Nyugat-Magyarországi Egyetem, Közgazdaságtudományi Kar

Horváth Róbert Szalay László

Matematika II

Sopron
2004

Bevezetés

Ezt a jegyzetet a Nyugat-Magyarországi Egyetem Közgazdaságtudományi Karán a 2003/04. tanév második félévében tartott matematika előadásunkhoz készítettük elsőéves nappali ill. levelező képzésben részt vevő gazdálkodás szakos hallgatóknak. Három fő témakör kerül tárgyalásra. Az elsőben az egy- és kétváltozós függvények integrálszámításával foglalkozunk, majd a differenciálegyenletek megoldásának alapvető módszereivel ismerkedhet meg az Olvasó. A félévet a lineáris algebra legfontosabb definícióival és tételeivel zárjuk nagy hangsúlyt helyezve az elemi bázistranszformáció alkalmazására különböző lineáris algebrai problémák megoldásában.

A jegyzet felépítésében a következő szerkesztési elvet követtük. Minden fejezet egy pár soros bevezetővel kezdődik, amely röviden előre vetíti a fejezet tartalmát. Egy fejezeten belül folytonosan számoztuk a definíciókat, tételeket és példákat. A bizonyítások végét ■, a példák végét □, míg a definíciók végét ♢ szimbólumokkal jelöltük. A bonyolultabb tételeket ill. definíciókat néhány egyszerű feladattal igyekeztünk érthetőbbé tenni. Ezeket a feladatokat, melyek részletes megoldása a fejezetek végen megtalálható, kerettel emeltük ki a szövegből. A fejezetek végén ellenőrző kérdésekre válaszolhat az Olvasó. Néhány fejezet végén javaslatokat teszünk önálló gyakorló feladatokra. Ezek a feladatok a *Denkinger Géza, Gyurkó Lajos: Analízis gyakorlatok, Tankönyvkiadó* (DGY) könyvből kerülnek ki.

A szerzők

1. fejezet

A határozatlan integrál

1.1. Primitív függvények és a határozatlan integrál

Az integrálás a deriválás "megfordítása". Integrálás során egy adott függvényhez keressük azon függvényeket, melyek deriváltjai az adott függvényt adják.

1.1.1. Definíció. Azt mondjuk, hogy az $F(x)$ függvény **primitív függvénye** az $f(x)$ függvénynek az $I \subset \mathbb{R}$ nyílt intervallumon, ha $F'(x) = f(x)$ minden $x \in I$ esetén. $\diamondsuit$

A jelölésben célszerű azt az elvet követni, hogy egy függvény primitív függvényét ugyanazon betű nagybetűs változatával jelöljük. Így pl. egy $G(x)$ függvényről sejthetjük majd, hogy annak deriváltja a $g(x)$ függvény lesz.

1.1.1. Feladat. Határozzuk meg az $f(x) = 2x$ és $g(x) = \cos x$ függvények primitív függvényeit!

Egy függvénynek nyilvánvalóan több primitív függvénye van, hiszen ha $F(x)$ primitív függvénye az $f(x)$ függvénynek, akkor $F(x) + c$ is primitív függvénye lesz $f(x)$-nek tetszőleges c konstans esetén, mert $(F(x) + c)' = f(x)$ minden $x \in I$ esetén. Könnyű látni, hogy így meg is kapjuk $f(x)$ összes primitív függvényét, hiszen ha mondjuk $F_1(x)$ és $F_2(x)$ is primitív függvénye $f(x)$-nek, akkor $(F_1(x) - F_2(x))' = f(x) - f(x) = 0$, ami csak úgy lehet, ha $F_1(x) - F_2(x)$ konstans függvény.

1.1.2. Definíció. Az $f(x)$ függvény $I \subset \mathbb{R}$ nyílt intervallumon vett összes primitív függvényének halmazát az $f(x)$ függvény I intervallumon vett **határozatlan integráljának** hívjuk. Jelölésére az $\int f(x)\ \mathrm{d}x$ formát alkalmazzuk, feltüntetve $\mathrm{d}x$ alakban a függvény változóját is. $\diamondsuit$

Ebben a jegyzetben, az egyszerűség kedvéért, az $\int f(x)\ \mathrm{d}x$ szimbólummal nem az $f(x)$ primitív függvényeinek halmazát fogjuk jelölni, hanem a primitív függvények valamelyikét. Tehát pl. az $f(x) = 2x$ függvény határozatlan integrálja

$$\int 2x\ \mathrm{d}x = \{F : \mathbb{R} \to \mathbb{R} \mid F(x) = x^2 + c,\ c \in \mathbb{R}\},$$

ami helyett a rövidebb

$$\int 2x\ \mathrm{d}x = x^2 + c$$

alakot fogjuk használni. Ez a jelölés azt jelenti, hogy a $2x$ függvény primitív függvényei az $x^2 + c$ alakú függvények. Azt az eljárást, amellyel egy függvény határozatlan integrálját határozzuk meg integrálásnak hívjuk.

1.1.2. FELADAT. Határozzuk meg az $f(x) = \cos x$ függvény határozatlan integrálját!

A határozatlan integrál definíciójából azonnal adódik a következő két tulajdonság. Legyen $F(x)$ egy primitív függvénye $f(x)$-nek az $I \subset \mathbb{R}$ nyílt intervallumon. Ekkor

$$\left(\int f(x)\,\mathrm{d}x\right)' = f(x), \qquad \int F'(x)\,\mathrm{d}x = F(x) + c.$$

A deriválási szabályokból következik, hogy ha az $f(x)$ és $g(x)$ függvényeknek van primitív függvényük egy $I \subset \mathbb{R}$ nyílt intervallumon, akkor az $f(x) \pm g(x)$ függvényeknek is van és

$$\int (f(x) \pm g(x))\,\mathrm{d}x = \int f(x)\,\mathrm{d}x \pm \int g(x)\,\mathrm{d}x,$$

valamint tetszőleges $\alpha \in \mathbb{R}$ mellett az $\alpha \cdot f(x)$ függvénynek is van primitív függvénye és

$$\int (\alpha \cdot f(x))\,\mathrm{d}x = \alpha \cdot \int f(x)\,\mathrm{d}x.$$

Az első egyenlőség azt fejezi ki, hogy két függvény összegének primitív függvényei a függvények primitív függvényeinek összegeként állnak elő, a második pedig azt, hogy egy függvény számszorosának primitív függvényei a függvény primitív függvényeinek számszorosai.

Az elemi függvények deriváltjait megadó táblázatot "visszafelé olvasva" máris meg tudjuk határozni néhány függvény határozatlan integrálját. A következő táblázatban ezeket az integrálási szabályokat, az ún. alapintegrálokat, soroljuk fel.

