció. *Legyenek* $\underline{a}, \underline{b} \in \mathcal{E}_n$ *($n = 2$ vagy 3). Az* $\underline{a}$ *és* $\underline{b}$ *vektor* $\underline{a} \cdot \underline{b}$*-vel jelölt skaláris szorzatán az*

$$\underline{a} \cdot \underline{b} := |\underline{a}| \cdot |\underline{b}| \cdot \cos \gamma$$

valós számot értjük, ahol γ *az* $\underline{a}$ *és* $\underline{b}$ *vektorok közrezárt szöge.*

A közrezárt szögen azt a legkisebb abszolút értékű forgásszöget értjük, amellyel az egyik vektor a másikba beforgatható.

Vegyük észre, hogy a skaláris szorzásnál két vektorhoz egy valós számot rendelünk hozzá. A skaláris szorzás művelete tehát egy $\mathcal{E}_n \times \mathcal{E}_n \to \mathbb{R}$ ($n = 2$ vagy 3) típusú leképezés.

Pl. 1. $|\underline{a}| = \sqrt{3},\ |\underline{b}| = 2,\ \gamma = 30^o$
$\underline{a} \cdot \underline{b} = \sqrt{3} \cdot 2 \cdot \cos 30^o = 3$

A definíció alapján a skaláris szorzatból meghatározható a közbezárt szög: $\cos\gamma = \dfrac{\underline{a} \cdot \underline{b}}{|\underline{a}| \cdot |\underline{b}|}$

Három esetet különböztethetünk meg:

1. $\underline{a} \cdot \underline{b} > 0 \Rightarrow \gamma$ hegyesszög;

2. $\underline{a} \cdot \underline{b} = 0 \Rightarrow \gamma$ derékszög;

3. $\underline{a} \cdot \underline{b} < 0 \Rightarrow \gamma$ tompa- vagy egyenesszög.

Pl. 2. Mekkora szöget zár be az $\underline{a} = \underline{i} + \underline{k}$ és a $\underline{b} = \underline{j} + \underline{k}$ vektor?
$|\underline{a}| = \sqrt{2} = |\underline{b}|,\ \underline{a} \cdot \underline{b} = 1 \quad \Rightarrow \quad \cos\gamma = \dfrac{1}{\sqrt{2}\sqrt{2}} = \dfrac{1}{2} \Rightarrow \gamma = 60^o$
Könnyen látható, hogy mindegyik descartes-i egységvektor önmagával vett skaláris szoraza 1, két különböző descartes-i egységvektor skaláris szorzata pedig 0.

Műveleti tulajdonságok: bármely $\underline{a}, \underline{b}, \underline{c} \in \mathcal{E}_n$ (n=2 vagy 3) vektorra és $\alpha \in \mathbb{R}$ számra

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

2. $(\underline{a} + \underline{b}) \cdot \underline{c} = \underline{a} \cdot \underline{c} + \underline{b} \cdot \underline{c}$

3. $(\alpha\underline{a}) \cdot \underline{b} = \alpha(\underline{a} \cdot \underline{b}) = \underline{a} \cdot (\alpha\underline{b})$

Igaz-e, hogy $\underline{a}(\underline{b} \cdot \underline{c}) = (\underline{a} \cdot \underline{b})\underline{c}$? Általában nem, ugyanis az előbbi vektor $\underline{a}$-val egyirányú, a második pedig $\underline{c}$-vel! Ha tehát $\underline{a}$ és $\underline{c}$ nem esik egy egyenesbe, akkor nem állhat fenn az egyenlőség.

A műveleti tulajdonságokat kihasználva megmutatható, hogy tetszőleges két vektor skaláris szorzata Descartes-koordinátákkal: $\underline{a} \cdot \underline{b} = \sum_{i=1}^{n} a_i b_i$.

3.2. Vektoriális szorzat

Az $\underline{a}$ és $\underline{b} \in \mathcal{E}_3$ térvektorok $\underline{a} \times \underline{b}$-vel jelölt vektoriális szorzatán azt a térvektort értjük, amelyre $|\underline{a} \times \underline{b}| = |\underline{a}| \cdot |\underline{b}| \cdot \sin\gamma$, továbbá $\underline{a} \times \underline{b} \perp \underline{a}$, $\underline{a} \times \underline{b} \perp \underline{b}$ úgy, hogy $\underline{a}$, $\underline{b}$, $\underline{a} \times \underline{b}$ jobbsodrású rendszert alkot.

A vektoriális szorzat tehát olyan vektor, amelynek hossza a két vektor által kifeszített paralelogramma területe, iránya merőleges mindkét vektorra, irányítását pedig a jobbkézszabállyal találhatjuk ki.

A vektoriális szorzásnál két térvektorhoz egy térvektort rendelünk hozzá. A vektoriális szorzás művelete tehát egy $\mathcal{E}_3 \times \mathcal{E}_3 \to \mathcal{E}_3$ típusú leképezés.

Műveleti tulajdonságok: bármely $\underline{a}, \underline{b}, \underline{c} \in \mathcal{E}_3$ vektorra és $\alpha \in \mathbb{R}$ számra

1. $\underline{a} \times \underline{b} = -\underline{b} \times \underline{a}$
2. $(\underline{a} + \underline{b}) \times \underline{c} = \underline{a} \times \underline{c} + \underline{b} \times \underline{c}$
3. $(\alpha \underline{a}) \times \underline{b} = \alpha(\underline{a} \times \underline{b}) = \underline{a} \times (\alpha \underline{b})$
4. $(\underline{a} \times \underline{b}) \times \underline{c} = (\underline{c} \cdot \underline{a})\underline{b} - (\underline{b} \cdot \underline{c})\underline{a}$ (kifejtési tétel)

A vektoriális szorzat a Descartes-koordinátákkal:
Ha $\underline{a} = (a_1, a_2, a_3)$, $\underline{b} = (b_1, b_2, b_3)$, akkor

$$\underline{a} \times \underline{b} = (a_2 b_3 - a_3 b_2)\underline{i} + (a_3 b_1 - a_1 b_3)\underline{j} + (a_1 b_2 - a_2 b_1)\underline{k}$$

Kiszámításához érdemes elkészíteni az alábbi táblázatot:

$$\begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Nevezzük a bal felső sarokból kiinduló átlót főátlónak, a másikat mellékátlónak. A szabály a következő: vesszük a főátlóbeli elemek szorzatát, majd hozzáadjuk az első oszlop legalsó elemének, az első sor második elemének, valamint a harmadik oszlop második elemének a szorzatát, illetve a kapott eredményhez ismét hozzáadjuk az eddig kimaradt elemek szorzatát. Ebből az eredményből ezután kivonjuk a mellékátló elemeinek szorzatát, majd – az előbbi módon – az első sor első elemének, a második sor utolsó elemének, illetve a harmadik sor középső (második) elemének a szorzatát, legvégül pedig az eddig kimaradt elemek szorzatát.

Pl. $\underline{a} = (3, 2, 1)$, $\underline{b} = (4, 5, 6)$

$$\begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ 3 & 2 & 1 \\ 4 & 5 & 6 \end{vmatrix}$$

$\underline{a} \times \underline{b} = (12 - 5)\underline{i} + (4 - 18)\underline{j} + (15 - 8)\underline{k} = (7, -14, 7)$

3.3. Vegyes (triadikus) szorzat

3.2 Definíció. *Az* $\underline{a}, \underline{b}, \underline{c} \in \mathcal{E}_3$ *térvektorok vegyes (vagy triadikus) szorzatán* $(\underline{a}, \underline{b}, \underline{c}) := (\underline{a} \times \underline{b}) \cdot \underline{c}$ *skalárt értjük.*

Ez a művelet tehát nem más mint egy $\mathcal{E}_3 \times \mathcal{E}_3 \times \mathcal{E}_3 \to \mathbb{R}$ típusú leképezés. A vegyes szorzat abszolút értékben véve a három vektor által kifeszített paralelepipedon térfogatát adja meg.

Vegyük észre, hogy a vektorok sorrendje befolyásolja a vegyes szorzat értékét. A három vektor összesen hat különböző sorrendben írható fel. A sorrend megváltoztatásával a vegyes szorzatnak csak az előjele változhat meg. A különböző sorrendű vegyes szorzatok közötti összefüggést fogalmazza meg a következő, ún. felcserélési tétel.

3.3 Tétel. *Az* $\underline{a}, \underline{b}, \underline{c} \in \mathcal{E}_3$ *vektorokra* $(\underline{a}, \underline{b}, \underline{c}) = (\underline{c}, \underline{a}, \underline{b}) = (\underline{b}, \underline{c}, \underline{a}) = -(\underline{b}, \underline{a}, \underline{c}) = -(\underline{a}, \underline{c}, \underline{b}) = -(\underline{c}, \underline{b}, \underline{a})$.

Az $(\underline{a}, \underline{b}, \underline{c})$ vegyes szorzat kiszámítása Descartes-koordinátákkal egyszerűen elvégezhető az alábbi táblázat segítségével, a vektoriális szorzásnál tanult szabály alkalmazásával:

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

Pl. $\underline{a} = (1, 0, 1)$, $\underline{b} = (2, -1, 6)$, $\underline{c} = (0, 2, 5)$

$$\begin{vmatrix} 1 & 0 & 1 \\ 2 & -1 & 6 \\ 0 & 2 & 5 \end{vmatrix}$$

$(\underline{a}, \underline{b}, \underline{c}) = 1 \cdot (-1) \cdot 5 + 0 \cdot 6 \cdot 0 + 2 \cdot 2 \cdot 1 - 1 \cdot (-1) \cdot 0 - 1 \cdot 2 \cdot 6 - 0 \cdot 2 \cdot 5 = -13$

Feladatok.

1. Lássuk be, hogy a vektoriális szorzás 2. tulajdonsága a következő sorrendben is fennáll:
$$\underline{a} \times (\underline{b} + \underline{c}) = \underline{a} \times \underline{b} + \underline{a} \times \underline{c}$$
Megoldás: $\underline{a} \times (\underline{b} + \underline{c}) =^{1.\ tul.} -(\underline{b} + \underline{c}) \times \underline{a} =^{2.\ tul.} -(\underline{b} \times \underline{a} + \underline{c} \times \underline{a}) = -\underline{b} \times \underline{a} - \underline{c} \times \underline{a} =^{1.\ tul.} \underline{a} \times \underline{b} + \underline{a} \times \underline{c}$

2. Lássuk be a következő azonosságot: $(\underline{a} \times \underline{b}) \times (\underline{c} \times \underline{d}) = (\underline{c}, \underline{d}, \underline{a})\underline{b} - (\underline{c}, \underline{d}, \underline{b})\underline{a}$
Megoldás: Jelölje $\underline{e}$ a $\underline{c} \times \underline{d}$ vektort.

$(\underline{a} \times \underline{b}) \times (\underline{c} \times \underline{d}) = (\underline{a} \times \underline{b}) \times \underline{e} = (\underline{e} \cdot \underline{a})\underline{b} - (\underline{b} \cdot \underline{e})\underline{a} = ((\underline{c} \times \underline{d}) \cdot \underline{a})\underline{b} - (\underline{b} \cdot (\underline{c} \times \underline{d}))\underline{a} = (\underline{c}, \underline{d}, \underline{a})\underline{b} - (\underline{c}, \underline{d}, \underline{b})\underline{a}$

4. A komplex számtest

Ebben a fejezetben egy kis kitérőt teszünk.

Láttuk, hogy ha az $\mathbb{R}^2$ halmazt ellátjuk az összeadás és a skalárral való szorzás műveletével, akkor az $\mathbb{R}^2$ ugyanazokkal a tulajdonságokkal rendelkezik, mint a geometriai vektorterek. Ezért $\mathbb{R}^2$-t ezen két művelettel ellátva vektortérnek neveztük. A skalárral való szorzás műveletét úgy is nevezhetjük, hogy külső számtesttel való szorzás, mert a számpárokat nem számpárokkal, hanem az $\mathbb{R}$ halmaz (számtest) elemeivel szoroztuk.

Most induljunk ki az $\mathbb{R}^2$ halmazból, és definiáljunk rajta más módon műveleteket:

1. Összeadáson értsük ugyanazt, mint a vektortérben: $(a, b) + (c, d) := (a + c, b + d)$.
2. Két elem egymással való szorzatán (belső szorzás!) értsük a következőt: $(a, b) \cdot (c, d) := (ac - bd, ad + bc)$

Az összeadás tulajdonságai így megegyeznek a valós számok összeadásának tulajdonságaival. A belső szorzást azért definiáljuk ilyen bonyolult módon, mert csak így lehet elérni, hogy ez a művelet egyszerre rendelkezzen a valós számok szorzásának következő négy tulajdonságával:

1. kommutatív
2. asszociatív
3. van olyan számpár (ún. egységelem), amellyel bármely (a, b) számpárt szorozva (a, b)-t kapjuk: ez az $(1, 0)$;
4. minden $(a, b) \neq (0, 0)$ számpárhoz létezik olyan számpár (ún. reciprok elem), amellyel megszorozva az $(1, 0)$-t kapjuk.

Mivel tehát ezekkel a műveletekkel az $\mathbb{R}^2$ elemei ugyanolyan tulajdonságokkal rendelkeznek, mint a valós számtest, ezért $\mathbb{R}^2$-t ezen két művelettel ellátva komplex számtestnek nevezzük. A komplex számtestet $\mathbb{C}$-vel jelöljük, és $\mathbb{C}$ elemeit komplex számoknak nevezzük. Az $z = (a, b)$ komplex számban a-t a komplex szám valós részének (jel.: $\operatorname{Re} z$), b-t a komplex szám képzetes részének ($\operatorname{Im} z$) nevezzük. A valós és a komplex számtest között fontos különbség, hogy a valós számokkal ellentétben a komplex számokon nincsen rendezés, azaz nincs kisebb és nagyobb komplex szám.

A valós számok a komplex számok részének tekinthetők: az a valós számot azonosíthatjuk az $(a,0)$ komplex számmal. Ezt az azonosítást az indokolja, hogy a valós számok és a nulla képzetes részű komplex számok között művelettartó bijekció van. Ezért egy a valós szám és az $(a,0)$ komplex szám közé a továbbiakban egyenlőségjelet teszünk.

Egy (a,b) komplex szám $\overline{(a,b)}$ konjugáltján az $(a,-b)$ komplex számot értjük. Pl. $\overline{(2,-3)} = (2,3)$.

Az (a,b) komplex szám abszolút értékén az $|(a,b)| := \sqrt{a^2+b^2}$ számot értjük.

Feladatok:

1. Számítsuk ki az $(5,2)$ és a $(4,3)$ komplex számok összegét és szorzatát.
   $(5,2)+(4,3)=(9,5)$, $(5,2)\cdot(4,3)=(5\cdot 4-2\cdot 3, 5\cdot 3+2\cdot 4)=(14,23)$

2. Nevezetes komplex szám a $(0,1)$ (ún. képzetes egység), amelyet i-vel jelölünk. Számítsuk ki i négyzetét!

   $$i^2=(0,1)\cdot(0,1)=(0\cdot 0-1\cdot 1, 0\cdot 1-1\cdot 0)=(-1,0)=-1$$

   (Az utolsó lépésben azt használtuk fel, hogy a kapott komplex szám a korábbiak értelmében azonosítható a -1 valós számmal, így a kettő közé egyenlőségjelet írtunk.) Van tehát olyan komplex szám, amelynek a négyzete -1, miközben a valós számok között nem lehetett ilyet találni!

3. Mutassuk meg, hogy egy (a,b) komplex szám felírható $a+ib$ alakban (algebrai alak)!

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

4. Az algebrai alakkal kényelmes számolni. Pl. az $(5,2)$ és a $(4,3)$ szám szorzatát így is kiszámíthatjuk:

   $$(5+2i)\cdot(4+3i)=5\cdot 4+2i\cdot 4+5\cdot 3i+2i\cdot 3i=20+23i+6i^2=20+23i-6=14+23i$$

5. Adjuk meg a következő komplex számok algebrai alakját:
   a.) $z=i^{1994}$ b.) $z=|\overline{1+i}|$ c.) $z=\dfrac{3-4i}{2-i}$
   Megoldás:
   a.) $(i^2)^{997}=(-1)^{997}=-1$
   b.) $\sqrt{2}$

c.) $\dfrac{3-4i}{2-i} = \dfrac{3-4i}{2-i} \cdot \dfrac{2+i}{2+i} = \dfrac{(3-4i)(2+i)}{4-i^2} = \dfrac{(3-4i)(2+i)}{5} = \dfrac{6-8i+3i-4i^2}{5} = \dfrac{10-5i}{5} = 2-i$

Komplex számokat ugyanúgy szokásos az abszolút érték és az irányszög segítségével megadni, ahogy a síkvektorokat. Ha az (a, b) komplex szám abszolút értéke r, szöge ϕ, akkor $(a, b) = (r\cos\phi, r\sin\phi) = r\cos\phi + ir\sin\phi = r(\cos\phi + i\sin\phi)$ (trigonometrikus alak). A trigonometrikus alakkal igen kényelmesen elvégezhető a szorzás és az osztás:

$r(\cos\alpha + i\sin\alpha) \cdot p(\cos\beta + i\sin\beta) = rp(\cos(\alpha+\beta) + i\sin(\alpha+\beta))$

$\dfrac{r(\cos\alpha + i\sin\alpha)}{p(\cos\beta + i\sin\beta)} = \dfrac{r}{p}(\cos(\alpha-\beta) + i\sin(\alpha-\beta))$

Pl. legyen $z_1 = 2(\cos 45^o + i\sin 45^o)$, $z_2 = 3(\cos 5^o + i\sin 5^o)$. Mivel egyenlő $z_1 \cdot z_2$ és $\dfrac{z_1}{z_2}$?

$z_1 \cdot z_2 = 6(\cos 50^o + i\sin 50^o)$, $\dfrac{z_1}{z_2} = \dfrac{2}{3}(\cos 40^o + i\sin 40^o)$

A szorzásból következik az egész kitevős hatvány képlete:

$$z^n = [r(\cos\phi + i\sin\phi)]^n = r^n(\cos n\phi + i\sin n\phi)$$

Egy komplex szám n-edik hatványa egyértelmű, n-edik gyökből azonban n db van (ezek között lehetnek egyenlők is):

$$\sqrt[n]{z} = \sqrt[n]{r}\left(\cos\frac{\phi + k2\pi}{n} + i\sin\frac{\phi + k2\pi}{n}\right), \quad k = 0, 1, \dots n-1$$

Pl. $\sqrt{i} = ?$
Mivel $i = \cos 90^o + i\sin 90^o$, így

$$\sqrt{i} = \sqrt{1}\left(\cos\frac{90^o + k \cdot 360^o}{2} + i\sin\frac{90^o + k \cdot 360^o}{2}\right), \quad k = 0, 1,$$

azaz $\sqrt{i} = \pm\left(\dfrac{\sqrt{2}}{2} + i\dfrac{\sqrt{2}}{2}\right)$.

Másodfokú egyenleteknek mindig van két megoldásuk a komplex számok halmazán. A szokásos megoldóképletet használhatjuk. Míg negatív diszkrimináns esetén nincs valós megoldás, komplex megoldásból kettő is van, mert egy negatív számnak két komplex négyzetgyöke van. A -1 gyökei pl. a $+i$ és a $-i$.

Pl. Oldjuk meg a $z^2+3z+4=0$ egyenletet!

$$z_{1,2}=\frac{-3\pm\sqrt{9-16}}{2}=\frac{-3\pm\sqrt{-1}\sqrt{7}}{2}=-\frac{3}{2}\pm\frac{\sqrt{7}}{2}i.$$

5. Mátrixszámítás

A mátrix a lineáris algebra egyik alapfogalma. Mint látni fogjuk, a mátrixok szoros kapcsolatban vannak bizonyos fajta lineáris leképezésekkel, és hasznosak – többek között – a lineáris egyenletrendszerek tárgyalásánál. Ezért először definiáljuk a lineáris leképezés (lineáris függvény) fogalmát.

5.1. Lineáris leképezések

Tekintsük azt a függvényt, amely minden valós számhoz hozzárendeli az ötszörösét!

$$f:\mathbb{R}\to\mathbb{R},\quad f(x)=5x$$

Ellenőrizzük, hogy mit rendel hozzá ez a függvény két szám összegéhez?

$$x_1,x_2\in\mathbb{R},\quad f(x_1+x_2)=5(x_1+x_2)=5x_1+5x_2=f(x_1)+f(x_2)$$

Azaz két szám összegéhez a megfelelő függvényértékek összegét rendeli.
Mit rendel egy szám λ-szorosához?

$$x\in\mathbb{R},\lambda\in\mathbb{R},\quad f(\lambda\cdot x)=5\cdot(\lambda\cdot x)=\lambda\cdot 5x=\lambda\cdot f(x).$$

Azaz egy x szám λ-szorosához az x-beli függvényérték λ-szorosát rendeli.

5.1 Definíció. *(Valós lineáris függvény)*
Ha egy $f:\mathbb{R}\to\mathbb{R}$ függvényre igaz az alábbi két tulajdonság:

1. $f(x_1+x_2)=f(x_1)+f(x_2)\quad\forall x_1,x_2\in\mathbb{R}$, *és*

2. $f(\lambda x)=\lambda f(x)\quad\forall x\in\mathbb{R},\forall\lambda\in\mathbb{R}$,

akkor az f függvényt lineáris függvénynek nevezzük.

Lineárisak-e a következő valós függvények?
a.) $f(x)=0\ \forall x\in\mathbb{R}$
Legyen $x_1,x_2,x\lambda\in\mathbb{R}$ tetszőleges. Ekkor $f(x_1+x_2)=0=f(x_1)+f(x_2)$, és $f(\lambda x)=0=$

$\lambda \cdot f(x)$. Tehát f lineáris.

b.) $f(x) = 1 \ \forall x \in \mathbb{R}$

Legyen $x_1, x_2 \in \mathbb{R}$. Ekkor $f(x_1 + x_2) = 1$, azonban $f(x_1) + f(x_2) = 1 + 1 = 2 \neq 1$. Azaz az azonosan 1 függvény nem lineáris.

c.) $f(x) = x^2$

Ez sem lineáris, hiszen pl. $f(2x) = (2x)^2 = 4x^2 \neq 2 \cdot f(x) = 2x^2$.

d.) $f(x) = \sin x$ - szintén nem lineáris, hiszen pl. $\sin(2 \cdot 90^o) = 0 \neq 2\sin(90^o) = 2$

e.) $f(x) = 5x + 2$

$f(x_1 + x_2) = 5(x_1 + x_2) + 2 = 5x_1 + 5x_2 + 2$, de $f(x_1) + f(x_2) = 5x_1 + 2 + 5x_2 + 2 = 5x_1 + 5x_2 + 4 \neq 5x_1 + 5x_2 + 2$

Tehát ez sem lineáris! Ezt a függvényt helyesen lineáris inhomogén függvénynek nevezzük, utalva arra, hogy a valódi lineáris függvénytől csak annyiban különbözik, hogy egy nemnulla konstanst hozzáadtunk.

Belátható, hogy az $\mathbb{R} \to \mathbb{R}$ függvények körében csak az $f(x) = a \cdot x$ alakúak lineárisak, ahol $a \in \mathbb{R}$ rögzített szám. (Az világos az előzőekből, hogy az ilyen függvény lineáris, de az is könnyen meggondolható, hogy ha egy valós függvény lineáris, akkor csak ilyen alakú lehet. Ugyanis ha f lineáris, akkor $f(x) = f(1 \cdot x) = f(1) \cdot x = a \cdot x$, ahol $a = f(1)$.)

Tekintsük most az $f : \mathbb{R}^2 \to \mathbb{R}$ leképezéseket. Ezek a számpárokhoz valós számokat rendelnek hozzá:

$$f : (x_1, x_2) \mapsto y \in \mathbb{R}$$

Az ilyen függvényeket a következőképpen szemléltethetjük. Az értelmezési tartomány (x_1, x_2) pontjának megfeleltetjük az xy-sík egy pontját, majd az erre merőleges z tengelyen felmérjük az $f(x_1, x_2)$ függvényértéket.

A lineáris függvény fogalma az $f : \mathbb{R}^2 \to \mathbb{R}$ függvényekre is kiterjeszthető: f-et akkor nevezzük lineárisnak, ha egyszerre teljesül rá a következő két tulajdonság:

1. $f(\underline{x} + \underline{\tilde{x}}) = f(\underline{x}) + f(\underline{\tilde{x}}) \quad \forall \underline{x}, \underline{\tilde{x}} \in \mathbb{R}^2$,
2. $f(\lambda \underline{x}) = \lambda f(\underline{x}) \quad \forall \underline{x} \in \mathbb{R}^2, \forall \lambda \in \mathbb{R}$.

5.2 Tétel. *Egy $f : \mathbb{R}^2 \to \mathbb{R}$ függvény pontosan akkor lineáris, ha $f(x_1, x_2) = a_1x_1 + a_2x_2$ alakú, ahol $a_1, a_2 \in \mathbb{R}$ rögzített számok.*

Biz.:

($\Leftarrow$) Ha az f függvény $f(x_1, x_2) = a_1x_1 + a_2x_2$ alakú, akkor

1.

$$f((x_1, x_2) + (\tilde{x}_1, \tilde{x}_2)) = f(x_1 + \tilde{x}_1, x_2 + \tilde{x}_2) = a_1(x_1 + \tilde{x}_1) + a_2(x_2 + \tilde{x}_2)$$

$$= (a_1x_1 + a_2x_2) + (a_1\tilde{x}_1 + a_2\tilde{x}_2) = f(x_1, x_2) + f(\tilde{x}_1, \tilde{x}_2)$$

$\forall (x_1, x_2), (\tilde{x}_1, \tilde{x}_2) \in \mathbb{R}^2$ esetén,
2.

$$f(\lambda \cdot (x_1, x_2)) = f(\lambda x_1, \lambda x_2) = a_1 \lambda x_1 + a_2 \lambda x_2 = \lambda(a_1 x_1 + a_2 x_2) = \lambda f(x_1, x_2)$$

$\forall (x_1, x_2) \in \mathbb{R}^2, \forall \lambda \in \mathbb{R}$ esetén. Tehát f lineáris.
($\Rightarrow$) Belátjuk, hogy ha f lineáris, akkor csak a fenti alakú lehet. Tudjuk, hogy $\mathbb{R}^2$-ben bázist alkotnak az $(1, 0)$, $(0, 1)$ elemek. Ezek segítségével felírhatjuk:

$$f(x_1, x_2) = f(x_1 \cdot (1, 0) + x_2 \cdot (0, 1))$$

Mivel f lineáris, ezért tagonként alkalmazhatjuk f-et, és az x_1 és x_2 szorzó kihozható:

$$f(x_1 \cdot (1, 0) + x_2 \cdot (0, 1)) = x_1 \cdot f(1, 0) + x_2 \cdot f(0, 1).$$

Vezessük be az $a_1 := f(1, 0), a_2 := f(0, 1)$ jelöléseket. Ezzel $f(x_1, x_2) = a_1 x_1 + a_2 x_2$.

Ez azt jelenti, hogy az $f : \mathbb{R}^2 \to \mathbb{R}$ lineáris leképezések megadhatók két számmal: a_1 és a_2, hozzátéve, hogy melyik az x_1 és melyik az x_2 együtthatója, vagyis egy rendezett számpárral: (a_1, a_2). (Itt megállapodás szerint előre írjuk az x_1 együtthatóját, és hátulra az x_2 együtthatóját. A sorrend fontos, hiszen pl. az $f(x_1, x_2) = 2x_1 + 3x_2$ és az $f(x_1, x_2) = 3x_1 + 2x_2$ függvény két különböző lineáris függvény!) Másképpen, az $\underline{x} := (x_1, x_2)$, $\underline{a} = (a_1, a_2)$ jelölések bevezetésével egy $f : \mathbb{R}^2 \to \mathbb{R}$ függvény pontosan akkor lineáris, ha $f(\underline{x}) = \underline{a} \cdot \underline{x}$ alakú, ahol a pont az R^2-beli skaláris szorzást jelenti. (Vegyük észre a valós lineáris függvényekkel való hasonlóságot!)

Térjünk át az $f : \mathbb{R}^2 \to \mathbb{R}^2$ leképezésekre. Ezek már számpárokhoz számpárokat rendelnek:

$$f(x_1, x_2) \mapsto (y_1, y_2)$$

Pl. $f(x_1, x_2) = (2x_1 + x_2, 3x_1 x_2^2)$.
Világos, hogy a képvektor y_1 és y_2 koordinátája is x_1 és x_2 függvénye:

$$y_1 = f_1(x_1, x_2), \quad y_2 = f_2(x_1, x_2),$$

ahol az f_1 és f_2 függvények $\mathbb{R}^2 \to \mathbb{R}$ típusúak, és úgy nevezzük őket, hogy f első és második koordináta-függvénye. (A megadott konkrét példában $f_1(x_1, x_2) = 2x_1 + x_2$, ill. $f_2(x_1, x_2) = 3x_1 x_2^2$.)
Nem nehéz látni, hogy a lineáris függvény definíciója az $f : \mathbb{R}^2 \to \mathbb{R}^2$ leképezésekre is átvihető. Mivel az $\mathbb{R}^2$-beli vektorok egyenlősége a megfelelő elemeik egyenlőségét jelenti, így egy $f : \mathbb{R}^2 \to \mathbb{R}^2$ függvény pontosan akkor lesz lineáris, ha f_1 és f_2 is lineáris, azaz

ha a függvény

$$f(x_1, x_2) = (a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)$$

alakú, ahol $a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R}$ rögzített számok.
Látható, hogy az $f : \mathbb{R}^2 \to \mathbb{R}^2$ lineáris leképezések négy számmal adhatók meg: a_{11}, a_{12}, a_{21}, a_{22}. Mivel fontos az is, hogy melyik szám hol áll, ezért ezeket a számokat egy két sorból és két oszlopból álló számtáblázattal, ún. 2×2-es mátrixszal adjuk meg:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Mindig az első sorba írjuk az y_1-ben szereplő együtthatókat: az első helyre az x_1, a második helyre az x_2 együtthatóját, és a második sorba az y_2 együtthatóit, szintén az első helyre az x_1 a második helyre az x_2 együtthatóját. Hansúlyozzuk, hogy minden $f : \mathbb{R}^2 \to \mathbb{R}^2$ lineáris függvényhez egyértelműen hozzá tudunk rendelni egy ilyen 2×2-es "táblázatot", és ez fordítva is igaz: minden 2×2-es táblázatnak egyértelműen megfeleltethető egy $f : \mathbb{R}^2 \to \mathbb{R}^2$ lineáris függvény.