$\int \alpha\,\mathrm{d}x = \alpha x + c, \quad (\alpha \in \mathbb{R});$	$\int x^\alpha\,\mathrm{d}x = \frac{x^{\alpha+1}}{\alpha+1} + c, \quad (\alpha \neq -1);$
$\int \sin x\,\mathrm{d}x = -\cos x + c;$	$\int \cos x\,\mathrm{d}x = \sin x + c;$
$\int \frac{1}{\cos^2 x}\,\mathrm{d}x = \mathrm{tg}x + c;$	$\int \frac{-1}{\sin^2 x}\,\mathrm{d}x = \mathrm{ctg}x + c;$
$\int \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \arcsin x + c;$	$\int \frac{1}{1+x^2}\,\mathrm{d}x = \mathrm{arctg}x + c;$
$\int \frac{1}{x}\,\mathrm{d}x = \ln\lvert x\rvert + c;$	$\int a^x\,\mathrm{d}x = \frac{a^x}{\ln a} + c, \quad (1 \neq a > 0);$
$\int \mathrm{sh}x\,\mathrm{d}x = \mathrm{ch}x + c;$	$\int \mathrm{ch}x\,\mathrm{d}x = \mathrm{sh}x + c;$
$\int \frac{1}{\sqrt{x^2+1}}\,\mathrm{d}x = \mathrm{arsh}x + c;$	$\int \frac{1}{\sqrt{x^2-1}}\,\mathrm{d}x = \mathrm{arch}x + c, \quad x > 1;$
$\int \frac{1}{\mathrm{ch}^2 x}\,\mathrm{d}x = \mathrm{th}x + c;$	$\int \frac{-1}{\mathrm{sh}^2 x}\,\mathrm{d}x = \mathrm{cth}x + c;$
$\int \frac{-1}{\sqrt{1-x^2}}\,\mathrm{d}x = \arccos x + c;$	$\int \frac{-1}{1+x^2}\,\mathrm{d}x = \mathrm{arcctg}x + c;$
$\int \frac{1}{1-x^2}\,\mathrm{d}x = \mathrm{arth}x + c, \quad \lvert x\rvert < 1;$	$\int \frac{1}{1-x^2}\,\mathrm{d}x = \mathrm{arcth}x + c, \quad \lvert x\rvert > 1.$

Az $\int x^\alpha \, \mathrm{d}x = x^{\alpha+1}/(\alpha+1) + c$ formula nyilvánvalóan csak akkor alkalmazható, ha $\alpha \neq -1$. Az $\alpha = -1$ esetben a primitív függvény nem polinomfüggvény lesz, hanem – ahogy a táblázat is mutatja – $\int x^{-1} \, \mathrm{d}x = \int (1/x) \, \mathrm{d}x = \ln|x| + c$.

Az $1/x$ függvénynek nincs primitív függvénye pl. a $(-1, 1)$ intervallumon, mert a primitív függvényt csak egy nyílt intervallumon értelmeztük, és $1/x$ nincs $x = 0$-ban értelmezve. Van viszont primitív függvény tetszőleges $I \subset \mathbb{R}^+$ nyílt intervallumon ($\int (1/x) \, \mathrm{d}x = \ln x + c$), és van primitív függvény tetszőleges $I \subset \mathbb{R}^-$ nyílt intervallumon is ($\int (1/x) \, \mathrm{d}x = \ln(-x) + c$). Ez a megállapítás van összefoglalva a táblázatban az $\int (1/x) \, \mathrm{d}x = \ln|x| + c$ képlettel.

Hasonlóan értelmezhetjük a táblázat utolsó sorát. A $(-\infty, -1)$ és az $(1, \infty)$ intervallumon az $1/(1-x^2)$ függvény primitív függvénye az $\operatorname{arcth} x + c$ függvény, míg a $(-1, 1)$ intervallumon az $\operatorname{arth} x + c$ függvény a primitív függvény.

Többször szükségünk lehet annak megállapítására, hogy egy adott függvény melyik függvény deriváltja, azaz hogy megkeressük a függvény egy primitív függvényét. Későbbi tanulmányainkból megemlíthetjük a határozott integrál fogalmát, amely nemnegatív folytonos függvények esetén a függvénygrafikon és az x-tengely egy szakasza által meghatározott tartomány területét adja meg, és nagyon könnyen számítható, ha ismerjük az adott függvény egy primitív függvényét. Egy másik fontos példa a valószínűségszámítás témakörében fog előjönni. Egy valószínűségi változó eloszlásfüggvénye a sűrűségfüggvényének egy primitív függvénye lesz. További példa a differenciálegyenletek megoldása, melyekről e félévben tanulunk majd részletesebben. Differenciálegyenletnek az olyan egyenleteket hívjuk, amelyek egy ismeretlen függvény és deriváltjai között állapítanak meg valamilyen összefüggést, és amelyekből az ismeretlen függvényt kell meghatároznunk. Lássunk erre egy példát.

Tegyük fel, hogy azon $f(x)$ függvényeket szeretnénk megkeresni, melyekre $f(x) > 0$ ($x \in \mathbb{R}$) és $f(x)$ deriváltfüggvénye megegyezik magával a függvénnyel. Ekkor az $f'(x) = f(x)$ ($x \in \mathbb{R}$) differenciálegyenletet kell megoldanunk. Osszuk el az egyenlet mindkét oldalát $f(x)$-szel:

$$\frac{f'(x)}{f(x)} = 1.$$

Az egyenlet bal oldalának egy primitív függvénye az $\ln(f(x))$ alakban írható, ugyanis

$$(\ln(f(x)))' = \frac{1}{f(x)} f'(x) = \frac{f'(x)}{f(x)}.$$

Emiatt

$$(\ln(f(x)))' = 1,$$

így $\ln(f(x))$ nem lehet más, mint $x + c$, hiszen a deriváltja 1. Ebből kapjuk, hogy $\ln(f(x)) = x + c$, azaz a keresett függvény

$$f(x) = e^{x+c},$$

ahol c tetszőleges valós szám lehet. Ez a függvény valóban pozitív és teljesül rá az $f'(x) = f(x)$ egyenlőség. A megoldás $f(x) = c_1 e^x$ alakban is írható, ahol $c_1 = e^c \in \mathbb{R}^+$.

Látható tehát, hogy fontos megismernünk azokat a módszereket, melyekkel megkereshetjük függvények primitív függvényeit. Ezen módszerekről szól a következő néhány fejezet. A primitív függvény keresés általában nehéz feladat. Csak az vigasztalhat minket, hogy ha már megtaláltuk, akkor könnyen leellenőrizhetjük, hogy jól számoltunk-e, hiszen ehhez csak deriválni kell a függvényt.

1.1.3. FELADAT. Az $f(x) = 1/(1+x^2)$ függvény melyik primitív függvénye halad át az $(1, 2\pi)$ koordinátájú ponton?

1.1.4. FELADAT. Határozzuk meg az $f(x) = (x+1)/\sqrt{x}$ és a $g(x) = \mathrm{tg}^2 x$ függvények határozatlan integrálját!

Ellenőrző kérdések

1. Mely függvényeket hívunk egy adott függvény primitív függvényeinek?
2. Milyen alakúak egy függvény primitív függvényei?
3. Értelmezzük a határozatlan integrál fogalmát!
4. Soroljuk fel a nevezetes integrálási szabályokat!

A feladatok megoldásai

1.1.1.
Ha az $F(x) = x^2$ függvényt deriváljuk, akkor $F'(x) = f(x) = 2x$ adódik, azaz az $F(x) = x^2$ függvény primitív függvénye az $f(x) = 2x$ függvénynek. Vegyük észre, hogy pl. az $F(x) = x^2 + 3$ is jó primitív függvénynek, sőt a 3 helyébe tetszőleges konstanst írhatnánk. Hasonlóan kapjuk, hogy a $G(x) = \sin x$ függvény primitív függvénye a $g(x) = \cos x$ függvénynek, de a $G(x) = \sin x + c$ függvény is megfelel primitív függvénynek (c tetszőleges konstans).

1.1.2.

$$\int \cos x \, \mathrm{d}x = \sin x + c$$

1.1.3.
Az alapintegrálok segítségével könnyen látható, hogy $\int 1/(1+x^2) \, \mathrm{d}x = \mathrm{arctg} x + c$, azaz az $\mathrm{arctg} 1 + c = 2\pi$ egyenlőségnek kell teljesülnie. Ebből a $\pi/4 + c = 2\pi$ egyenlőséget kapjuk, így $c = 7\pi/4$, azaz a keresett primitív függvény $F(x) = \mathrm{arctg} x + 7\pi/4$.

1.1.4.
Mivel egyik függvény sem szerepel a integrálási szabályok táblázatában, ezért az integrálandó függvényeket át kell alakítanunk.

$$\int \frac{x+1}{\sqrt{x}} \, \mathrm{d}x = \int \left(\sqrt{x} + \frac{1}{\sqrt{x}} \right) \, \mathrm{d}x = \int (x^{1/2} + x^{-1/2}) \, \mathrm{d}x = \frac{2}{3} x^{3/2} + 2x^{1/2} + c.$$

$$\int \mathrm{tg}^2 x \, \mathrm{d}x = \int \left(\frac{1}{\cos^2 x} - 1 \right) \, \mathrm{d}x = \mathrm{tg} x - x + c.$$

Önálló munkára kijelölt feladatok: DGY 63. oldal 1-12. feladat.

1.2. Integrálás helyettesítéssel

A helyettesítéses integrálás egy integrálási módszer, amely az összetett függvények deriválási szabályán alapszik.

Az előző félévben tanultuk az összetett függvények deriválási szabályát. Eszerint $(f(g(x)))' = f'(g(x)) \cdot g'(x)$, feltéve hogy a szereplő deriváltak léteznek.

1.2.1. Tétel. *Ha a $g(x)$ függvény differenciálható egy $I \subset \mathbb{R}$ nyílt intervallumon és az $f(x)$ függvény értelmezve van $g(x)$ értékkészletén és ott van primitív függvénye (legyen ez $F(x)$), akkor az $f(g(x)) \cdot g'(x)$ függvénynek is van primitív függvénye I-n és*

$$\int f(g(x)) \cdot g'(x)\,\mathrm{d}x = F(g(x)) + c. \tag{1.2.1}$$

Bizonyítás. Az összetett függvény deriválási szabályát alkalmazva kapjuk, hogy

$$(F(g(x)) + c)' = F'(g(x)) \cdot g'(x) = f(g(x)) \cdot g'(x). \blacksquare$$

1.2.2. Példa. Példaként keressük meg az

$$x \mapsto \frac{e^x}{1 + e^{2x}}$$

függvény határozatlan integrálját. Kisebb átalakítással,

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

látható, hogy az előző tételt alkalmazhatjuk az $f(x) = 1/(1 + x^2)$ és $g(x) = e^x$ választással. Tehát

$$\int \frac{e^x}{1 + e^{2x}}\,\mathrm{d}x = \operatorname{arctg}(e^x) + c.$$

□

1.2.1. Feladat. Határozzuk meg az $x \mapsto 1/(x \cdot \ln x)$ függvény határozatlan integrálját!

1.2.2. Feladat. Határozzuk meg az $x \mapsto \sin(2x)$ függvény határozatlan integrálját!

Az 1.2.1. tétel három speciális esetét külön is érdemes kiemelnünk, mert gyakran találkozunk velük feladatmegoldás során. Az első eset az, amikor $f(x) = 1/x$. Ekkor

$$\int \frac{g'(x)}{g(x)}\,\mathrm{d}x = \ln|g(x)| + c. \tag{1.2.2}$$

A második eset, amikor $f(x) = x^\alpha$ $(\alpha \neq -1)$. Ekkor

$$\int (g(x))^\alpha \cdot g'(x)\,\mathrm{d}x = \frac{(g(x))^{\alpha+1}}{\alpha + 1} + c. \tag{1.2.3}$$

A harmadik esetben, ha $g(x) = ax + b$, $0 \neq a, b \in \mathbb{R}$, akkor

$$\int f(ax + b)\,\mathrm{d}x = \frac{1}{a}F(ax + b) + c. \tag{1.2.4}$$

1.2.3. Feladat. Oldjuk meg az 1.2.1 és 1.2.2 feladatokat újra az (1.2.2) és (1.2.3) képletek alkalmazásával!

1.2.4. Feladat. Határozzuk meg a következő integrálokat!

$$\int e^{3x+2}\,\mathrm{d}x, \qquad \int \sin(5x-1)\,\mathrm{d}x$$

Nagyon hasznos integrálási módszerhez juthatunk abban az esetben, ha az (1.2.1) képletben szereplő g függvény invertálható. Helyettesítsük a képlet mindkét oldalába az $x = \tilde{g}(t)$ ($\tilde{g}$ a g függvény inverze) kifejezést. Ekkor a jobb oldal az

$$F(g(\tilde{g}(t))) + c = F(t) + c = \int f(t)\,\mathrm{d}t$$

alakban írható, a bal oldalon pedig csak jelöljük, hogy kiszámítása után x helyébe az $x = \tilde{g}(t)$ függvényt kell behelyettesíteni. Így az oldalak felcserélése után az

$$\int f(t)\,\mathrm{d}t = \int f(g(x)) \cdot g'(x)\,\mathrm{d}x|_{x=\tilde{g}(t)}$$

formulához jutunk. Eredményünket tétel formájában is megfogalmazzuk. A hagyományok miatt a tételben x és t szerepét felcseréljük.

1.2.3. Tétel. *Ha a $g(t)$ függvény invertálható és differenciálható egy $I \subset \mathbb{R}$ nyílt intervallumon és az $f(x)$ függvény értelmezve van $g(t)$ értékkészletén és ott van primitív függvénye, akkor az $f(g(t)) \cdot g'(t)$ függvénynek is van primitív függvénye I-n és*

$$\int f(x)\,\mathrm{d}x = \int f(g(t)) \cdot g'(t)\,\mathrm{d}t|_{t=\tilde{g}(x)} \tag{1.2.5}$$

Bizonyítás. A bizonyítás a tétel előtti bekezdésben található. ■

Szokás x-et régi változónak és t-t új változónak hívni. A g függvény a régi változót fejezi ki az új változóval, míg $\tilde{g}$ (g inverze) az új változót fejezi ki a régivel. A helyettesítéses integrálás úgy történik, hogy a függvény régi változójának helyébe (x) beírjuk annak az új változóval való előállítását ($x = g(t)$), majd a függvényt beszorozzuk a helyettesítő függvény deriváltjával ($g'(t)$). Ezután a kapott függvényt integráljuk, majd t helyébe visszaírjuk a $t = \tilde{g}(x)$ kifejezést.