A gyakorlatban fontos az $f : \mathbb{R}^n \to \mathbb{R}^m$ típusú leképezések vizsgálata, ahol $n, m \in \mathbb{N}$ tetszőleges. Ilyenkor a függvény n-változós, és értéke minden R^n-beli pontban egy m-dimenziós vektor:

$$f : (x_1, x_2, \ldots, x_n) \mapsto (y_1, y_2, \ldots, y_m).$$

A képvektor mindegyik koordinátája $x_1, x_2, \ldots, x_n$ függvénye, azaz

$$\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_n) \\ y_2 &= f_2(x_1, x_2, \ldots, x_n) \\ &\vdots \\ y_m &= f_m(x_1, x_2, \ldots, x_n) \end{aligned}$$

ahol az $f_i, i = 1, 2, \ldots, m$ függvények $\mathbb{R}^n \to \mathbb{R}$ típusúak (f koordináta-függvényei).

A linearitás az ilyen típusú függvények körében is értelmezhető, és belátható, hogy egy $f : \mathbb{R}^n \to \mathbb{R}^m$ függvény pontosan akkor lineáris, amikor

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{aligned}$$

alakú, ahol a_{ij}, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ rögzített valós számok. Itt már $m \cdot n$ darab valós szám határozza meg a leképezést, pontosabban egy $m \times n$-es mátrix:

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

Úgy is szokás fogalmazni, hogy egy $f : \mathbb{R}^n \to \mathbb{R}^m$ lineáris leképezés reprezentálható egy $m \times n$-es mátrixszal.
Az $m \times n$-es mátrixok halmazát $\mathbb{R}^{m \times n}$ jelöli. A mátrixokat nagybetűvel (pl. A), a mátrix i-edik sorának j-edik elemét pedig a megfelelő kisbetűvel és indexelve (a_{ij}) szokásos jelölni.

Megjegyezzük, hogy a lineáris függvény fogalma egészen általánosan a következőképpen értelmezhető:

5.3 Definíció. *Legyen $(X_1, \oplus_1, \odot_1)$ és $(X_2, \oplus_2, \odot_2)$ két tetszőleges vektortér. Az $f : X_1 \to X_2$ leképezést lineárisnak nevezzük, ha igaz rá a következő két tulajdonság:*

1. $f(x_1 \oplus_1 x_2) = f(x_1) \oplus_2 f(x_2) \quad \forall x_1, x_2 \in V_1,$

2. $f(\lambda \odot_1 x) = \lambda \odot_2 f(x) \quad \forall x \in X_1, \forall \lambda \in \mathbb{R}.$

5.2. Speciális mátrixok

- Az $1 \times n$-es mátrixokat **sormátrix**nak nevezzük.
- Az $n \times 1$-es mátrixokat **oszlopmátrix**nak nevezzük.
- Az $n \times n$-es mátrixokat **négyzetes (kvadratikus)** mátrixnak nevezzük. A gyakorlatban ez utóbbiak a legfontosabbak.

A négyzetes mátrixokon belül további speciális mátrixokat definiálunk.

- **Diagonális mátrix**: $a_{ij} = 0$, ha $i \neq j$, pl. **identitásmátrix** vagy **egységmátrix** (a főátlóban egyesek)
- **Szimmetrikus mátrix**: $a_{ij} = a_{ji}$, $i, j = 1, 2 \ldots, n$ (a főátlóra szimmetrikusak az elemek)
- **Antiszimmetrikus mátrix**: $a_{ij} = -a_{ji}$, $i, j = 1, 2 \ldots, n$ (a főátlóra szimmetrikusan elhelyezkedő elemek egymás (-1)-szeresei, a főátlóban nullák vannak!)
- **Háromszögmátrix**ok: felsőháromszög-mátrix: $a_{ij} = 0$, ha $i > j$, ill. alsóháromszögmátrix: $a_{ij} = 0$, ha $i < j$.

5.3. Műveletek mátrixokkal

A mátrixok körében többféle műveletet értelmezünk. Láttuk, hogy a mátrixok azért fontosak, mert segítségükkel egyszerűen megadhatjuk az $f : \mathbb{R}^n \to \mathbb{R}^m$ ható lineáris leképezéseket. Az ilyen leképezéseknek létezik összege, skalárszorosa, kompozíciója. Éppen ezért a mátrixok közötti műveleteket nem öncélúan értelmezzük, hanem úgy, hogy szoros kapcsolatban legyenek a lineáris leképezések körében végzett műveletekkel.

5.3.1. Mátrixok összeadása

Célunk: úgy értelmezni az összeadást mátrixok között, hogy két mátrix összege a megfelelő lineáris leképezések összegét reprezentálja.

Tekintsünk két $\mathbb{R}^2 \to \mathbb{R}^2$ lineáris leképezést:

$$f(x_1, x_2) = (a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2),$$

és

$$g(x_1, x_2) = (b_{11}x_1 + b_{12}x_2, b_{21}x_1 + b_{22}x_2).$$

(Így tehát az f leképezés mátrixa $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, a g leképezésé pedig $\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$.)

Képezzük f és g összegét, szem előtt tartva, hogy két függvényt mindig úgy adunk össze, hogy az értelmezési tartomány minden egyes pontjában összeadjuk a két függvényértéket:

$$(f+g)(x_1, x_2) = f(x_1, x_2)+g(x_1, x_2) = (a_{11}x_1+a_{12}x_2, a_{21}x_1+a_{22}x_2, b_{11}x_1+b_{12}x_2, b_{21}x_1+b_{22}x_2).$$

Látható, hogy eredményül szintén $f : \mathbb{R}^2 \to \mathbb{R}^2$ lineáris leképezést kaptunk, amelynek mátrixa

$$\begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}.$$

Vegyük észre, hogy az f és g leképezés mátrixának megfelelő elemei összeadódtak. Ez általánosabban, $\mathbb{R}^n \to \mathbb{R}^m$ lineáris leképezésekre is igaz. Ez motiválja a következő definíciót:

5.4 Definíció. *Az $A, B \in \mathbb{R}^{m\times n}$ mátrixok $A + B$ összegén azt a $C \in \mathbb{R}^{m\times n}$ mátrixot értjük, amelynek elemeire*

$$c_{ij} := a_{ij} + b_{ij}, \quad i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n.$$

Egyszerűbben fogalmazva: $m \times n$-es mátrix csak $m \times n$-es mátrixszal adható össze, és az összeadást elemenként végezzük. A mátrixok összeadása a következő műveleti tulajdonságokkal rendelkezik:

- Kommutatív, azaz $A+B=B+A \quad \forall A,B \in \mathbb{R}^{m\times n}$
- Asszociatív, azaz $(A+B)+C=A+(B+C) \quad \forall A,B,C \in \mathbb{R}^{m\times n}$
- A csupa nulla elemű $m\times n$-es mátrix (jelölje most 0) rendelkezik azzal a tulajdonsággal, hogy bármely mátrixhoz adva az illető mátrixok kapjuk vissza: $A+0=A$.
- Minden $A=(a_{ij})\in \mathbb{R}^{m\times n}$ mátrixhoz létezik negatív mátrix: $-A=(-a_{ij})\in \mathbb{R}^{m\times n}$.

Vegyük észre, hogy ezek a tulajdonságok megegyeznek a vektorok összeadásának tulajdonságaival.

5.5 Megjegyzés. *Csak azonos méretű mátrixokat lehet összeadni, ugyanis a megfelelő f,g függvények összeadásának csak így van értelme!*

5.3.2. Mátrixok szorzása skalárral

Célunk úgy értelmezni a mátrix skalárszorosát, hogy az a mátrixnak megfelelő lineáris leképezés skalárszorosát reprezentálja.

Képezzük az

$$f(x_1,x_2)=(a_{11}x_1+a_{12}x_2, a_{21}x_1+a_{22}x_2)$$

lineáris leképezés $\lambda\in\mathbb{R}$ skalárszorosát. Figyelembe véve, hogy egy függvény skalárral való szorzása azt jelenti, hogy a függvényértéket minden helyen megszorozzuk az adott számmal:

$$\begin{aligned}(\lambda\cdot f)(x_1,x_2) &= \lambda\cdot f(x_1,x_2)=\lambda\cdot(a_{11}x_1+a_{12}x_2,a_{21}x_1+a_{22}x_2)=\\ &= (\lambda a_{11}x_1+\lambda a_{12}x_2,\lambda a_{21}x_1+\lambda a_{22}x_2).\end{aligned}$$

Látható, hogy lineáris leképezést kaptunk ($\mathbb{R}^2\to\mathbb{R}^2$), amelynek mátrixa

$$\begin{bmatrix}\lambda a_{11} & \lambda a_{12}\\ \lambda a_{21} & \lambda a_{22}\end{bmatrix}.$$

Vagyis az f lineáris leképezés mátrixának minden eleme λ-val szorzódik. Ugyanezt a megfigyelést tehetjük, ha tetszőleges, $\mathbb{R}^n\to\mathbb{R}^m$ lineáris leképezés skalárszorosát vizsgáljuk. Ez motiválja a következő definíciót:

5.6 Definíció. *Az $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ mátrixnak a $\lambda \in \mathbb{R}$ skalárszorosán azt az $\tilde{A} = (\tilde{a}_{ij}) \in \mathbb{R}^{m\times n}$ mátrixot értjük, amelyre $\tilde{a}_{ij} = \lambda \cdot a_{ij}, i = 1, 2 \dots, m, j = 1, 2, \dots, n$. Jelölése: $\lambda \cdot A$ vagy egyszerűbben λA.*

Mátrixot tehát úgy szorzunk valós számmal, hogy a mátrix minden elemét megszorozzuk az adott számmal. A mátrixok skalárral való szorzása az alábbi műveleti tulajdonságokkal rendelkezik.

- $\lambda(A + B) = \lambda A + \lambda B \quad \forall A, B \in \mathbb{R}^{m\times n},\ \forall \lambda \in \mathbb{R}$;
- $(\lambda + \mu)A = \lambda A + \mu A \quad \forall A \in \mathbb{R}^{m\times n},\ \forall \lambda, \mu \in \mathbb{R}$;
- $\lambda(\mu A) = (\lambda\mu)A \quad \forall A \in \mathbb{R}^{m\times n},\ \forall \lambda, \mu \in \mathbb{R}$.

Vegyük észre, hogy ezek megegyeznek a vektorok skalárral való szorzásának tulajdonságaival. Megállapíthatjuk tehát, hogy $\mathbb{R}^{m\times n}$ az előbb értelmezett összeadással és skalárral valós szorzással vektorteret alkot.

5.3.3. Mátrixszorzás

A mátrixok közötti szorzásműveletet úgy szeretnénk értelmezni, hogy két mátrix szorzata a megfelelő lineáris leképezések kompozícióját reprezentálja.

Nyilvánvalóan, ha egy f lineáris (vagy nem lineáris) leképezés pl. $\mathbb{R}^n$-ből $\mathbb{R}^m$-be képez, akkor ennek csak abban az esetben képezhetjük a kompozícióját egy másik hasonló típusú leképezéssel $g \circ f$ sorrendben, ha $g : \mathbb{R}^m \to \mathbb{R}^l$. Itt l tetszőleges pozitív egész lehet, de m nem tetszőleges: az $\mathbb{R}^m$ az f függvény képtere. Mit lehet mondani egy $f : \mathbb{R}^n \to \mathbb{R}^m$ és egy $g : \mathbb{R}^m \to \mathbb{R}^l$ lineáris leképezés kompozíciójáról? Az egyszerűség kedvéért azt az esetet nézzük meg részletesebben, amikor $n = m = l = 2$, azaz mindkét lineáris leképezés $\mathbb{R}^2 \to \mathbb{R}^2$ típusú. Legyen tehát

$$f(x_1, x_2) = (a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)$$

és

$$g(y_1, y_2) = (b_{11}y_1 + b_{12}y_2, b_{21}y_1 + b_{22}y_2).$$

A $g \circ f$ kompozíció felírásához a g függvény argumentumában y_1 és y_2 helyébe helyettesítsük be az (x_1, x_2) f-képét:

$$(g\circ f)(x_1, x_2) = (b_{11}(a_{11}x_1+a_{12}x_2)+b_{12}(a_{21}x_1+a_{22}x_2), b_{21}(a_{11}x_1+a_{12}x_2)+b_{22}(a_{21}x_1+a_{22}x_2))$$

Az eredmény koordináta-függvényeiben x_1 és x_2 együtthatóit leolvasva látható, hogy $g \circ f$ szintén lineáris leképezés $\mathbb{R}^2 \to \mathbb{R}^2$, és a következő mátrixszal azonosítható:

$$\begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{bmatrix}.$$

Ennek a mátrixnak az i-edik sorában és j-edik oszlopában $(i, j = 1, 2)$ lévő elemét úgy kaphatjuk meg, hogy a B mátrix i-edik sorában lévő vektort skalárisan szorozzuk az A mátrix j-edik oszlopában lévő vektorral. Általában pedig, ha $f : \mathbb{R}^n \to \mathbb{R}^m$ és $g : \mathbb{R}^m \to \mathbb{R}^l$ lineáris leképezések, akkor $(g \circ f) : \mathbb{R}^n \to \mathbb{R}^l$ lineáris leképezés lesz, amelynek mátrixa $l \times n$-es, és i-edik sorának j-edik eleme a $\sum_{k=1}^{m} b_{ik}a_{kj}, i = 1, \ldots, l,\ j = 1, \ldots n$ képlettel adható meg.

Két mátrix egymással való szorzatát csak arra az esetre értelmezzük, amikor a megfelelő lineáris leképezéseknek értelmes a kompozíciója, és úgy definiáljuk, hogy a szorzatmátrix a kompozícióleképezést reprezentálja.

5.7 Definíció. *Az $A \in \mathbb{R}^{l \times m}$ és $B \in \mathbb{R}^{m \times n}$ mátrixok $A \cdot B$ szorzatán azt a $C \in \mathbb{R}^{l \times n}$ mátrixot értjük, amelynek elemeire*

$$c_{ij} := \sum_{k=1}^{m} a_{ik}b_{kj}, \quad i = 1, 2, \ldots, l,\ j = 1, 2, \ldots, n.$$

Vegyük észre, hogy ha adva van egy f lineáris leképezés $\mathbb{R}^n$-ből $\mathbb{R}^m$-be, akkor egy $\underline{x} \in \mathbb{R}^n$ vektor képét (vagyis az $f(\underline{x})$ vektort) is meg tudjuk kapni mátrixszorzás segítségével: a leképezés mátrixával megszorozzuk az $\underline{x}$ vektort mint $n \times 1$-es oszlopmátrixot. Egy $\mathbb{R}^n$-ből $\mathbb{R}^m$-be képező lineáris leképezés tehát nem más, mint egy $m \times n$-es mátrixszal való szorzás.

A mátrixok szorzása az alábbi műveleti tulajdonságokkal rendelkezik.

- Nem kommutatív, vagyis nem igaz az, hogy minden $A \in \mathbb{R}^{l \times m}$ és $B \in \mathbb{R}^{m \times n}$ mátrixra $AB = BA$.
- Asszociatív, azaz $(AB)C = A(BC)$ minden olyan A, B, C mátrixra, amelyekre értelmes a két oldal.
- Disztributív, azaz $(A + B)C = AC + BC$ és $C(A + B) = CA + CB$ minden olyan A, B, C mátrixra, amelyekre értelmes a két oldal.

5.4. Az inverz mátrix

A mátrixok szorzása kapcsán bevezetjük az inverz mátrix fogalmát.

5.8 Definíció. *Az $A \in \mathbb{R}^{n\times n}$ mátrixot* **invertálható***nak nevezzük, ha létezik olyan $X \in \mathbb{R}^{n\times n}$ mátrix, amelyre $A \cdot X = I$, ahol I az $n \times n$-es identitásmátrix. Ekkor az X mátrixot az A mátrix* **inverz***ének nevezzük és A^{-1}-gyel jelöljük.*

Megjegyzés. Belátható, hogy ha $A \cdot X = I$, akkor $X \cdot A = I$ is automatikusan teljesül, tehát az inverz mátrixra $A \cdot A^{-1} = I$ és $A^{-1} \cdot A = I$. Ha az A mátrix az $f : \mathbb{R}^n \to \mathbb{R}^n$ lineáris leképezért reprezentálja, akkor A^{-1} ezen függvény inverzét reprezentálja, vagyis azt az $f^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ függényt, amelyre $f \circ f^{-1} = \mathrm{id}$ és $f^{-1} \circ f = \mathrm{id}$.

Az inverz mátrix létezésének feltételeivel és elemeinek kiszámításával később (a determináns definiálása után) foglalkozunk.

5.5. A transzponált mátrix

5.9 Definíció. *Az $A \in \mathbb{R}^{m\times n}$ mátrix* **transzponált***jának nevezzük azt az $A^T \in \mathbb{R}^{n\times m}$ mátrixot, amelynek elemeire*

$$a^T_{ij} := a_{ji}, \quad i = 1, 2, \ldots n, \ j = 1, 2, \ldots, m.$$

Ez tehát az a mátrix, amelyet A sorainak és oszlopainak a felcserélésével kapunk.

6. Lineáris algebrai egyenletrendszerek

6.1 Definíció. *Lineáris algebrai egyenletrendszernek (röviden lineáris egyenletrendszernek) nevezzük az*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{sn}x_n &= b_m \end{aligned} \tag{1}$$

alakú egyenletrendszereket, ahol $x_1, x_2, \ldots, x_n$ az ismeretlen valós (vagy komplex) számok, a_{ij} $(i = 1, \ldots, m, j = 1, \ldots, n)$ (az ismeretlenek együtthatói), és $b_1, b_2 \ldots, b_m$ (az ún. szabad tagok) adott valós (vagy komplex) számok.

Pl. a következő egyenletrendszer egy két egyenletből álló $(m = 2)$ kétismeretlenes $(n = 2)$ lineáris egyenletrendszer:

$$\begin{aligned} x_1 + x_2 &= 3 \\ 2x_1 + 2x_2 &= 5 \end{aligned}$$

6.2 Definíció. *Az (1) lineáris egyenletrendszer egy megoldásának nevezzük az n számból álló $\lambda_1, \lambda_2, \ldots, \lambda_n$ rendszert, ha az $x_i = \lambda_i$ $(i = 1, 2, \ldots, n)$ helyettesítéssel mindegyik egyenletben azonosságot kapunk.*

A lineáris egyenletrendszerekkel kapcsolatban a következő alapkérdéseket fogalmazhatjuk meg:

1. Megoldható-e az egyenletrendszer?

2. Ha igen, hány megoldása van? (Megoldani az egyenletrendszert azt jelenti, hogy meghatározzuk az összes megoldását!)

3. Hogyan határozható(k) meg a megoldás(ok)?

6.3 Definíció. *(Homogén és inhomogén lineáris egyenletrendszer)*
Ha $b_i = 0$, $i = 1, 2, \ldots, m$ (vagyis az összes szabad tag nulla az egyenletrendszerben), akkor **homogén**, *egyébként* **inhomogén** *lineáris egyenletrendszerről beszélünk.*

A továbbiakban gyakran alkalmazzuk az egyenletrendszer rövidített jelölésmódját.

6.4 Definíció. *Az*

$$A := \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

mátrixot az (1) lineáris egyenletrendszer együtthatómátrixának nevezzük.

Vezessük be az

$$\underline{x} := \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \quad \text{és} \quad \underline{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

jelöléseket is. Ezek segítségével (1) felírható

$$A\underline{x} = \underline{b} \tag{2}$$

alakban, és az egyenletrendszer megoldása azt jelenti, hogy meghatározzuk mindazon $\underline{x}$ vektorokat, amelyek kielégítik a (2) egyenlőséget.

Pl. 1. Mely egyenletrendszernek felel meg az $A\underline{x} = \underline{b}$ egyenlet, ha

$$A = \begin{bmatrix} 3 & 1 & 4 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{és} \quad \underline{b} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}?$$

Megoldás:

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

2. Írjuk fel az

$$\begin{aligned} x_2 &= 2 \\ x_1 + x_2 &= 3 \end{aligned}$$

egyenletrendszer együtthatómátrixát!
Megoldás:

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

Az egyenletrendszer megoldhatóságának vizsgálatában szükségünk lesz egy, a négyzetes mátrixokhoz rendelhető szám, az ún. determináns ismeretére. A továbbiakban ennek értelmezésével foglalkozunk.

6.1. Másod- és harmadrendű determinánsok

Tekintsük az

$$a_{11}x_1 + a_{12}x_2 = b_1 \tag{3}$$
$$a_{21}x_1 + a_{22}x_2 = b_2 \tag{4}$$

lineáris egyenletrendszert. Ennek együtthatómátrixa az

$$A := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

mátrix.

Középiskolából ismert megoldási módszer az állandó együtthatók módszere: kiküszöböljük az egyik ismeretlent a következő módon: az egyenleteket megfelelő számokkal szorozva elérjük, hogy az egyik ismeretlen együtthatója mindkét egyenletben ugyanaz legyen, és az egyenleteket kivonjuk egymásból. Küszöböljük ki az x_2-t. Ehhez az első egyenletet a_{22}-vel, a másodikat a_{12}-vel kell megszoroznunk, és eredményül az

$$(a_{11}a_{22} - a_{21}a_{12})x_1 = b_1a_{22} - b_2a_{12} \tag{5}$$

egyenletet kapjuk. Ha x_1-et küszöböljük ki $((4) \cdot a_{11} - (3) \cdot a_{21})$, akkor az

$$(a_{11}a_{22} - a_{21}a_{12})x_2 = b_2a_{11} - b_1a_{21} \tag{6}$$

egyenlethez jutunk.

Látható: az (3)-(4) egyenletrendszer megoldhatósága szempontjából három eset van:

- Ha $a_{11}a_{22} - a_{21}a_{12} \neq 0$, akkor x_1 és x_2 kifejezhető, egy (azaz egyértelmű) megoldás van.
- Ha $a_{11}a_{22} - a_{21}a_{12} = 0$, akkor két eset lehetséges:
  1. Ha (5) és (6) jobb oldala is 0, akkor végtelen sok megoldás van. (Vigyázat, ez nem jelenti azt, hogy bármilyen x_1 és x_2 szám megoldás!)
  2. Ha legalább az egyik egyenlet jobb oldala nem nulla, akkor nincs megoldás.

Foglalkozzunk azzal az esettel, amikor egyértelmű megoldás van. Ekkor a megoldás:

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}}, \quad x_2 = \frac{b_2a_{11} - b_1a_{21}}{a_{11}a_{22} - a_{21}a_{12}}.$$

A nevező csak az együtthatómátrix elemeit tartalmazza. Úgy kapjuk, hogy az A mátrix főátlóbeli elemeinek a szorzatából kivonjuk a másik két elem szorzatát. Az ezen szabály szerint egy 2×2-es mátrixhoz rendelt számnak külön nevet is adunk.

6.5 Definíció. *(2×2-es mátrix determinánsa)*
Az $a_{11}a_{22} - a_{21}a_{12}$ számot az

$$A := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

mátrix **determinláns***ának nevezzük. Jelölése:* $|A|$, $\det A$ *vagy* $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$.

A 2×2-es mátrixok determinánsát másodrendű determinánsnak is nevezzük.

Ezen új fogalmat felhasználva tehát az x_1 és x_2 nevezőjében álló A együtthatómátrix determinánsa. Vegyük észre továbbá, hogy az x_1 számlálójában álló kifejezés nem más, mint az

$$A_1 := \begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix}$$

mátrix determinánsa, az x_2 számlálója pedig az

$$A_2 := \begin{bmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{bmatrix}$$

mátrix determinánsa.
Összefoglalva: a (3)–(4) egyenletrendszernek pontosan akkor létezik egyértelmű megoldása, ha $\det A \neq 0$, és ekkor a megoldás

$$x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A}.$$

A megoldás kiszámításának ezen módját Cramer-szabálynak nevezzük.

A továbbiakban a determináns fogalmát kiterjesztjük nagyobb méretű mátrixokra, és látni fogjuk, hogy a Cramer-szabály nagyobb méretű lineáris egyenletrendszerekre is alkalmazható.

Tekintsük az

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned} \tag{7}$$

lineáris egyenletrendszert. Ennek együtthatómátrixa az

$$A := \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3} \tag{8}$$

mátrix.

6.6 Definíció. *Az $a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$ számot az (8) mátrix determinánsának nevezzük (harmadrendű determináns). Jelölése:* $|A|$, $\det A$ *vagy* $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$.

6.7 Megjegyzés. *A definícióban szereplő hat tagú összeg a vektoriális és a vegyes szorzat Descartes-koordinátás alakjánál már látott módon számítódik az A mátrix elemeiből.*

Az ismeretlenek kiküszöbölésével itt is megmutatható, hogy a (7) egyenletrendszernek pontosan akkor létezik egyértelmű megoldása, ha $\det A \neq 0$, és a megoldás:

$$x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A}, \quad x_3 = \frac{\det A_3}{\det A},$$

ahol A_i $(i = 1, 2, 3)$ az a mátrix, amelyet A-ból úgy kapunk, hogy az i-edik oszlopát kicseréljük $\underline{b}$-re, pl.

$$A_2 = \begin{bmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{bmatrix}.$$

A Cramer-szabály tehát itt is érvényes. Ha $\det A = 0$, akkor nincs egyértelmű megoldás (nincs megoldás vagy végtelen sok van.)

6.2. n-edrendű determinánsok

Vegyük észre, hogy a 3×3-as mátrixok determinánsa összerakható 2×2-es determinánsokból:

$$\det A = a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \cdot \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Azt mondjuk, hogy a determinánst kifejtettük az első sor szerint. Ez általánosabban is igaz: Legyen i egy tetszőleges index $(i \in \{1, 2, 3\})$. Ekkor

$$\det A = \sum_{j=1}^{3} a_{ij} \cdot (-1)^{i+j} \cdot |A_{ij}|,$$

ahol $|A_{ij}|$ annak a mátrixnak a determinánsa, amelyet A-ból úgy kapunk, hogy az a_{ij} elemet tartalmazó sort és oszlopot elhagyjuk. (A determinánst kifejtjük az i-edik sor szerint.) Itt $|A_{ij}|$ neve: az a_{ij} elemhez adjungált aldetermináns. Az $(-1)^{i+j} \cdot |A_{ij}|$ kifejezést előjeles adjungált alderemninánsnak nevezzük, és erre az $|A_{ij}|^{\pm}$ jelölést fogjuk alkalmazni. Megjegyezzük, hogy nemcsak bármelyik sor, hanem bármelyik oszlop szerint is kifejthető a determináns.

Ez az eljárás, amellyel a harmadrendű determináns kiszámítását visszavezetjük másodrendű determinánsok kiszámítására, lehetőséget ad az n-edrendű determináns értelmezésére, ahol $n \in \mathbb{N}$ tetszőleges. (Nyilvánvalóan, egy n-ed rendű determináns n! tagú összegből áll.)

6.8 Definíció. *Az $A \in \mathbb{R}^{n\times n}$ mátrix determinánsának nevezzük a*

$$\det A := \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|A_{ij}|$$

számot, ahol $i \in \{1,2,\ldots,n\}$ tetszőleges sorindex.

A fenti képletben az A_{ij} mátrix $(n-1)\times(n-1)$-es. (Belátható, hogy valóban tetszőleges sorindexet választva ugyanahhoz a számhoz jutunk, tehát értelmes a definíció. Sőt, oszlop szerint kifejtve is ugyanahhoz a számhoz jutunk.)
A determinánst csak négyzetes mátrixokra értelmezzük. Tetszőleges mátrixnak vannak viszont aldeterminánsai.

6.9 Definíció. *Az $A \in \mathbb{R}^{m\times n}$ mátrix k-adrendű aldeterminánsának nevezünk minden olyan k-adrendű determinánst, amelyet úgy kapunk, hogy az A mátrixból kiválasztunk k darab sort és k darab oszlopot ($k \leq \min(m,n)$), és az ezek kereszteződéseiben lévő elemekből összeállított mátrix determinánsát képezzük.*

6.2.1. A determináns tulajdonságai

Tetszőleges $A \in \mathbb{R}^{n\times n}$ mátrixok determinánsára igazak a következő tulajdonságok.

1. $\det A = \det A^T$, vagyis a transzponálással nem változik meg a mátrix determinánsa.
   Biz.: $\det A$-t az i-edik sor és $\det A^T$-t az i-edik oszlop szerint kifejtve ugyanazt kapjuk.

2. Ha az A mátrix valamely sora vagy oszlopa csak nullákat tartalmaz, akkor $\det A = 0$.
   Biz.: Ha a csupa nulla sor vagy oszlop szerint fejtjük ki a mátrixot, nullát kapunk.

3. Két sor felcserélésével a determináns előjele ellentétesre változik.

4. Két egyforma sort tartalmazó mátrix determinánsa 0.
   Biz.: Tegyük fel, hogy az A mátrix i-edik és j-edik sora egyenlő, $i \neq j$. Cseréljük fel ezt a két sort, ekkor a determináns $-\det A$ lesz. Egyenlő sorokat cseréltünk fel, ezért a determináns nem változott, azaz $-\det A = \det A$. Ez csak úgy lehet, ha $\det A = 0$.

5. Ha egy sor (oszlop) minden elemét megszorozzuk ugyanazzal a $\lambda \in \mathbb{R}$ számmal, akkor a determináns a λ-szorosára változik.

6. Ha a mátrix egyik sora (oszlopa) előáll egy másik sor (oszlop) számszorosaként, akkor a mátrix determinánsa 0.

7. Ha A valamelyik sora (oszlopa) a többi sorok (oszlopok) lineáris kombinációja, akkor $\det A = 0$.

8. A determináns nem változik, ha valamely sorának (oszlopának) minden eleméhez hozzáadjuk egy másik sor (oszlop) számszorosát.
Pl.
$$\begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 7+1 & 8+2 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 8 & 10 \end{vmatrix}$$

9. $A, B \in \mathbb{R}^{n\times n}$ esetén $\det(A \cdot B) = \det A \cdot \det B$.

6.2.2. Speciális mátrixok determinánsa

- Az $I \in \mathbb{R}^{n\times n}$ identitásmátrixra $\det I = 1$.
- Diagonális és háromszögmátrixok determinánsa a főátlóbeli elemek szorzata.

6.3. Az inverz mátrix létezésének feltétele és elemeinek megadása

A determináns segítségével meg tudjuk határozni egy mátrix inverzének elemeit. (Erre szükségünk lesz a lineáris egyenletrendszerek megoldásának felírásakor.) Ezért ebben az alfejezetben ezzel a kérdéssel foglalkozunk.