1.2.4. Példa. Példaként oldjuk meg az 1.2.2. példát újra:

$$\int \frac{e^x}{1+e^{2x}}\,\mathrm{d}x = ?$$

Helyettesítsük e^x-et t-vel. Az $e^x = t$ egyenlőségből következik, hogy $\tilde{g}(x) = e^x$ és $g(t) = \ln t$ ($g(t)$ természetesen differenciálható és invertálható is). Az 1.2.3. tétel alkalmazásához szükségünk van még a $g'(t) = 1/t$ függvényre.

$$\int \frac{e^x}{1+e^{2x}}\,\mathrm{d}x = \int \frac{t}{1+t^2} \cdot \frac{1}{t}\,\mathrm{d}t|_{t=e^x} =$$

$$= \int \frac{1}{1+t^2}\,\mathrm{d}t|_{t=e^x} = \operatorname{arctg} t + c|_{t=e^x} = \operatorname{arctg}(e^x) + c.$$

□

1.2.5. Példa. Határozzuk meg az

$$\int x^2 \sqrt[3]{1-x}\,\mathrm{d}x$$

integrált a $t = (1-x)^{1/3}$ helyettesítés segítségével. Rögtön láthatjuk, hogy $x = 1-t^3$, azaz $g(t) = 1-t^3$ és $g'(t) = -3t^2$. Először az 1.2.3. tétel szerint az

$$\int f(g(t))\cdot g'(t)\,\mathrm{d}t = \int (1-t^3)^2\cdot t\cdot(-3t^2)\,\mathrm{d}t = -3\cdot\int (t^3-2t^6+t^9)\,\mathrm{d}t = -3\left(\frac{t^4}{4}-\frac{2t^7}{7}+\frac{t^{10}}{10}\right)+c$$

integrált kell kiszámítani, majd t helyébe a $t = (1-x)^{1/3}$ kifejezést írni. Így kapjuk, hogy

$$\int x^2 \sqrt[3]{1-x}\,\mathrm{d}x = -3\left(\frac{(1-x)^{4/3}}{4}-\frac{2(1-x)^{7/3}}{7}+\frac{(1-x)^{10/3}}{10}\right)+c.$$

□

1.2.5. Feladat. Az $\int \sqrt{4-x^2}\,\mathrm{d}x$ határozatlan integrált szeretnénk kiszámítani a $]-2,2[$ intervallumon a helyettesítés módszerével. Melyik helyettesítés alkalmazható? (a) $x = \sin t,\ t \in]-\pi/2,\pi/2[$; (b) $x = 2\sin t,\ t \in]0,\pi/2[$; (c) $x = 2\sin t,\ t \in]-\pi/2,\pi/2[$; (d) $x = 2\cos t,\ t \in]0,\pi/2[$; (e) $x = 2\cos t,\ t \in]\pi,2\pi[$.

Az 1.4. fejezetben részletesebben foglalkozunk majd a trigonometrikus és irracionális függvények integrálásával. Ezen fejezetekben ismertetjük majd, hogy az egyes függvényeknél mikor melyik helyettesítés a célravezető.

Ellenőrző kérdések

1. Melyik deriválási szabályon alapul a helyettesítéses integrálás? Ismertessük ezt a deriválási szabályt!

2. Hogyan kaphatjuk meg az $f^\alpha \cdot f'$ ($\alpha \in \mathbb{R}$) alakú függvények primitív függvényeit?

3. Ismertesse a helyettesítéses integrálás képletét, és a képlet alkalmazhatóságának feltételeit!

A feladatok megoldásai

1.2.1.
Az 1.2.1. tételt alkalmazhatjuk az $f(x) = 1/x$ és $g(x) = \ln x$ választással. Így

$$\int \frac{1}{x\cdot\ln x}\,\mathrm{d}x = \ln|\ln x| + c.$$

1.2.2.
Egy trigonometrikus átalakítás után itt is az 1.2.1. tételt alkalmazhatjuk az $f(x) = x$ és $g(x) = \sin x$ választással.

$$\int \sin(2x)\,\mathrm{d}x = 2\int \sin x\cdot\cos x\,\mathrm{d}x = 2\frac{\sin^2 x}{2}+c = \sin^2 x + c.$$

Egy másfajta átalakítással az $f(x) = \sin x$ és $g(x) = 2x$ a megfelelő választás

$$\int \sin(2x)\,\mathrm{d}x = \frac{1}{2}\int \sin(2x)\cdot 2\,\mathrm{d}x = -\frac{1}{2}\cos(2x) + c.$$

Vegyük észre, hogy a fenti két esetben a $\sin(2x)$ függvény primitív függvényeit két különböző alakban kaptuk meg. Mint tudjuk, ezek csak egy konstansban térhetnek el egymástól, azaz a két primitív függvény különbségének konstanst kell adnia. Egy így is van, mert

$$-\frac{1}{2}\cos(2x) - \sin^2 x = -\frac{1}{2}(\cos^2 x - \sin^2 x) - \sin^2 x = -\frac{1}{2}\cos^2 x - \frac{1}{2}\sin^2 x = -\frac{1}{2}.$$

1.2.3.

$$\int \frac{1}{x\cdot \ln x}\,\mathrm{d}x = \int \frac{\frac{1}{x}}{\ln x}\,\mathrm{d}x = \int \frac{(\ln x)'}{\ln x}\,\mathrm{d}x = \ln|\ln x| + c.$$

$$\int \sin(2x)\,\mathrm{d}x = 2\int (\sin x)^1\cdot\cos x\,\mathrm{d}x = 2\int (\sin x)^1\cdot(\sin x)'\,\mathrm{d}x = \sin^2 x + c.$$

1.2.4.
Alkalmazzuk az (1.2.4) képletet.

$$\int e^{3x+2}\,\mathrm{d}x = \frac{1}{3}e^{3x+2} + c,\ \int \sin(5x-1)\,\mathrm{d}x = -\frac{1}{5}\cos(5x-1) + c.$$

1.2.5.
(a) Itt $x \in]-1,1[$, azaz nem kapjuk vissza a $]-2,2[$ intervallumot, ez a helyettesítés nem alkalmazható. (b) Itt $x \in]0,2[$, azaz itt sem kapjuk vissza a $]-2,2[$ intervallumot, ez a helyettesítés nem alkalmazható. (c) Ez a helyettesítés megfelelő. (d) Itt $x \in]0,2[$, azaz itt sem kapjuk vissza a $]-2,2[$ intervallumot, ez a helyettesítés nem alkalmazható. (e) Ez a helyettesítés alkalmazható.

Önálló munkára kijelölt feladatok: DGY 64. oldal 13-24. feladat, 65. oldal 25-46. feladat, 66. oldal 72-85. feladat

1.3. Parciális integrálás

A parciális integrálás egy integrálási módszer, amely a függvények szorzatának $(f(x)\cdot g(x))' = f'(x)\cdot g(x) + f(x)\cdot g'(x)$ deriválási szabályán alapszik.

Az előző fejezetben láttuk, hogy az összetett függvény deriválási szabályának segítségével hogyan lehet integrálási formulákhoz jutni. Ebben a fejezetben a szorzat deriválási szabályából vezetünk le integrálási módszereket.

1.3.1. Tétel. *Tegyük fel, hogy az $u(x)$ és $v(x)$ függvények differenciálhatóak egy $I \subset \mathbb{R}$ intervallumon, és hogy az $u'(x)\cdot v(x)$ függvénynek van primitív függvénye. Ekkor az $u(x)\cdot v'(x)$ függvénynek is van primitív függvénye és*

$$\int u(x)\cdot v'(x)\,\mathrm{d}x = u(x)\cdot v(x) - \int u'(x)\cdot v(x)\,\mathrm{d}x.$$

Bizonyítás. Azt kell csak megmutatnunk, az $u(x)\cdot v(x) - \int u'(x)\cdot v(x)\,\mathrm{d}x$ alakú függvények deriváltja $u(x)\cdot v'(x)$. Alkalmazzuk az összeg és szorzat deriválási szabályát és a határozatlan integrál definícióját:

$$\left(u(x)\cdot v(x) - \int u'(x)\cdot v(x)\,\mathrm{d}x\right)' = u'(x)\cdot v(x) + u(x)\cdot v'(x) - u'(x)\cdot v(x) = u(x)\cdot v'(x).$$

■

A fenti integrálási módszer segítségével a keresett integrál felírható egy függvény és egy másik integrál különbségeként. Ezt a másik integrált persze még meg kell határozni, de ha jól választjuk meg az $u(x)$ és $v'(x)$ függvényeket, akkor ez az integrál már egyszerűbben kiszámítható, mint az eredeti. Mivel az eredeti függvényt ezzel a módszerrel csak "részben" integráljuk, ezért kapta az eljárás a parciális (azaz magyarul részleges) integrálás elnevezést. A módszert általában szorzatok integráljának kiszámítására használjuk. Most végignézzük azt a három fő esetet, amikor a parciális integrálás célravezető.

Az első eset az, amikor az integrálandó függvény tartalmaz egy $\ln x$, $\mathrm{arctg}x$, $\mathrm{arcctg}x$, $\arcsin x$ vagy $\arccos x$ alakú tényezőt. Ebben az esetben $u(x)$-nek a fent felsorolt függvények egyikét választjuk, $v'(x)$-nek pedig a szorzat többi tényezőjét.

1.3.2. Példa. Határozzuk meg az

$$\int \mathrm{arctg}x\,\mathrm{d}x$$

integrált! Ez az integrál nyilvánvalóan a fenti esethez tartozik, $u(x) = \mathrm{arctg}x$ és $v'(x) = 1$ választással alkalmazható a parciális integrálási módszer ($u'(x) = 1/(1+x^2)$, $v(x) = x$).

$$\int \mathrm{arctg}x\,\mathrm{d}x = \int \mathrm{arctg}x\cdot(x)'\,\mathrm{d}x =$$

$$= \mathrm{arctg}x\cdot x - \int (\mathrm{arctg}x)'\cdot x\,\mathrm{d}x = \mathrm{arctg}x\cdot x - \int \frac{x}{1+x^2}\,\mathrm{d}x = x\cdot\mathrm{arctg}x - \frac{1}{2}\ln(1+x^2) + c.$$

Az utolsó lépésben az integrált lényegében az (1.2.2) képlettel számítottuk ki. □

A második eset az, amikor $p(x)\cdot e^{ax}$, $p(x)\cdot\sin(ax)$ vagy $p(x)\cdot\cos(ax)$ alakú függvényeket kell integrálnunk, ahol $a \in \mathbb{R}\setminus\{0\}$ és $p(x)$ valamilyen polinomja az x változónak. Ekkor $p(x)$-et érdemes $u(x)$-nek választani és a másik tényezőt $v'(x)$-nek. A parciális integrálás

szabályát alkalmazva így újra egy hasonló integrálhoz jutunk, de az abban szereplő polinom fokszáma eggyel kisebb lesz az eredeti fokszámnál. Ha az eljárást annyiszor alkalmazzuk, hogy a polinom fokszáma nullára csökkenjen, akkor már az integrál az alapintegrálok segítségével könnyen meghatározható.

1.3.3. Példa. Határozzuk meg az

$$\int x^2 \cdot e^{-x}\,\mathrm{d}x$$

integrált! Ez az integrál nyilvánvalóan a fenti esethez tartozik, $u(x) = x^2$ és $v'(x) = e^{-x}$ választással alkalmazható a parciális integrálási módszer ($u'(x) = 2x$, $v(x) = -e^{-x}$).

$$\int x^2 \cdot e^{-x}\,\mathrm{d}x = \int x^2 \cdot \left(-e^{-x}\right)'\,\mathrm{d}x = x^2 \cdot (-e^{-x}) - \int 2x \cdot (-e^{-x})\,\mathrm{d}x =$$

$$= -x^2e^{-x} + 2\int x \cdot e^{-x}\,\mathrm{d}x. \tag{1.3.1}$$

A jobb oldalon most egy újra ebbe a kategóriába tartozó integrált kapunk, tehát itt megint az $u(x) = x$ és $v'(x) = e^{-x}$ ($u'(x) = 1$, $v(x) = -e^{-x}$) választás a megfelelő.

$$\int x \cdot e^{-x}\,\mathrm{d}x = \int x \cdot \left(-e^{-x}\right)'\,\mathrm{d}x = x \cdot (-e^{-x}) - \int 1 \cdot (-e^{-x})\,\mathrm{d}x =$$

$$= -xe^{-x} + \int e^{-x}\,\mathrm{d}x = -xe^{-x} - e^{-x}.$$

Eredményünket az (1.3.1) képletbe visszahelyettesítve kapjuk, hogy

$$\int x^2 \cdot e^{-x}\,\mathrm{d}x = -x^2e^{-x} + 2\int x \cdot e^{-x}\,\mathrm{d}x = -x^2e^{-x} + 2\left(-xe^{-x} - e^{-x}\right) =$$

$$= -x^2e^{-x} - 2xe^{-x} - 2e^{-x} + c.$$

Nyilvánvalóan elegendő csak a számolás végén feltüntetni a c konstanst. □

Harmadik esetként a parciális integrálás az $e^{ax}\cdot\sin(bx)$, $e^{ax}\cdot\cos(bx)$, $\sin(\ln x)$ és $\cos(\ln x)$ függvényeknél hasznos még ($a, b \neq 0$). A parciális integrálási szabályt kétszer alkalmazva a keresett integrálok kifejezhetők. Az első két esetben a $v'(x) = e^{ax}$, míg az utóbbi két esetben a $v'(x) = 1$ választás célravezető.

1.3.4. Példa. Határozzuk meg az

$$\int \sin(\ln x)\,\mathrm{d}x$$

integrált! Ez az integrál nyilvánvalóan a fenti esethez tartozik, $u(x) = \sin(\ln x)$ és $v'(x) = 1$ választással alkalmazható a parciális integrálási módszer ($u'(x) = \cos(\ln x)/x$, $v(x) = x$).

$$\int \sin(\ln x)\,\mathrm{d}x = \int \sin(\ln x) \cdot (x)'\,\mathrm{d}x = \sin(\ln x) \cdot x - \int (\sin(\ln x))' \cdot x =$$

$$= \sin(\ln x) \cdot x - \int \cos(\ln x)\,\mathrm{d}x = \sin(\ln x) \cdot x - \int \cos(\ln x) \cdot (x)'\,\mathrm{d}x =$$

$$= \sin(\ln x)\cdot x - \left(\cos(\ln x) \cdot x - \int (-\sin(\ln x))\,\mathrm{d}x\right) = \sin(\ln x)\cdot x - \cos(\ln x)\cdot x - \int \sin(\ln x)\,\mathrm{d}x.$$

Az egyenlőség elején és végén is szerepel a keresett integrál, ahonnét az integrált kifejezhetjük:

$$\int \sin(\ln x)\,\mathrm{d}x = \frac{x}{2}(\sin(\ln x) - \cos(\ln x)) + c.$$

□

Ellenőrző kérdések

1. Mondjuk ki a parciális integrálás szabályát! Milyen feltételek mellett alkalmazható ez a szabály?

2. Melyik deriválási szabályon alapul a parciális integrálás szabálya?

3. Milyen típusú függvények integrálásához jó használni a parciális integrálási módszert? Mit kell ezen esetekben $u(x)$-nek és $v'(x)$-nek választani?

Önálló munkára kijelölt feladatok: DGY 65. oldal 47-64. feladat.

1.4. Nevezetes függvények integrálása

Ebben a fejezetben néhány nevezetes függvény integrálásával ismerkedünk meg. Szó lesz egyszerűbb racionális törtfüggvények, irracionális függvények ill. trigonometrikus függvények integrálásáról.

Polinomfüggvények és egyszerűbb racionális törtfüggvények integrálása

A polinomfüggvények integrálása tagonként elvégezhető, ahogy ezt a következő példa is mutatja.