Láttuk, hogy egy $A \in \mathbb{R}^{n\times n}$ mátrix pontosan akkor invertálható, ha létezik olyan $X \in \mathbb{R}^{n\times n}$, amelyre $A \cdot X = I$ és $X \cdot A = I$. X helyett az A^{-1} jelölést alkalmazzuk. Könnyen meggondolható, hogy ha A-nak létezik inverze, akkor az egyértelmű. Tegyük fel ugyanis, hogy A-nak két inverze is létezik, A^{-1} és $\tilde{A}^{-1}$. Ekkor

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

Szorozzuk meg az egyenlőség mindkét oldalát balról a másik inverzzel, azaz A^{-1}-gyel:

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

A jobb oldalon A^{-1} áll, ugyanis az identitásmátrixszal tetszőleges $M \in \mathbb{R}^{n\times n}$ mátrixot szorozva (bármilyen sorrendben) M-et kapjuk. A bal oldalon $A^{-1} \cdot A \cdot \tilde{A}^{-1} = (A^{-1} \cdot A) \cdot \tilde{A}^{-1} = I \cdot \tilde{A}^{-1} = \tilde{A}^{-1}$. Ebből $A^{-1} = \tilde{A}^{-1}$, vagyis a két inverz mátrix nem lehet különböző, csak egy inverz mátrix létezik.

Eddig nem foglalkoztunk azzal a kérdéssel, hogy mikor létezik A^{-1}.

6.10 Definíció. *Az $A \in \mathbb{R}^{n\times n}$ mátrixot szingulárisnak nevezzük, ha* $\det A = 0$, *és regu-lárisnak, ha* $\det A \neq 0$.

6.11 Állítás. *Szinguláris mátrixnak nem létezik inverze.*

Biz.: Ha létezne inverze, akkor a determináns 9. tulajdonsága miatt $\det A \cdot \det A^{-1} = \det(A \cdot A^{-1}) = \det I = 1$ lenne, de ha $\det A = 0$, akkor ez nem lehetséges.

Megmutatjuk, hogy ha A reguláris, akkor létezik A^{-1}, és meg is adjuk az elemeit.

6.12 Definíció. *Legyen $A \in \mathbb{R}^{n \times n}$. Jelölje $A^* = (a^*_{ij})$ azt az $n \times n$-es mátrixot, amelyre $a^*_{ij} = |A_{ji}|^{\pm}$. Ezt az A^* mátrixot az A mátrix adjungált mátrixának nevezzük.*

$$A^* = \begin{bmatrix} |A_{11}|^{\pm} & |A_{21}|^{\pm} & \dots & |A_{n1}|^{\pm} \\ |A_{12}|^{\pm} & |A_{21}|^{\pm} & \dots & |A_{n1}|^{\pm} \\ \vdots & & . & \vdots \\ |A_{1n}|^{\pm} & |A_{2n}|^{\pm} & \dots & |A_{nn}|^{\pm} \end{bmatrix}$$

6.13 Állítás. *Az A mátrixnak és az A^* adjungált mátrixnak a szorzata olyan diagonálmátrix, amelynek főátlóbeli elemei az A determinánsával egyenlők, azaz*

$$A \cdot A^* = A^* \cdot A = \begin{bmatrix} \det A & 0 & . & 0 \\ 0 & \det A & 0 & . \\ . & . & . & 0 \\ 0 & . & 0 & \det A \end{bmatrix}.$$

Biz.: Végezzük el a mátrixszorzást! A szorzatmátrix i-edik sorában és j-edik oszlopában a

$$c_{ij} := \sum_{k=1}^{n} a_{ik} |A_{jk}|^{\pm}$$

szám áll.

- Ha $i = j$, akkor $c_{ij} := \sum_{k=1}^{n} a_{ik} |A_{ik}|^{\pm}$, ami egyenlő $\det A$-val, mivel ez az összeg éppen az A mátrix determinánsának az i-edik sor szerinti kifejtése.
- Ha $i \neq j$, akkor $c_{ij} = 0$, ugyanis ekkor a c_{ij} összeg azon mátrix determinánsának a kifejtése, amelyet A-ból úgy kapunk, hogy a j-edik sorába az i-edik sorát tesszük. Ennek a mátrixnak i-edik és j-edik sorában megegyeznek az elemek, ezért a determinánsa nulla.

6.14 Következmény. *Ha* $\det A \neq 0$, *akkor fennáll az*

$$\frac{1}{\det A} \cdot A \cdot A^* = \frac{1}{\det A} \cdot A^* \cdot A = I$$

egyenlőség, amelyből azonnal látható, hogy A-nak létezik inverze, és

$$A^{-1} = \begin{bmatrix} \frac{|A_{11}|^{\pm}}{\det A} & \frac{|A_{21}|^{\pm}}{\det A} & . & \frac{|A_{n1}|^{\pm}}{\det A} \\ \frac{|A_{12}|^{\pm}}{\det A} & . & . & . \\ . & . & . & . \\ \frac{|A_{1n}|^{\pm}}{\det A} & . & . & \frac{|A_{nn}|^{\pm}}{\det A} \end{bmatrix}.$$

Tehát A^{-1} i-edik sorában az A mátrix i-edik oszlopában lévő elemekhez tartozó előjeles adjungált aldeterminánsok állnak, $\det A$*-val osztva.*

Ha $A \in \mathbb{R}^{2\times 2}$, akkor az adjungált mátrixát egyszerű felírni. Számítsuk ki általánosan az adjungált márix elemeit. Legyen

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Az adjungált aldeterminánsok 1×1-es mátrixok determinánsai lesznek. Az 1×1-es mátrixok a valós számok, az ilyen mátrix determinánsa saját maga. Ennek megfelelően az adjungált mátrix elemei:

$$\begin{aligned} a^*_{11} &= |A_{11}|^{\pm} = (-1)^{1+1}|a_{11}| = a_{22} \\ a^*_{12} &= |A_{21}|^{\pm} = (-1)^{2+1}|a_{21}| = -a_{12} \\ a^*_{21} &= |A_{12}|^{\pm} = (-1)^{1+2}|a_{12}| = -a_{21} \\ a^*_{22} &= |A_{22}|^{\pm} = (-1)^{2+2}|a_{22}| = a_{11} \end{aligned}$$

Tehát az adjungált mátrix:

$$A^* = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Azt vehetjük tehát észre, hogy a 2×2-es A mátrix adjungált mátrixát úgy írhatjuk fel, hogy megcseréljük a főátlóban lévő számokat, a másik kettőt pedig az ellentettjére változtatjuk.

6.4. A Cramer-szabály

Megmutatjuk, hogy a Cramer-szabály $n \times n$-es mátrixú lineáris egyenletrendszerek megoldására is alkalmazható, amennyiben az együtthatómátrix reguláris.

Korábban bevezettük a lineáris egyenletrendszerek rövidített jelölésmódját. Ennek alapján egy $n \times n$-es mátrixú lineáris egyenletrendszer az

$$A\underline{x} = \underline{b} \tag{9}$$

alakba írható, ahol $A \in \mathbb{R}^{n\times n}$, és $\underline{x}, \underline{b} \in \mathbb{R}^n$.

Tegyük fel, hogy az A mátrix reguláris, azaz $\det A \neq 0$. Ekkor a korábbiak értelmében létezik A^{-1}. Szorozzuk meg az (9) egyenletet balról az A^{-1} mátrixszal! Így az

$$\underline{x} = A^{-1}\underline{b}$$

egyenlethez jutunk.

A jobb oldalon álló vektor j-edik elemét ($j = 1, 2, \dots, n$) az

$$\frac{|A_{1j}|^{\pm}}{\det A}b_1 + \frac{|A_{2j}|^{\pm}}{\det A}b_2 + \ldots + \frac{|A_{nj}|^{\pm}}{\det A}b_n = \frac{1}{\det A}(|A_{1j}|^{\pm}b_1 + \ldots + |A_{nj}|^{\pm}b_n) = \frac{\det A_j}{\det A}$$

összefüggés adja meg, azaz a j-edik ismeretlen előáll

$$x_j = \frac{\det A_j}{\det A}$$

alakban.

Összefoglalva, ha a lineáris egyenletrendszerben az együtthatómátrix négyzetes (tehát ugyanannyi egyenlet van, ahány ismeretlen), és reguláris az együtthatómátrix, akkor az egyenletrendszernek létezik egyetlen megoldása, és az felírható a Cramer-szabály segítségével determinánsok hányadosaként.

7. A lineáris algebrai egyenletrendszerek általános elmélete

Az előző fejezetben bevezettük a lineáris algebrai egyenletrendszerekkel kapcsolatos legfontosabb alapfogalmakat. A lineáris egyenletrendszerek megoldásáról azonban még kevés szó esett: csak négyzetes mátrixú egyenletrendszer megoldásával foglalkoztunk, és csak az egyértelmű megoldás létezésének feltételét fogalmaztuk meg. Ebben a fejezetben választ kapunk arra a kérdésre, hogy egy tetszőleges lineáris egyenletrendszerről hogyan dönthetjük el, hogy megoldható-e vagy sem, és ha megoldható, hány megoldása van.

7.1. A mátrix rangja

A mátrix rangjának a fogalma a lineáris egyenletrendszerek megoldhatóságának vizsgálatában lesz segítségünkre.

7.1 Definíció. *Az $A \in \mathbb{R}^{m\times n}$ mátrix rangján az egymással lineárisan független rendszert alkotó oszlopvektorok maximális számát értjük a mátrixban. Jelölése:* $\mathrm{rang}(A)$.

Pl. mennyi a rangja a következő mátrixoknak?

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Az A mátrix három oszlopvektora lineárisan független rendszert alkot, ezért $\mathrm{rang}(A) = 3$. A B mátrixban a három oszlopvektor összefüggő (köztük van a nullvektor), ezért a rang csak 3-nál kisebb lehet. Az $(1,0,0)$ és a $(0,0,1)$ oszlopvektor lineárisan független, ezért a rang 2. A C mátrixban az $(1,0)$ és $(0,1)$ oszlop lineárisan független, a három oszlopvektor azonban már lineárisan összefüggő, így $\mathrm{rang}(C) = 2$.
Nyilvánvalóan, ha $A = 0$ (nullmátrix), akkor a rangja 0.

A rangot a legtöbb esetben nehéz a fenti definíció alapján kiszámítani. Ezért hasznos a következő tétel, amely lehetővé teszi, hogy egy mátrix rangját az aldeterminánsainak a vizsgálatával határozzuk meg.

7.2 Tétel. *(Mátrixok rangszámtétele)*
Az $A \in \mathbb{R}^{m\times n}$ mátrixban a legmagasabb rendű nemnulla aldetermináns rendje egyenlő az A rangjával.

Biz.: Jelölje r a legmagasabb rendű nemnulla aldetermináns rendjét. (Ez azt jelenti, hogy a mátrixban található olyan $r \times r$-es aldetermináns, amelynek az értéke nemnulla, de minden annál nagyobb méretű aldeterminánsa nulla.) Az egyszerűség kedvéért tegyük fel, hogy az egyik ilyen r-edrendű aldetermináns a mátrix bal felső sarkában foglal helyet. (A bizonyítás menetéből látható, hogy ez nem jelenti az általánosság megszorítását.) Jelölje ezt a nullától különböző aldeterminánst D. Abból, hogy $D \neq 0$, következik, hogy a mátrix első r oszlopa lineárisan független. (Ha ez nem lenne igaz, akkor a D aldetermináns oszlopai is lineárisan összefüggők volnának, és ekkor $D = 0$ lenne.)

Legyen $l \in \{r+1, \ldots, n\}$ tetszőleges index, és mutassuk meg, hogy az l-edik oszlop kifejezhető az első r oszlop lineáris kombinációjaként. (Ez éppen azt jelenti, hogy a rang r.) Legyen $i \in \{1, 2, \ldots, s\}$ tetszőleges index, és jelölje Δ_i a

$$\Delta_i = \begin{vmatrix} a_{11} & . & . & a_{1r} & a_{1l} \\ . & & & & . \\ . & & & & . \\ a_{r1} & . & . & a_{rr} & a_{rl} \\ a_{i1} & . & . & a_{ir} & a_{il} \end{vmatrix}$$

determinánst. Ennek értéke nulla, ugyanis
– ha $i > r$, akkor Δ_i az A mátrix $(r+1)$-edrendű aldeterminánsa, amiről feltettük, hogy nulla;
– ha pedig $i \le r$, akkor Δ_i két sora egyenlő.
Fejtsük ki a Δ_i determinánst az utolsó sora szerint. Írjuk fel először a kifejtésben szereplő előjeles adjungált aldeterminánsokat:

$$|A_{ij}| = (-1)^{(r+1)+j} \cdot \begin{vmatrix} a_{11} & . & a_{1,j-1} & a_{1,j+1} & . & a_{1r} a_{1l} \\ . & & . & . & & . \quad . \\ . & & . & . & & . \quad . \\ a_{r1} & . & a_{r,j-1} & a_{r,j+1} & . & a_{rr} a_{rl} \end{vmatrix}, \ j = 1, \ldots, r,$$

$$|A_{il}| = (-1)^{2(r+1)} \cdot D = D.$$

Mivel az $|A_{ij}|$, $j = 1, \ldots, r$ összegek függetlenek i-től, ezért $|A_{ij}|$ helyett alkalmazzuk az $|A_{(j)}|$ jelölést. Így a Δ_i determináns felírható a

$$\Delta_i = a_{i1}|A_{(1)}| + a_{i2}|A_{(2)}| + \ldots + a_{ir}|A_{(r)}| + a_{il}D$$

alakban. Ebből, felhasználva, hogy $D \neq 0$, az

$$a_{il} = -\frac{|A_{(1)}|}{D}a_{i1} - \frac{|A_{(2)}|}{D}a_{i2} - \ldots \frac{|A_{(r)}|}{D}a_{ir}$$

összefüggést nyerjük. Mivel az itt szereplő együtthatók i-től függetlenek, ezért az egész l-edik oszlop az első r oszlop $-\frac{|A_{(1)}|}{D}, -\frac{|A_{(2)}|}{D}, \ldots, -\frac{|A_{(r)}|}{D}$ együtthatókkal vett lineáris kombinációja.

A bizonyításból látszik, hogy ha találtunk egy r-edrendű nemnulla aldeterminánst, és az ezt tartalmazó $(r+1)$-edrendű aldeterminánsok már mind nullák, akkor a mátrix rangja r. Ezért a mátrix rangját úgy számítjuk ki, hogy az alacsonyabb rendű aldeterminánsoktól a magasabb rendűek felé haladunk.

7.3 Megjegyzés. *A rangszámtételből és abból, hogy a mátrix transzponálásával nem változnak meg az aldeterminánsai, következik, hogy a rang definíciójában oszlopvektorok helyett sorvektorokat is írhatunk. (Vagyis a lineárisan független oszlopvektorok maximális száma mindig pontosan egyenlő a lineárisan független sorvektorok maximális számával.)*

Mindezekből következik, hogy egy mátrixnak nem lehet nagyobb a rangja, mint ahány sora és ahány oszlopa van, azaz $\mathrm{rang}(A) \le \min\{m, n\}$. Egy 2×3-as mátrixnak például legfeljebb 2 lehet a rangja, egy 5×3-asnak legfeljebb 3, és egy $n \times n$-esnek legfeljebb n.

7.2. Lineáris egyenletrendszerek megoldhatósága

Ebben az alfejezetben megfogalmazzuk, hogy mi annak szükséges és elégséges feltétele, hogy egy lineáris egyenletrendszernek létezzen megoldása.

Tekintsük az

$$A\underline{x} = \underline{b},\ A \in \mathbb{R}^{m \times n},\ \underline{x} \in \mathbb{R}^n,\ \underline{b} \in \mathbb{R}^m$$

lineáris egyenletrendszert.