1.4.1. Példa.

$$\int (x^3 - 2x^2 + 12x - 3)\,\mathrm{d}x = \frac{x^4}{4} - 2\frac{x^3}{3} + 12\frac{x^2}{2} - 3x = \frac{x^4}{4} - \frac{2x^3}{3} + 6x^2 - 3x + c.$$

□

A két polinom hányadosaként előálló függvényeket racionális törtfüggvényeknek hívjuk. A továbbiakban néhány egyszerűbb racionális törtfüggvény integrálásával ismerkedünk meg.

Az (1.2.2) képlet segítségével egyszerűen integrálhatók azok a racionális törtfüggvények, melyek számlálója a nevező deriváltja.

1.4.2. Példa.

$$\int \frac{4x^3 - 8x + 1}{x^4 - 4x^2 + x - 2}\,\mathrm{d}x = \ln|x^4 - 4x^2 + x - 2| + c.$$

□

Az (1.2.2) és (1.2.3) képletek felhasználásával igazolhatjuk, hogy

$$\int \frac{A}{x - a}\,\mathrm{d}x = A \cdot \ln|x - a| + c \tag{1.4.1}$$

és

$$\int \frac{A}{(x - a)^\alpha}\,\mathrm{d}x = \frac{A}{1 - \alpha} \cdot \frac{1}{(x - a)^{\alpha - 1}} + c, \quad (\alpha \neq 1). \tag{1.4.2}$$

1.4.3. Példa.

$$\int \frac{4}{x - 3}\,\mathrm{d}x = 4\ln|x - 3| + c,$$

$$\int \frac{4}{(x - 3)^{10}}\,\mathrm{d}x = \int 4(x - 3)^{-10}\,\mathrm{d}x = 4\frac{(x - 3)^{-9}}{-9} = -\frac{4}{9}\frac{1}{(x - 3)^9} + c.$$

□

Tekintsük most az olyan racionális törtfüggvényeket, melyek számlálójában egy elsőfokú, míg nevezőjében egy másodfokú kifejezés áll. Ezek a függvények általánosan az

$$\frac{Mx + N}{x^2 + Bx + C}$$

alakban írhatók (M, N, B és C tetszőleges valós számok). A nevezőben szereplő polinom főegyütthatóját az általánosság megszorítása nélkül választhatjuk egynek, mert ha nem egy lenne, akkor a számlálót és a nevezőt is végigoszthatjuk ezzel az együtthatóval. A

függvény integrálásánál három esetet különböztetünk meg aszerint, hogy a nevezőnek két különböző valós ($B^2-4C>0$), két egybeeső valós ($B^2-4C=0$) ill. két komplex gyöke van ($B^2-4C<0$).

I. $B^2-4C>0$. Ebben az esetben a nevező az $x^2+Bx+C=(x-a_1)\cdot(x-a_2)$ alakban írható, ahol a_1 és a_2 a nevező két különböző valós gyöke. A másodfokú egyenletek megoldóképlete alapján

$$a_{1,2}=\frac{-B\pm\sqrt{B^2-4C}}{2}.$$

Közös nevezőre hozással ellenőrizhető, hogy

$$\frac{Mx+N}{(x-a_1)\cdot(x-a_2)}=\frac{\frac{Ma_1+N}{a_1-a_2}}{x-a_1}+\frac{\frac{Ma_2+N}{a_2-a_1}}{x-a_2}, \tag{1.4.3}$$

azaz az eredeti függvény felírható két olyan törtfüggvény összegeként, melyek számlálója konstans, a nevezőjükben pedig egy elsőfokú polinom szerepel. A számlálókban szereplő konstansok nagyon egyszerűen meghatározhatók. Az $x-a_1$ nevezőjű tört számlálóját úgy kapjuk, hogy az eredeti függvényben letakarjuk a nevezőben szereplő $(x-a_1)$ tényezőt, és a "maradék képletben" x helyébe a_1-et helyettesítünk. Hasonlóan az $x-a_2$ nevezőjű tört számlálóját úgy kapjuk, hogy az eredeti függvényben letakarjuk a nevezőben szereplő $(x-a_2)$ tényezőt, és a "maradék képletben" x helyébe a_2-t helyettesítünk. A fent ismertetett eljárást általánosan parciális törtekre bontásnak hívják.

Az eredeti racionális törtfüggvény integrálása úgy történik, hogy az (1.4.3) képlet segítségével parciális törtek összegeként állítjuk elő, majd azokat az (1.4.1) képlet segítégével külön-külön integráljuk.

$$\int\frac{Mx+N}{(x-a_1)\cdot(x-a_2)}\,\mathrm{d}x=\frac{Ma_1+N}{a_1-a_2}\int\frac{1}{x-a_1}\,\mathrm{d}x+\frac{Ma_2+N}{a_2-a_1}\int\frac{1}{x-a_2}\,\mathrm{d}x= \tag{1.4.4}$$

$$=\frac{Ma_1+N}{a_1-a_2}\ln|x-a_1|+\frac{Ma_2+N}{a_2-a_1}\ln|x-a_2|+c.$$

II. $B^2-4C=0$. Ebben az esetben a nevező az $x^2+Bx+C=(x-a)^2$ alakban írható, ahol a a nevező kétszeres valós gyöke, azaz $a=-B/2$. Közös nevezőre hozással ellenőrizhető, hogy

$$\frac{Mx+N}{(x-a)^2}=\frac{Ma+N}{(x-a)^2}+\frac{M}{x-a}, \tag{1.4.5}$$

azaz az eredeti függvény felírható két olyan törtfüggvény összegeként, melyek számlálója konstans, a nevezőjükben pedig egy elsőfokú ill. annak négyzete szerepel. A számlálókban szereplő konstansok nagyon egyszerűen meghatározhatók. Az $(x-a)^2$ nevezőjű tört számlálóját úgy kapjuk, hogy az eredeti függvényben számlálójában x helyébe a-t helyettesítünk. Az $x-a$ nevezőjű tört számlálójában pedig M szerepel. Ez az eljárás is parciális törtekre bontás.

Az eredeti racionális törtfüggvény integrálása ezután úgy történik, hogy az (1.4.5) képlet segítségével parciális törtek összegeként állítjuk elő, majd azokat az (1.4.1) és (1.4.2) képletek segítégével külön-külön integráljuk.

$$\int\frac{Mx+N}{(x-a)^2}\,\mathrm{d}x=(Ma+N)\int\frac{1}{(x-a)^2}\,\mathrm{d}x+M\int\frac{1}{x-a}\,\mathrm{d}x= \tag{1.4.6}$$

$$=-(Ma+N)\frac{1}{x-a}+M\ln|x-a|+c.$$

III. $B^2-4C<0$. Ebben az esetben a nevező már nem bontható fel elsőfokú polinomok szorzatára. Az integrálás úgy történik, hogy először a számlálót $(M/2)(2x+B)+N-MB/2$ alakban írjuk fel. Így megjelenik a számlálóban a nevező deriváltja. Ezután a tört két tört összegeként írható fel, melynek első tagja az (1.2.2) képlet segítségével integrálható. A második tag integrálása, nevezőjének teljes négyzetté való alakítása után, az (1.2.4) képlettel történik.

$$\int \frac{Mx+N}{x^2+Bx+C}\,\mathrm{d}x = \int \frac{(M/2)(2x+B)+N-MB/2}{x^2+Bx+C}\,\mathrm{d}x = \tag{1.4.7}$$

$$= \frac{M}{2}\ln|x^2+Bx+C| + \int \frac{N-MB/2}{x^2+Bx+C}\,\mathrm{d}x =$$

$$= \frac{M}{2}\ln|x^2+Bx+C| + \frac{1}{2}\int \frac{2N-MB}{\left(x+\frac{B}{2}\right)^2+C-\frac{B^2}{4}}\,\mathrm{d}x =$$

$$= \frac{M}{2}\ln|x^2+Bx+C| + \frac{1}{2}\int \frac{\frac{2N-MB}{C-B^2/4}}{\left(\frac{x+\frac{B}{2}}{\sqrt{C-B^2/4}}\right)^2+1}\,\mathrm{d}x =$$

$$= \frac{M}{2}\ln|x^2+Bx+C| + \frac{2N-MB}{2\sqrt{C-B^2/4}}\cdot \operatorname{arctg}\left(\frac{x+\frac{B}{2}}{\sqrt{C-B^2/4}}\right)+c.$$

A fenti képleteket nem érdemes megjegyezni. Sokkal hasznosabb az integrálás fő lépéseit megtanulni, és azokat sorra végrehajtani egy feladat megoldása során.