7.4 Definíció. *Az*

$$A_b := \begin{bmatrix} a_{11} & a_{12} & . & a_{1n} & b_1 \\ . & & & & . \\ . & & & & . \\ a_{m1} & . & . & a_{mn} & b_m \end{bmatrix} \tag{10}$$

mátrixot a (10) lineáris egyenletrendszer kibővített mátrixának nevezzük.

Mivel ezt a mátrixot az A mátrixból úgy kaptuk, hogy egy oszloppal kiegészítettük, ezért vagy megegyezik a két mátrix rangja, vagy a kibővített mátrixé eggyel nagyobb, mint az együtthatómátrixé.

A fő tételünk előtt belátunk egy segédtételt.

7.5 Lemma *Ha az $\underline{a}_1, \dots \underline{a}_p$ vektorok maximális elemszámú lineárisan független rendszert alkotnak az $\{\underline{a}_1, \dots \underline{a}_p, \dots, \underline{a}_n\}$ vektorhalmazban, akkor az $\underline{a}_{p+1}, \dots \underline{a}_n$ vektorok mindegyike kifejezhető az $\underline{a}_1, \dots \underline{a}_p$ vektorok lineáris kombinációjaként.*

Biz.: Legyen $j \in \{p+1, \dots, n\}$ egy tetszőleges index. Mivel $\{\underline{a}_1, \dots \underline{a}_p\}$ maximális elemszámú lineárisan független rendszer, ezért az $\{\underline{a}_1, \dots \underline{a}_p, \underline{a}_j\}$ rendszer már lineárisan összefüggő. Ez azt jelenti, hogy léteznek olyan, nem mind nullával egyenlő $\alpha_1, \dots \alpha_p, \alpha_j$ számok, amelyek mellett

$$\alpha_1 \underline{a}_1 + \ldots + \alpha_p \underline{a}_p + \alpha_j \underline{a}_j = \underline{0}.$$

Ekkor azonban az α_j együtthatónak különböznie kell nullától, mert $\alpha_j = 0$ esetén a lineárisan független $\underline{a}_1, \dots \underline{a}_p$ vektorok közül az egyik nemnulla együtthatóval szerepelne a nullértékű lineáris kombinációban, ami ellentmondás. Tehát $\alpha_j \neq 0$, azaz $\underline{a}_j$ kifejezhető a többi vektor lineáris kombinációjaként:

$$\underline{a}_j = -\frac{\alpha_1}{\alpha_j}\underline{a}_1 - \ldots - \frac{\alpha_p}{\alpha_j}\underline{a}_p.$$

7.6 Tétel. *(Kronecker–Capelli-tétel)*
A (10) lineáris egyenletrendszernek pontosan akkor létezik megoldása, ha $\operatorname{rang}(A) = \operatorname{rang}(A_b)$.

Biz.: Jelölje az A együtthatómátrix i-edik oszlopát $\underline{a}_i$. Azt kell belátnunk, hogy az egyenletrendszernek pontosan akkor létezik megoldása, ha az $\{\underline{a}_1, \dots \underline{a}_n\}$ és az $\{\underline{a}_1, \dots \underline{a}_n, \underline{b}\}$ vektorhalmazban azonos a lineárisan független vektorok maximális száma. A bizonyítás két részből áll.
1. ($\Rightarrow$) Tegyük fel, hogy létezik megoldás. Ekkor $\underline{b} = \sum_{i=1}^{n} \lambda_i \underline{a}_i$, ahol a λ_i, $i = 1, 2, \dots, n$ rendszer az egyenletrendszer valamely megoldása. Tegyük fel, hogy az A mátrixban legfeljebb p darab oszlop alkot lineárisan független rendszert, pl. az $\underline{a}_1, \dots \underline{a}_p$ vektorok. Ekkor az előző lemmából következik, hogy A fennmaradó oszlopai mind előállnak az $\underline{a}_1, \dots \underline{a}_p$ vektorok lineáris kombinációjaként. Mivel $\underline{b} = \sum_{i=1}^{n} x_i \underline{a}_i$ alakú, így $\underline{b}$ is előáll az $\underline{a}_1, \dots \underline{a}_p$ vektorok lineáris kombinációjaként. Ez azt jelenti, hogy az $\{\underline{a}_1, \dots \underline{a}_n, \underline{b}\}$ vektorhalmaz lineárisan összefüggő, így nem tartalmaz több lineárisan független vektort, mint az $\{\underline{a}_1, \dots \underline{a}_n\}$ halmaz.
2. ($\Leftarrow$) Tegyük fel, hogy $\mathrm{rang}(A) = \mathrm{rang}(A_b)$. Megmutatjuk, hogy ekkor az egyenletrendszernek létezik megoldása.
Jelölje az $\{\underline{a}_1, \dots \underline{a}_n\}$ és az $\{\underline{a}_1, \dots \underline{a}_n, \underline{b}\}$ vektorhalmazban lineárisan független rendszert alkotó vektorok közös maximális számát p. Ha az $\underline{a}_1, \dots \underline{a}_n$ vektorok között $\{\underline{a}_1, \dots \underline{a}_p\}$ lineárisan független, akkor, mivel az $\{\underline{a}_1, \dots \underline{a}_n, \underline{b}\}$ rendszerben sincs több független vektor, $\underline{b}$ az előző lemma értelmében előáll $\underline{b} = \alpha_1 \underline{a}_1 + \dots + \alpha_p \underline{a}_p$ alakban. Ekkor a

$$\underline{b} = \alpha_1 \underline{a}_1 + \dots + \alpha_p \underline{a}_p + 0 \cdot \alpha_{p+1} + \dots + 0 \cdot \underline{a}_n$$

egyenlőség is fennáll, ami azt jelenti, hogy az $\alpha_1, \dots, \alpha_p, 0, \dots, 0$ számok kielégítik az egyenletrendszert.

A Kronecker–Capelli-tétel csak a megoldás létezéséről szól, a megoldások számáról nem. Vizsgáljuk meg, hogy mitől függ a megoldások száma!
Legyen $\mathrm{rang}(A) = \mathrm{rang}(A_b) = r$. Ekkor az A és az A_b mátrixban egyaránt r darab lineárisan független sor van, ami egyben azt is jelenti, hogy az egyenletrendszerben csak r darab egyenlet független egymástól; a maradék $m - r$ darab egyenlet nem jelent új feltételt, és ilyen értelemben felesleges. Az m darab egyenletből álló (10) rendszer tehát helyettesíthető egy r darab egyenletet tartalmazó

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rn}x_n &= b_r
\end{aligned} \tag{11}$$

alakú egyenletrendszerrel. Elegendő ennek a megoldásait meghatározni.

Nyilvánvalóan r legfeljebb az ismeretlenek számával egyenlő, azaz $r \leq n$. A (11) megoldásainak számát illetően két eset lehetséges:

– Ha $r = n$, akkor a (11) rendszerben ugyanannyi egyenlet van, mint ahány ismeretlen, és az együtthatómátrix determinánsa nemnulla, következésképpen az egyenletrendszernek létezik megoldása, és a megoldás egyértelmű.

– Ha $r < n$, akkor tegyük fel, hogy a nemnulla r-edrendű aldetermináns például az első r darab ismeretlen együtthatóiból áll. Vigyük át a jobb oldalra az $x_{r+1}, \ldots, x_n$ ismeretleneket tartalmazó tagokat, és adjunk az $x_{r+1}, \ldots, x_n$ ismeretleneknek tetszőleges $c_{r+1}, \ldots, c_n$ értékeket. Így az

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r &= b_1 - a_{1,r+1}c_{r+1} - \ldots - a_{1n}c_n \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2r}x_r &= b_2 - a_{2,r+1}c_{r+1} - \ldots - a_{2n}c_n \\
&\vdots \\
a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r &= b_r - a_{r,r+1}c_{r+1} - \ldots a_{rn}c_n
\end{aligned} \tag{12}$$

egyenletrendszerhez jutunk, és ennek már létezik egyetlen $x_1, \ldots, x_r$ megoldása. Ezek a tetszőlegesen megválasztott $c_{r+1}, \ldots, c_n$ paraméterek függvényei, ezért ebben az esetben végtelen sok megoldás van.

Vegyük észre, hogy más eset nem lehetséges, tehát ha megoldható a lineáris egyenletrendszer, akkor vagy egy megoldása van, vagy végtelen sok megoldása van.

A fenti eredményeket a következőkben foglalhatjuk össze.

- Ha $\text{rang}(A) \neq \text{rang}(A_b)$, akkor az egyenletrendszernek nincs megoldása.
- Ha $\text{rang}(A) = \text{rang}(A_b) = n$, azaz a közös rang egyenlő az ismeretlenek számával, akkor az egyenletrendszernek egyértelmű megoldása van.
- Ha $\text{rang}(A) = \text{rang}(A_b) < n$, azaz a közös rang kisebb az ismeretlenek számánál, akkor az egyenletrendszernek végtelen sok megoldása van.

7.3. Homogén lineáris egyenletrendszerek

Tekintsük az

$$A\underline{x} = \underline{0}, \; A \in \mathbb{R}^{m \times n}, \; \underline{x} \in \mathbb{R}^n, \; \underline{b} \in \mathbb{R}^m \tag{13}$$

homogén lineáris egyenletrendszert.

A Kronecker–Capelli-tételből következik, hogy (13) mindig megoldható, mivel a $\underline{0}$ oszlop hozzávétele nem növelheti a mátrix rangját. Könnyen látható az is, hogy az $x_1 = x_2 =$

$\ldots = x_n = 0$ számok mindig kielégítik az egyenletrendszert; ez az egyenletrendszer ún. triviális megoldása.
Legyen az A mátrix rangja r.
Ha $r = n$, akkor az előző fejezet értelmében egyetlen megoldás van, azaz csak a triviális megoldás létezik.
Ha $r < n$, akkor végtelen sok megoldás van, vagyis ekkor léteznek nemtriviális megoldások is.
Speciálisan, n egyenletből álló n ismeretlenes homogén lineáris egyenletrendszernek pontosan akkor létezik nemtriviális megoldása, amikor $\det A = 0$ (mivel ez azt jelenti, hogy $\mathrm{rang}(A) < n$). Ha pedig $m < n$, azaz kevesebb egyenlet van, mint ahány ismeretlen, akkor $\mathrm{rang}(A) < n$ miatt vannak nemtriviális megoldások is.

A homogén lineáris egyenletrendszerek a következő speciális tulajdonságokkal rendelkeznek:

1. Ha a $\lambda_1, \ldots, \lambda_n$ rendszer megoldása (13)-nak, akkor tetszőleges $\alpha \in \mathbb{R}$ esetén $\alpha \cdot \lambda_1, \ldots, \alpha \cdot \lambda_n$ is megoldás.
   Biz.: Jelölje $\underline{\lambda}$ a $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ vektort. $\lambda_1, \ldots, \lambda_n$ megoldása (13)-nak $\Rightarrow A\underline{\lambda} = \underline{0}$. Ekkor tetszőleges $\alpha \in \mathbb{R}$ esetén $A(\alpha\underline{\lambda}) = \alpha(A\underline{\lambda}) = \alpha \cdot \underline{0} = \underline{0}$, azaz $\alpha \cdot \lambda_1, \ldots \alpha \cdot \lambda_n$ is megoldás.

2. Ha $\lambda_1, \ldots, \lambda_n$ és $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n$ is megoldása (13)-nak, akkor $\lambda_1 + \tilde{\lambda}_1, \ldots, \lambda_n + \tilde{\lambda}_n$ is megoldás.
   Biz.: Legyen $\underline{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ és $\underline{\tilde{\lambda}} = (\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_n)$. $\lambda_1, \ldots, \lambda_n$ és $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n$ is megoldása (13)-nak $\Rightarrow A\underline{\lambda} = \underline{0}$ és $A\underline{\tilde{\lambda}} = \underline{0}$. $\Rightarrow A(\underline{\lambda} + \underline{\tilde{\lambda}}) = A\underline{\lambda} + A\underline{\tilde{\lambda}} = \underline{0} + \underline{0} = \underline{0}$.

3. A fentiek következménye, hogy egy homogén lineáris egyenletrendszer megoldásainak tetszőleges lineáris kombinációja is megoldása az egyenletrendszernek.

7.4. Kapcsolat a homogén és az inhomogén lineáris egyenletrendszerek megoldásai között

Tekintsük az

$$A\underline{x} = \underline{b},\ A \in \mathbb{R}^{m \times n},\ \underline{x} \in \mathbb{R}^n,\ \underline{b} \in \mathbb{R}^m \tag{14}$$

lineáris egyenletrendszert. Kapcsolatot keresünk ezen inhomogén rendszer és az A együtthatómátrixú

$$A\underline{x} = \underline{0} \tag{15}$$

homogén rendszer megoldásai között.

7.7 Állítás. *A (15) egy tetszőleges megoldásának és a (14) egy tetszőleges megoldásának összege megoldása (14)-nek.*

Biz.: Legyen $\lambda_1, \ldots, \lambda_n$ megoldása (14)-nek, $\mu_1, \ldots, \mu_n$ pedig (15)-nek. Ekkor a $\underline{\lambda} = (\lambda_1, \ldots, \lambda_n)$ és $\underline{\mu} = (\mu_1, \ldots, \mu_n)$ jelöléssel $A\underline{\lambda} = \underline{b}$ és $A\underline{\mu} = \underline{0}$. $\Rightarrow A(\underline{\lambda} + \underline{\mu}) = A\underline{\lambda} + A\underline{\mu} = \underline{b} + \underline{0} = \underline{b}$, vagyis a $\underline{\lambda} + \underline{\mu}$ vektor elemei, a $\lambda_1 + \mu_1, \ldots \lambda_n + \mu_n$ számok megoldása (14)-nek.

7.8 Állítás. *(14) bármely két megoldásának különbsége megoldása (15)-nek.*

Biz.: Legyenek $\lambda_1, \ldots, \lambda_n$ és $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n$ megoldásai (14)-nek. Ekkor a $\underline{\lambda} = (\lambda_1, \ldots, \lambda_n)$ és $\underline{\tilde{\lambda}} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n)$ jelöléssel $A\underline{\lambda} = \underline{b}$ és $A\underline{\tilde{\lambda}} = \underline{b}$. $\Rightarrow A(\underline{\lambda} - \underline{\tilde{\lambda}}) = A\underline{\lambda} - A\underline{\tilde{\lambda}} = \underline{b} - \underline{b} = \underline{0}$ azaz $\lambda_1 - \tilde{\lambda}_1, \ldots \lambda_n - \tilde{\lambda}_n$ megoldása (15)-nek.