1.4.4. Példa. Határozzuk meg az

$$\int \frac{x+5}{x^2-1}\,\mathrm{d}x$$

integrált! Az integrált az (1.4.4) képlet segítségével számítjuk ki, azaz a törtet először felírjuk parciális törtek összegeként, majd ezeket a törteket külön-külön integráljuk.

$$\int \frac{x+5}{x^2-1}\,\mathrm{d}x = \int \frac{x+5}{(x-1)(x+1)}\,\mathrm{d}x = \int \left(\frac{3}{x-1}+\frac{-2}{x+1}\right)\,\mathrm{d}x =$$

$$= 3\ln|x-1| - 2\ln|x+1| = \ln\frac{|x-1|^3}{|x+1|^2}+c.$$

□

1.4.5. Példa. Határozzuk meg az

$$\int \frac{x+5}{(x-1)^2}\,\mathrm{d}x$$

integrált! Az integrált az (1.4.6) képlet segítségével számítjuk ki, azaz a törtet először felírjuk parciális törtek összegeként, majd ezeket a törteket külön-külön integráljuk.

$$\int \frac{x+5}{(x-1)^2}\,\mathrm{d}x = \int \left(\frac{6}{(x-1)^2}+\frac{1}{x-1}\right)\,\mathrm{d}x = -\frac{6}{x-1}+\ln|x-1|+c.$$

□

1.4.6. Példa. Határozzuk meg az

$$\int \frac{x+5}{x^2+x+\frac{9}{4}}\,\mathrm{d}x$$

integrált! Az integrált az (1.4.7) képlet segítségével számítjuk ki.

$$\int \frac{x+5}{x^2+x+\frac{9}{4}}\,\mathrm{d}x = \int \frac{\frac{1}{2}(2x+1)+\frac{9}{2}}{x^2+x+\frac{9}{4}}\,\mathrm{d}x = \frac{1}{2}\ln\left|x^2+x+\frac{9}{4}\right| + \frac{9}{2}\int \frac{1}{x^2+x+\frac{9}{4}}\,\mathrm{d}x =$$

$$= \frac{1}{2}\ln\left|x^2+x+\frac{9}{4}\right| + \frac{9}{2}\int \frac{1}{(x+\frac{1}{2})^2+2}\,\mathrm{d}x = \frac{1}{2}\ln\left|x^2+x+\frac{9}{4}\right| + \frac{9}{2}\int \frac{\frac{1}{2}}{\left(\frac{x+\frac{1}{2}}{\sqrt{2}}\right)^2+1}\,\mathrm{d}x =$$

$$= \frac{1}{2}\ln\left|x^2+x+\frac{9}{4}\right| + \frac{9\sqrt{2}}{4}\mathrm{arctg}\left(\frac{x+\frac{1}{2}}{\sqrt{2}}\right) + c.$$

□

1.4.1. Feladat. Integráljuk az

$$\frac{x}{x^2-x-6}, \qquad \frac{2x-3}{x^2+2x+2}$$

racionális törtfüggvényeket!

Néhány trigonometrikus függvény integrálása

Az elemi trigonometrikus függvények integráljai az alapintegrálok közé tartoznak. Bonyolultabb esetekben trigonometrikus átalakításokkal az integrált egyszerűbb alakra hozhatjuk, és így végezhetjük el az integrálást. A legfontosabbak a

$$\sin^2 x + \cos^2 x = 1,$$

$$\sin(2x) = 2\cdot\sin x\cdot\cos x$$

és a

$$\cos(2x) = \cos^2 x - \sin^2 x$$

azonosságok és a belőlük következő

$$\cos^2 x = \frac{1+\cos(2x)}{2}, \qquad \sin^2 x = \frac{1-\cos(2x)}{2}$$

összefüggések.

Megjegyezzük, hogy hiperbolikus függvények esetén is nagyon hasonló azonosságok érvényesek:

$$\mathrm{ch}^2 x - \mathrm{sh}^2 x = 1,$$

$$\mathrm{sh}(2x) = 2\cdot\mathrm{sh}x\cdot\mathrm{ch}x,$$

$$\mathrm{ch}(2x) = \mathrm{ch}^2 x + \mathrm{sh}^2 x,$$

$$\mathrm{ch}^2 x = \frac{1+\mathrm{ch}(2x)}{2}, \qquad \mathrm{sh}^2 x = \frac{\mathrm{ch}(2x)-1}{2}.$$

1.4.7. Példa.

$$\int \sin^2 x \, \mathrm{d}x = \int \frac{1-\sin(2x)}{2} \, \mathrm{d}x = \int \frac{1}{2} \, \mathrm{d}x - \int \frac{\sin(2x)}{2} \, \mathrm{d}x = \frac{x}{2} + \frac{\cos(2x)}{4} + c.$$

□

1.4.8. Példa.

$$\int \sin^3 x \, \mathrm{d}x = \int (1-\cos^2 x)\sin x \, \mathrm{d}x = \int \sin x + \int (-\sin x)\cos^2 x \, \mathrm{d}x = -\cos x + \frac{\cos^3 x}{3} + c.$$

□

Trigonometrikus függvényeket tartalmazó integrálok kiszámítására nincs általános szabály. Sok esetben azonban a helyettesítéses integrálás jól alkalmazható. Erre mutatunk most két példát.

1.4.9. Példa. Számítsuk ki az

$$\int \frac{1}{2\sin x - \cos x + 5} \, \mathrm{d}x$$

integrált a $t = \mathrm{tg}(x/2)$ helyettesítéssel! A helyettesítő függvény az $x = g(t) = 2\mathrm{arctg}t$ függvény és $g'(t) = 2/(1+t^2)$. Fejezzük ki $\sin x$ és $\cos x$ értékét a t változó segítségével.

$$\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2} = \frac{2\sin(x/2)}{\frac{1}{\cos(x/2)}} = \frac{\frac{2\sin x/2}{\cos(x/2)}}{\frac{\cos^2(x/2)+\sin^2(x/2)}{\cos^2(x/2)}} = \frac{2t}{1+t^2}.$$

$$\cos x = \cos\left(2\cdot\frac{x}{2}\right) = \cos^2(x/2) - \sin^2(x/2) = \frac{\frac{\cos^2(x/2)-\sin^2(x/2)}{\cos^2(x/2)}}{\frac{\cos^2(x/2)+\sin^2(x/2)}{\cos^2(x/2)}} = \frac{1-t^2}{1+t^2}.$$

Ez után alkalmazzuk a helyettesítéses integrálás szabályát.