7.9 Következmény. *Az inhomogén lineáris egyenletrendszer összes megoldását megkapjuk, ha egy tetszőleges megoldásához hozzáadjuk a megfelelő homogén rendszer minden egyes megoldását.*

8. Gyakorlati módszerek lineáris egyenletrendszerek megoldására

A reguláris négyzetes mátrixú lineáris egyenletrendszerek egyik megoldási módszeréről már esett szó: ez a Cramer-szabály, amelynek segítségével az ismeretleneket közvetlenül kiszámíthatjuk az egyenletrendszer együtthatóiból és szabad tagjaiból. Ezt a módszert a gyakorlatban szinte sohasem alkalmazzák, mivel az ismeretlenek determinánsok hányadosaiként állnak elő, és a determinánsszámítás – különösen nagyméretű mátrixok esetén – rendkívül műveletigényes. Egy darab $n \times n$-es determináns kifejtésekor ugyanis, ha csak az aldeterminánsok számításigényét nézzük, $n!$ számú szorzást kell elvégeznünk. $n = 50$ esetén ez 50!, azaz több mint 10^{64} darab szorzást jelent. Ennek a számnak az érzékeltetésére végezzünk el egy egyszerű gondolatkísérletet! Képzeljünk el egy olyan számítógépet, amely 10^{-10} méter méretű számolóegységekből áll, és a számolást az elképzelhető legnagyobb sebességgel: fénysebességgel végzi. Egy szorzás ideje legyen az az idő, amely alatt a fény átfut egy számolóegységen: ez $1/3 \cdot 10^{-18}$ másodperc. Akkor 10^{64} darab szorzás

$$t = \frac{1}{3} \cdot 10^{-18} \cdot 10^{64} = \frac{1}{3} \cdot 10^{46} s$$

időt igényel. Ha az eredményt átszámítjuk évekbe, akkor azt kapjuk, hogy ez az idő több mint $\frac{1}{3} \cdot 10^{37}$ év!

Az alkalmazások során (pl. az időjárás-előrejelzésben) jóval nagyobb méretű egyenletrendszerek is előfordulnak. Világos, hogy ezeket a Cramer-szabállyal lehetetlen reális időn belül megoldani. Ezért a gyakorlatban más módszerek terjedtek el. Ezek közül három alapvetővel foglalkozunk: a Gauss-eliminációval, a Gauss–Jordan-eliminációval és a faktorizációs eljárásokkal. Mindezek az ún. direkt módszerek közé tartoznak, ami azt jelenti, hogy pontos számolással a pontos megoldást kapjuk meg. A másik lehetőség valamilyen iterációs módszer alkalmazása: a megoldás közelítését állítjuk elő egy hozzá konvergáló vektorsorozat segítségével.

8.1. Kiküszöbölési (eliminációs) módszerek

A kiküszöbölési eljárások azon az észrevételen alapulnak, hogy egyes speciális egyenletrendszerek könnyen megoldhatók. A legegyszerűbb eset az, amikor az A együtthatómátrix diagonális. Tekintsük például az

$$\begin{aligned} 5x_1 &= 10 \\ 2x_2 &= 4 \\ 3x_3 &= -1 \end{aligned}$$

egyenletrendszert. Ebben az esetben

$$A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad \text{és} \quad b = \begin{bmatrix} 10 \\ 4 \\ -1 \end{bmatrix},$$

és az ismeretlenek az egyenletekből közvetlenül kifejezhetők:

$$\begin{aligned} x_1 &= 2 \\ x_2 &= 2 \\ x_3 &= -\tfrac{1}{3}. \end{aligned}$$

Ha az együtthatómátrix az identitásmátrix, akkor a jobb oldalon a megoldás áll. Könnyű dolgunk van háromszögmátrixú egyenletrendszerek esetén is. Az

$$\begin{array}{rrrcr} x_1 & +2x_2 & +2x_3 & = & 7 \\ & 4x_2 & +5x_3 & = & 17 \\ & & 9x_3 & = & 27 \end{array}$$

egyenletrendszer mátrixa felsőháromszög-mátrix. Az utolsó egyenletből kifejezhető az x_3 ismeretlen: $x_3 = 27/9 = 3$. Ezt behelyettesítjük a második egyenletbe, és kifejezzük x_2-t: $x_2 = 1/2$. Végül x_1 az első egyenletből adódik x_2 és x_3 behelyettesítésével: $x_1 = 7 - 2 \cdot 3 - 2 \cdot \frac{1}{2} = 0$.

Legyen $A\underline{x} = \underline{b}$ négyzetes mátrixú egyenletrendszer, ahol A se nem diagonális, se nem háromszögmátrix. A kiküszöbölési eljárások lényege, hogy a rendszert átalakítjuk háromszögmátrixú vagy diagonálmátrixú rendszerré. Az átalakítás ekvivalens átalakítást jelent, azaz olyat, amely nem változtatja meg az egyenletrendszer megoldását. Így például

1. egyenleteket felcserélhetünk;
2. egyenleteket bármilyen nullától különböző számmal szorozhatunk;
3. bármelyik egyenlethez hosszáadhatjuk egy másik egyenlet tetszőleges számszorosát.

A következőkben két kiküszöbölési eljárást ismertetünk.

8.1.1. Gauss-elimináció

A Gauss-elimináció során az egyenletrendszert az előző részben ismertetett 1., 2. ill. 3. műveletek segítségével felsőháromszög-mátrixú rendszerré transzformáljuk. Ehhez célszerű az együtthatókat és a szabad tagokat táblázatba foglalni, és az átalakításokat ezen a táblázaton nyomon követni, hiszen az ismeretleneket fölösleges minden egyes lépésben kiírnunk. Az elemek nullára transzformálása során felülről lefelé és balról jobbra haladunk. Így elkerülhetjük, hogy már lenullázott elem újra feltöltődjön.

Alkalmazzuk ezt a módszert a

$$\begin{array}{cccc} 2x_1 & +x_2 & +x_3 & = & 4 \\ x_1 & +3x_2 & +2x_3 & = & 6 \\ x_1 & +2x_2 & +2x_3 & = & 5 \end{array}$$

egyenletrendszerre! Készítsük el a együtthatók és a szabad tagok táblázatát:

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 1 & 3 & 2 & 6 \\ 1 & 2 & 2 & 5 \end{array}\right].$$

Lépések:
1. Először a bal felső elem alatti számokat transzformáljuk nullára. Ehhez a 2. ill. a 3. egyenletből kivonjuk az első egyenlet $\frac{1}{2}$-szeresét:

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 5/2 & 3/2 & 4 \\ 0 & 3/2 & 3/2 & 3 \end{array}\right].$$

2. A főátló alatt már csak egy nemnulla elem van: a 3. sor 2. eleme. Ezért a 3. sorból kivonjuk a 2. sor $\frac{3}{5}$-szörösét:

$$\left[\begin{array}{rrr|r} 2 & 1 & 1 & 4 \\ 0 & 5/2 & 3/2 & 4 \\ 0 & 0 & 3/5 & 3/5 \end{array}\right].$$

Tehát a megoldandó egyenletrendszer ekvivalens a

$$\begin{array}{rrrcl} 2x_1 & +x_2 & +x_3 & = & 4 \\ & \frac{5}{2}x_2 & +\frac{3}{2}x_3 & = & 4 \\ & & \frac{3}{5}x_3 & = & \frac{3}{5} \end{array}$$

egyenletrendszerrel. Ebből az ismeretleneket alulról fölfelé kiküszöbölve az

$$x_1 = 1, \quad x_2 = 1, \quad x_3 = 1$$

megoldáshoz jutunk.

Ha egy egyenletrendszernek több megoldása is van, akkor a rendszert megoldani annyit jelent, mint meghatározni az összes megoldását. A következő egyenletrendszernek végtelen sok megoldása van. Határozzuk meg az összes megoldását a Gauss-elimináció technikáját alkalmazva!

$$\begin{array}{rrrcl} x_1 & +x_2 & +3x_3 & = & 1 \\ 2x_1 & +6x_2 & -x_3 & = & 6 \\ 3x_1 & +7x_2 & +2x_3 & = & 7 \end{array}$$

Megoldás:

$$\left[\begin{array}{rrr|r} 1 & 1 & 3 & 1 \\ 2 & 6 & -1 & 6 \\ 3 & 7 & 2 & 7 \end{array}\right] \to^{(2.)-(1.)\cdot 2} \left[\begin{array}{rrr|r} 1 & 1 & 3 & 1 \\ 0 & 4 & -7 & 4 \\ 3 & 7 & 2 & 7 \end{array}\right]$$

$$\to^{(3.)-(1.)\cdot 3} \left[\begin{array}{rrr|r} 1 & 1 & 3 & 1 \\ 0 & 4 & -7 & 4 \\ 0 & 4 & -7 & 4 \end{array}\right] \to^{(3.)-(2.)} \left[\begin{array}{rrr|r} 1 & 1 & 3 & 1 \\ 0 & 4 & -7 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Az utolsó előtti lépésben megjelenő két egyforma sor már mutatja, hogy az egyenletek nem voltak függetlenek egymástól. Ha a harmadik egyenletből kivonjuk a másodikat, egy csupa nulla sort kapunk. Ez a $0x_1 + 0x_2 + 0x_3 = 0$ egyenletnek felel meg, amelynek minden számhármas eleget tesz. Az eredeti egyenletrendszer tehát ekvivalens egy olyannal, amely csak két egyenletből áll. Ennek végtelen sok megoldása van. Ezeket felírhatjuk pl. a következőképpen: legyen $x_3 = p \in \mathbb{R}$ tetszőleges. A maradék ismeretlenek értéke természetesen már függ ettől a p paramétertől: az x_2-t a második egyenletből határozhatjuk meg: $x_2 = 1 + \frac{7}{4}p$, végül az első egyenletből meghatározható az első ismeretlen: $x_1 = -\frac{19}{4}p$. Tehát az egyenletrendszer összes megoldását a következőképpen adhatjuk meg:

$$x_1 = -\frac{19}{4}p,\ x_2 = 1 + \frac{7}{4}p,\ x_3 = p,\ p \in \mathbb{R} \text{ tetszőleges.}$$

Megjegyezzük, hogy a Gauss-elimináció technikája nem csak négyzetes mátrixú egyenletrendszerekre alkalmazható, hanem tetszőleges lineáris egyenletrendszerre. Az átalakítások során arra törekedjünk, hogy az azonos sor-és oszlopindexű elemek alatt nullák legyenek.

8.1.2. Gauss–Jordan-elimináció

Ez a módszer olyan négyzetes mátrixú egyenletrendszerekre alkalmazható, amelyekben az együtthatómátrix determinánsa nem nulla (egyértelmű megoldás van). Ennél az eljárásnál a Gauss-féle kiküszöbölés lépései után a táblázat átalakítását tovább folytatjuk egészen addig, amíg az együtthatómátrix helyén az identitásmátrixot nem kapjuk. Ekkor a szabad tagok helyén leolvasható a megoldás.

Pl. Az előző 1. példában szereplő egyenletrendszert oldjuk meg a Gauss–Jordan-féle eljárással.

Megoldás: A Gauss-elimináció lépései után tovább folytatjuk az egyenletrendszer átalakítását. A lépéseket a következő táblázatok mutatják:

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & \frac{5}{2} & \frac{3}{2} & 4 \\ 0 & 0 & \frac{3}{5} & \frac{3}{5} \end{array}\right] \to^{(3.)\cdot\frac{5}{3},\ (2.)\cdot\frac{2}{5}} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 1 & \frac{3}{5} & \frac{8}{5} \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\to^{(2.)-(3.)\cdot\frac{3}{5}} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \to^{(1.)-(3.)} \left[\begin{array}{ccc|c} 2 & 1 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\to^{(1.)-(2.)} \left[\begin{array}{ccc|c} 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \to^{(1.)/2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Az utolsó résztáblázat jobb oldali oszlopa szerint az összes ismeretlen 1-gyel egyenlő.

8.1.3. Mátrixinvertálás Gauss- és Gauss–Jordan-eliminációval

A Gauss- és a Gauss–Jordan-féle eljárás mátrixinvertálásra is alkalmazható. Legyen $A \in \mathbb{R}^{n \times n}$ olyan négyzetes mátrix, amelyre $\det A \neq 0$ (láttuk, hogy az ilyen mátrixoknak létezik inverzük). Keressük azt az $X \in \mathbb{R}^{n \times n}$ mátrixot, amely kielégíti $AX = I$ mátrixegyenletet. Az egyenlet X megoldása nem más, mint a keresett A^{-1} inverz mátrix. Vezessük be az

$$\underline{x}_i := \begin{pmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{ni} \end{pmatrix} \qquad i = 1, 2, \ldots, n$$

jelölést a keresett mátrix i-edik oszlopvektorára. Az A mátrixszal megszorozva az $\underline{x}_i$ oszlopvektort, az $AX = I$ szorzatmátrix i-edik oszlopát kell megkapnunk, vagyis az $\underline{e}_i \in \mathbb{R}^n$ oszlopvektort. (Itt $\underline{e}_i$ jelöli azt az $\mathbb{R}^n$-beli vektort, amelynek i-edik eleme 1, a többi nulla.) A feladat tehát az n darab

$$A\underline{x}_i = \underline{e}_i \quad i = 1, 2 \ldots, n$$

egyenletrendszer megoldása. Mivel ezek együtthatómátrixa közös, ezért a Gauss- és a Gauss–Jordan-elimináció esetén is mind az n egyenletrendszerben ugyanazokat az átalakításokat végezzük el az A mátrix elemein. Ezért célszerű az n db egyenletrendszert egyszerre megoldani, azaz a táblázatba mind az n db jobb oldalt belefoglalni. A Gauss-elimináció alkalmazásakor a bal oldali $n \times n$-es blokkot felsőháromszög-mátrixúvá alakítjuk, és ezután mind az n jobb oldallal visszahelyettesítünk. A Gauss–Jordan-elimináció esetén a bal oldalt az identitásmátrixszá transzformáljuk, ekkor a megfelelő átalakítások nyomán a jobb oldali blokkban magát az inverz mátrixot kapjuk meg.

Pl. Invertáljuk az

$$A = \begin{bmatrix} 2 & -4 \\ -8 & 6 \end{bmatrix}$$

mátrixot!

Megoldás: 1. Gauss-eliminációval

$$\left[\begin{array}{cc|cc} 2 & -4 & 1 & 0 \\ -8 & 6 & 0 & 1 \end{array}\right] \to^{(2.)+(1.)\cdot 4} \left[\begin{array}{cc|cc} 2 & -4 & 1 & 0 \\ 0 & -10 & 4 & 1 \end{array}\right]$$

A jobb oldali első oszlophoz a következő egyenletrendszer tartozik:

$$\begin{aligned} 2x_{11} - 4x_{21} &= 1 \\ -10x_{21} &= 4 \end{aligned}$$

Ebből $x_{21} = -\frac{2}{5}$ és $x_{11} = -\frac{3}{10}$. A második oszlophoz tartozó egyenletrendszer a következő:

$$\begin{aligned} 2x_{12} - 4x_{22} &= 0 \\ -10x_{22} &= 1 \end{aligned}$$

Ebből $x_{22} = -\frac{1}{10}$ és $x_{12} = -\frac{1}{5}$. Tehát a keresett inverz mátrix:

$$X = A^{-1} = \begin{bmatrix} -\frac{3}{10} & -\frac{1}{5} \\ -\frac{2}{5} & -\frac{1}{10} \end{bmatrix}$$

2. Gauss–Jordan-eliminációval: a fenti lépések után a táblázatot tovább alakítjuk:

$$\rightarrow^{(2.)/(-10)} \left[\begin{array}{cc|cc} 2 & -4 & 1 & 0 \\ 0 & 1 & -\frac{2}{5} & -\frac{1}{10} \end{array}\right] \rightarrow^{(1.)/2} \left[\begin{array}{cc|cc} 1 & -2 & 1/2 & 0 \\ 0 & 1 & -\frac{2}{5} & -\frac{1}{10} \end{array}\right]$$

$$\rightarrow^{(1.)+(2.)\cdot 2} \left[\begin{array}{cc|cc} 1 & 0 & -\frac{3}{10} & -\frac{1}{5} \\ 0 & 1 & -\frac{2}{5} & -\frac{1}{10} \end{array}\right]$$

Látható, hogy az utolsó lépésben a jobb oldali blokkban az előbb kiszámított inverz mátrixot kaptuk.

Néhány kiegészítés a Gauss- és a Gauss–Jordan-eliminációhoz

1. Mindkét eliminációs módszer műveletigénye sokkal kisebb, mint a Cramer-szabályé: egy n-ismeretlenes egyenletrendszernél a Gauss-elimináció során hozzávetőlegesen $n^3/3$, a Gauss–Jordan-eliminációnál $n^3/2$ darab szorzást (vagy osztást) kell elvégezni.

2. Ha az egyenletrendszer megoldása során kerekítünk, az eliminációt érdemes az ún. **főelem-kiválasztás**sal végezni. Láthattuk, hogy a szabályos Gauss-eliminációs algoritmus során az aktuális táblázat i-edik oszlopában általános esetben úgy tudjuk nullára transzformálni az a_{ii} alatti elemeket, hogy az i-edik sort először elosztjuk a_{ii}-vel, majd az így kapott sor megfelelő számszorosát kivonjuk vagy hozzáadjuk a lejjebb lévő sorokhoz. Mivel kis számmal való osztásnál nagy lehet a kerekítési hiba hatása, ezért kedvezőtlen, ha az a_{ii} elem kis abszolút értékű. Ennek elkerülésére szolgál a főelem-kiválasztás.

A **részleges főelem-kiválasztás** során megvizsgáljuk, hogy az adott oszlopban a főátló alatt van-e a főátlóbeli elemnél nagyobb abszolút értékű szám, és ha igen, akkor sorcserével a főátlóba hozzuk.
Alább láthatunk egy példát részleges főelem-kiválasztásra. Mielőtt nullára transzformálnánk a 2. oszlopban lévő elemeket a főátló alatt, felcseréljük a 2. és a 3. sort. Ezzel a főátlóba kerül az oszlop legnagyobb abszolút értékű főátló alatti eleme.

$$\left[\begin{array}{cccc|c} 2 & 3 & 6 & 2 & 7 \\ 0 & \mathbf{1} & 5 & 4 & 1 \\ 0 & \underline{\mathbf{500}} & 5 & 10 & -15 \\ 0 & \mathbf{3} & 2 & 5 & -9 \end{array}\right] \to \left[\begin{array}{cccc|c} 2 & 3 & 6 & 2 & 7 \\ 0 & \underline{500} & 5 & 10 & -15 \\ 0 & 1 & 5 & 4 & 1 \\ 0 & 3 & 2 & 5 & -9 \end{array}\right]$$

Még jobban csökkenthető a számítási hiba a **teljes főelem-kiválasztás**sal. Ilyenkor a táblázatnak a főátlóbeli elemből jobbra és lefelé kiinduló legnagyobb négyzetes blokkjában keressük a legnagyobb abszolút értékű elemet (főelem). Ennek főátlóba hozásához esetleg sor- és oszlopcserét is végre kell hajtanunk. Az utóbbinál arra kell vigyázni, hogy megváltozik az oszlopok számozása. (Pl. ha a 4. oszlopot áthozzuk a 2. oszlopba, akkor onnantól kezdve a 2. oszlop az x_4 ismeretlen együtthatóit fogja tartalmazni!)
Alább láthatunk egy példát teljes főelem-kiválasztásra. Megkeressük a legnagyobb abszolút értékű elemet a vastagon szedett blokkban. Ezt úgy hozhatjuk a főátlóba, hogy felcseréljük a 2. és a 3. sort, majd a 2. és a 4. oszlopot. Mostantól a 2. oszlop tartalmazza az x_4 együtthatóit, és a 4. oszlop az x_2 együtthatóit.

$$\left[\begin{array}{cccc|c} 2 & 3 & 6 & 2 & 7 \\ 0 & \mathbf{1} & \mathbf{5} & \mathbf{4} & 1 \\ 0 & \mathbf{5} & \mathbf{5} & \underline{\mathbf{100}} & -15 \\ 0 & \mathbf{3} & \mathbf{2} & \mathbf{5} & -9 \end{array}\right] \to \left[\begin{array}{cccc|c} 2 & 3 & 6 & 2 & 7 \\ 0 & 5 & 5 & \underline{100} & -15 \\ 0 & 1 & 5 & 4 & 1 \\ 0 & 3 & 2 & 5 & -9 \end{array}\right] \to \left[\begin{array}{cccc|c} 2 & 2 & 6 & 3 & 7 \\ 0 & \underline{100} & 5 & 5 & -15 \\ 0 & 4 & 5 & 1 & 1 \\ 0 & 5 & 2 & 3 & -9 \end{array}\right]$$

Az alábbi példán bemutatjuk a főelem-kiválasztás hatását a megoldás pontosságára. Tegyük fel, hogy Gauss-eliminációval szeretnénk megoldani a

$$\begin{array}{rcrcl} 0,00031x_1 & & +x_2 & = & 3 \\ x_1 & & +x_2 & = & 7 \end{array}$$

egyenletrendszert. (Eláruljuk, hogy ennek pontos megoldása (kerekítve) $x_1 = 4,001$, $x_2 = 2,999$.) Alkalmazzuk a Gauss-eliminációt először főelem-kiválasztása nélkül, és mindig kerekítsünk négy értékes jegyre:

$$\left[\begin{array}{cc|c} 0,00031 & 1 & 3 \\ 1 & 1 & 7 \end{array}\right] \to^{(2.)-(1.)\cdot\frac{1}{0,00031}} \left[\begin{array}{cc|c} 0,00031 & 1 & 3 \\ 0 & -3225 & -9670 \end{array}\right]$$

Ezzel a

$$\begin{array}{rcl} 0,00031x_1 \quad +x_2 & = & 3 \\ 3225x_2 & = & 9670 \end{array}$$

egyenletrendszerhez jutottunk (a második egyenletet megszoroztuk (-1)-gyel). Ennek megoldása négy értékes jegyre kerekítve: $x_1 = 6,452,\ x_2 = 2,998$. Vegyük észre, hogy az első ismeretlen értékét nagyon nagy hibával kaptuk meg!
Ha főelem-kiválasztással számolunk, akkor a

$$\left[\begin{array}{cc|c} 0,00031 & 1 & 3 \\ 1 & 1 & 7 \end{array}\right]$$

táblázatból indulunk ki. Ekkor a megoldás kerekítve $x_1 = 4,002,\ x_2 = 2,998$, azaz most már x_1-et is megfelelő pontossággal kaptuk meg. És ehhez csak fel kellett cserélnünk a két sort.

Pontos számolásnál a főelem-kiválasztásnak természetesen nincs jelentősége.

8.1.4. Felbontási (faktorizációs) eljárások

Faktorizációról akkor beszélünk, ha egy mátrixot két mátrix szorzatára bontunk fel. Tegyük fel, hogy az egyenletrendszer $A \in \mathbb{R}^{n\times n}$ együtthatómátrixa, amelyről ismét feltesszük, hogy reguláris, felírható

$$A = B \cdot C, \quad B, C \in \mathbb{R}^{n\times n}$$

alakban. Ekkor az $A\underline{x} = \underline{b}$ egyenletrendszer másképpen $(B \cdot C)\underline{x} = \underline{b}$, azaz $B(C\underline{x}) = \underline{b}$ alakú. Jelölje a $C\underline{x}$ szorzatot $\underline{y}$. Így az $A\underline{x} = \underline{b}$ egyenletrendszer megoldása ekvivalens a

$$B\underline{y} = \underline{b}$$

és a

$$C\underline{x} = \underline{y}$$

egyenletrendszerek egymás utáni megoldásával.
Nyilvánvalóan ennek csak akkor van értelme, ha a két rendszert külön-külön könnyebb

megoldani, mint az eredeti egyenletrendszert, azaz ha B és C „jó struktúrájú" mátrixok. Ez a helyzet például, ha háromszögmátrixokról van szó. Különösen akkor előnyös faktorizációs módszert alkalmazni, ha ugyanazon együtthatómátrixszal, de különböző jobb oldalakkal is meg akarjuk oldani az egyenletrendszert. Az eljárás legmunkaigényesebb része ugyanis a felbontás (B és C kiszámítása), amit ha egyszer már kiszámítottunk és elraktároztunk, akkor bármikor felhasználhatjuk. (A B és C mátrix ismeretében a fenti két egyenletrendszer már könnyen megoldható.)

A következőkben két konkrét faktorizációs módszert ismertetünk.

1. LU-felbontás. Ennek a felbontási módszernek az a lényege, hogy az A mátrixot $L \cdot U$ alakban írjuk fel, ahol L a főátlóban egyeseket tartalmazó alsóháromszög-mátrix, U pedig felsőháromszög-mátrix (a főátlóban nem feltétlenül egyesekkel). Ekkor a megoldandó egyenletrendszerek:

$$L\underline{y} = \underline{b}$$

és

$$U\underline{x} = \underline{y}.$$

Az L és U meghatározásához jelölje a két mátrix kiszámítandó elemeit l_{ij} ill. u_{ij}. Annak kell teljesülnie, hogy a mátrixszorzást elvégezve az A mátrixot kapjuk, így

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ l_{n1} & l_{n2} & \dots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix}.$$

A bal oldalon elvégezzük a szorzást, és az elemeket egyenlővé tesszük az A mátrix megfelelő elemeivel.

Példa. Oldjuk meg LU-felbontással a

$$\begin{array}{rrrcl} 2x_1 & +x_2 & +x_3 & = & 4 \\ x_1 & +3x_2 & +2x_3 & = & 6 \\ x_1 & +2x_2 & +2x_3 & = & 5 \end{array}$$

egyenletrendszert.

Megoldás: Először meghatározzuk az

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

együtthatómátrix LU-felbontását. Ehhez a

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

egyenlőségnek kell fennállnia. A szorzatmátrix első oszlopát az A első oszlopával összehasonlítva az

$$u_{11} = 2,$$

$$l_{21}u_{11} = 1, \quad \text{azaz} \quad l_{21} = \frac{1}{2},$$

$$l_{31}u_{11} = 1, \quad \text{vagyis} \quad l_{31} = \frac{1}{2}$$

összefüggések adódnak. A második oszlopaik összehasonlításából

$$u_{12} = 1,$$

$$l_{21}u_{12} + u_{22} = 3, \quad \text{azaz} \quad u_{22} = 3 - \frac{1}{2} \cdot 1 = \frac{5}{2},$$

$$l_{31}u_{12} + l_{32}u_{22} = 2, \quad \text{vagyis} \quad l_{32} = \frac{2}{5} \cdot \left(1 - \frac{1}{2} \cdot 1\right) = \frac{3}{5},$$

és végül a harmadik oszlopaik összehasonlításából azonnal adódik, hogy

$$u_{13} = 1,$$

$$l_{21}u_{13} + u_{23} = 2, \quad \text{amiből} \quad u_{23} = \frac{3}{2},$$

$$l_{31}u_{13} + l_{32}u_{23} + u_{33} = 2, \quad \text{vagyis} \quad u_{33} = \frac{3}{5}.$$

A keresett mátrixok tehát

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{3}{5} & 1 \end{bmatrix} \quad \text{és} \quad U = \begin{bmatrix} 2 & 1 & 1 \\ 0 & \frac{5}{2} & \frac{3}{2} \\ 0 & 0 & \frac{3}{5} \end{bmatrix}.$$

Ezért az $L\underline{y} = \underline{b}$ egyenletrendszer a következő:

$$\begin{array}{llllll} y_1 & & & = & 4 \\ \frac{1}{2}y_1 & +y_2 & & = & 6 \\ \frac{1}{2}y_1 & +\frac{3}{5}y_2 & +y_3 & = & 5 \end{array}$$

Az első, a második és a harmadik egyenletből

$$y_1 = 4, \quad y_2 = 4, \quad y_3 = \frac{3}{5}.$$

Végül az $U\underline{x} = \underline{y}$, azaz a

$$\begin{array}{llllll} 2x_1 & +x_2 & +x_3 & = & 4 \\ & +\frac{5}{2}x_2 & +\frac{3}{2}x_3 & = & 4 \\ & & +\frac{3}{5}x_3 & = & \frac{3}{5} \end{array}$$

egyenletrendszer megoldása az utolsó egyenletből kiindulva:

$$x_3 = 1, \quad x_2 = 1, \quad x_1 = 1.$$

2. Cholesky-felbontás. Ez a módszer akkor használható, ha A szimmetrikus, pozitív definit mátrix. (A pozitív definitség azt jelenti, hogy az $(A\underline{x}) \cdot \underline{x}$ skaláris szorzat semmilyen $\underline{x}$ vektorra nem negatív, és csak akkor nulla, ha $\underline{x} = \underline{0}$.) Megmutatható, hogy egy szimmetrikus mátrix pozitív definit, ha az összes bal felső sarokaldeterminánsa pozitív. Ilyenkor A felírható $L \cdot L^T$ alakban, ahol L alsóháromszög-mátrix. Ekkor a megoldandó egyenletrendszerek:

$$L\underline{y} = \underline{b}$$

és

$$L^T\underline{x} = \underline{y}.$$

Pl. Adjuk meg az

$$A = \begin{bmatrix} 9 & 6 & 3 \\ 6 & 5 & 3 \\ 3 & 3 & 6 \end{bmatrix}$$

mátrix Cholesky-felbontását! Oldjuk meg az $A\underline{x} = \underline{b}$ egyenletrendszert, ahol $\underline{b} = (1, 1, 1)^T$.

Megoldás: A

$$\begin{bmatrix} 9 & 6 & 3 \\ 6 & 5 & 3 \\ 3 & 3 & 6 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

mátrixegyenlőségből:

$$l_{11}^2 = 9 \to l_{11} = 3 \quad \text{(-3 is választható)}$$

$$l_{21}l_{11} = 6 \;\Rightarrow l_{21} = 2, \quad l_{31}l_{11} = 3 \;\Rightarrow l_{31} = 1$$

$$l_{21}^2 + l_{22}^2 = 5 \;\Rightarrow l_{22}^2 = 1 \;\Rightarrow l_{22} = 1 \text{ (vagy -1)}$$

$$l_{31}l_{21} + l_{32}l_{22} = 3 \;\Rightarrow l_{32} = 1$$

$$l_{31}^2 + l_{32}^2 + l_{33}^2 = 6 \;\Rightarrow l_{33}^2 = 4 \;\Rightarrow l_{33} = 2 \text{ (vagy -2)}$$

Tehát

$$L = \begin{bmatrix} 3 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{és} \quad L^T = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

1. lépés: Megoldjuk az $L\underline{y} = \underline{b}$ egyenletrendszert:

$$\left.\begin{array}{llll} 3y_1 & & & = \;1 \\ 2y_1 & +y_2 & & = \;1 \\ y_1 & +y_2 & +2y_3 & = \;1 \end{array}\right\} \Rightarrow y_1 = \frac{1}{3},\; y_2 = \frac{1}{3},\; y_3 = \frac{1}{6}$$

2. lépés: Megoldjuk az $L^T\underline{x} = \underline{y}$ egyenletrendszert:

$$\left.\begin{array}{rrrl} 3x_1 & +2x_2 & +x_3 & = \;\frac{1}{3} \\ & x_2 & +x_3 & = \;\frac{1}{3} \\ & & 2x_3 & = \;\frac{1}{6} \end{array}\right\} \Rightarrow x_3 = \frac{1}{12},\; x_2 = \frac{1}{4},\; x_1 = -\frac{1}{12}$$