$$\int \frac{1}{2\sin x - \cos x + 5} \, \mathrm{d}x = \int \frac{1}{2\frac{2t}{1+t^2} - \frac{1-t^2}{1+t^2} + 5} \cdot \frac{2}{1+t^2} \, \mathrm{d}t =$$

$$= \int \frac{2}{4t - 1 + t^2 + 5 + 5t^2} \, \mathrm{d}t = \int \frac{2}{6t^2+4t+4} \, \mathrm{d}t = \int \frac{\frac{1}{3}}{t^2+\frac{2}{3}t+\frac{2}{3}} \, \mathrm{d}t =$$

$$= \int \frac{\frac{1}{3}}{\left(t+\frac{1}{3}\right)^2 + \frac{5}{9}} \, \mathrm{d}t = \int \frac{\frac{3}{5}}{1+\left(\frac{3t+1}{\sqrt{5}}\right)^2} \, \mathrm{d}t = \frac{1}{\sqrt{5}}\mathrm{arctg}\left(\frac{3t+1}{\sqrt{5}}\right) =$$

$$= \frac{1}{\sqrt{5}}\mathrm{arctg}\left(\frac{3\cdot\mathrm{tg}(x/2)+1}{\sqrt{5}}\right) + c$$

□

1.4.10. Példa. Számítsuk ki az

$$\int \frac{\sin^3 x}{\cos^4 x} \, \mathrm{d}x$$

integrált a $t = \cos x$ helyettesítéssel! Ebben az esetben

$$x = \arccos t, \quad \sin x = \sqrt{1-t^2}, \quad (\arccos t)' = \frac{-1}{\sqrt{1-t^2}}$$

és

$$\int \frac{\sin^3 x}{\cos^4 x}\,\mathrm{d}x = \int \frac{(\sqrt{1-t^2})^3}{t^4}\cdot\frac{-1}{\sqrt{1-t^2}}\,\mathrm{d}t =$$

$$= \int \frac{t^2-1}{t^4}\,\mathrm{d}t = \frac{-1}{t}+\frac{1}{3t^3} = -\frac{1}{\cos x}+\frac{1}{3\cos^3 x}+c.$$

Vegyük észre, hogy a fenti helyettesítéssel x a $[0,\pi]$ intervallumba esik, azaz a formula csak a $]0,\pi[$ intervallumon adja meg a primitív függvényt (természetesen $x \neq \pi/2$). □

Néhány irracionális függvény integrálása

Irracionális függvényeknek hívjuk azokat a függvényeket, melyekben az ismeretlen változó gyökös kifejezései szerepelnek. Nem célunk az összes irracionális függvény integrálására egy egységes módszert adni (nem is lehetne), megelégszünk azzal, hogy csak néhány speciális függvénytípust vizsgálunk meg.

Első példaként tekintsük azokat az integrálokat, melyeknél egy másodfokú polinom gyökét kell integrálni. Ezek általános alakja

$$\int \sqrt{Ax^2+Bx+C}\,\mathrm{d}x.$$

Az integrálás úgy kezdődik, hogy a másodfokú kifejezést teljes négyzetté alakítjuk:

$$Ax^2+Bx+C = A\left(x+\frac{B}{2A}\right)^2 + C - \frac{B^2}{4A},$$

majd a kifejezésből kiemeljük az utolsó tagot

$$\left(C-\frac{B^2}{4A}\right)\cdot\left(\frac{A}{C-\frac{B^2}{4A}}\left(x+\frac{B}{2A}\right)^2+1\right) = \left(\frac{4AC-B^2}{4A}\right)\cdot\left(\frac{4A^2}{4AC-B^2}\left(x+\frac{B}{2A}\right)^2+1\right).$$

Ha ezután bevezetjük a

$$t = \left|\frac{4A^2}{4AC-B^2}\right|\cdot\left(x+\frac{B}{2A}\right)$$

új változót, akkor az eredeti integrál átírható a következő integrálok valamelyikének alakjába

$$\int \sqrt{1-t^2}\,\mathrm{d}t,\quad \int \sqrt{t^2-1}\,\mathrm{d}t,\quad \int \sqrt{1+t^2}\,\mathrm{d}t.$$

Az első integrál $t=\sin u$ vagy $t=\cos u$, a második $t=\mathrm{ch}u$, míg a harmadik integrál $t=\mathrm{sh}u$ helyettesítéssel számítható ki.

1.4.11. Példa. Határozzuk meg az

$$\int \sqrt{1-x^2}\,\mathrm{d}x$$

integrált! Itt egyszerre alkalmazható az $x=\sin t$ ($t=\arcsin x$, $(\sin t)'=\cos t$) helyettesítés. Ekkor

$$\int \sqrt{1-(\sin t)^2}\cdot\cos t\,\mathrm{d}t = \int \cos^2 t\,\mathrm{d}t =$$

$$= \int \frac{1+\cos(2t)}{2}\,\mathrm{d}t = \frac{t}{2}+\frac{\sin(2t)}{4} = \frac{\arcsin x}{2}+\frac{\sin(2\arcsin x)}{4} =$$

$$\frac{\arcsin x}{2}+\frac{x\sqrt{1-x^2}}{2}+c.$$

□

1.4.12. Példa. Határozzuk meg az

$$\int \sqrt{x^2+6x+8}\,\mathrm{d}x$$

integrált! Alakítsuk teljes négyzetté a gyök alatti kifejezést,

$$x^2+6x+8=(x+3)^2-1,$$

majd helyettesítsük az $x+3$ kifejezést $\mathrm{ch}u$-val. Ekkor

$$x=\mathrm{ch}u-3,\quad (\mathrm{ch}u-3)'=\mathrm{sh}u$$

és

$$\int \sqrt{x^2+6x+8}\,\mathrm{d}x=\int \sqrt{(x+3)^2-1}\,\mathrm{d}x=\int \sqrt{\mathrm{ch}^2u-1}\cdot \mathrm{sh}u\,\mathrm{d}u=\int \mathrm{sh}^2u\,\mathrm{d}u=$$

$$=\int \frac{\mathrm{ch}(2u)-1}{2}\,\mathrm{d}u=\frac{\mathrm{sh}(2u)}{4}-\frac{u}{2}=\frac{\mathrm{sh}(2(\mathrm{arch}(x+3)))}{4}-\frac{(\mathrm{arch}(x+3))}{2}+c.$$

□

Az

$$\int \frac{1}{\sqrt{Ax^2+Bx+C}}\,\mathrm{d}x$$

integrálnál a teljes négyzetté alakítás után, szintén a

$$t=\left|\frac{4A^2}{4AC-B^2}\right|\cdot\left(x+\frac{B}{2A}\right)$$

helyettesítéssel és a $\left(C-\frac{B^2}{4A}\right)$ konstanstól eltekintve az

$$\int \frac{1}{\sqrt{1-t^2}}\,\mathrm{d}t,\quad \int \frac{1}{\sqrt{t^2-1}}\,\mathrm{d}t,\quad \int \frac{1}{\sqrt{1+t^2}}\,\mathrm{d}t$$

integrálok valamelyikéhez jutunk. Ezek az integrálok mind elemi integrálok. Az első primitív függvénye $\arcsin t$, a másodiké $\mathrm{arch}t$, míg a harmadiké $\mathrm{arsh}t$.

1.4.13. Példa. Határozzuk meg az

$$\int \frac{1}{\sqrt{x^2+4x-12}}\,\mathrm{d}x$$

integrált! Végezzük el a teljes négyzetté alakítást a nevezőben

$$\int \frac{1}{\sqrt{x^2+4x-12}}\,\mathrm{d}x=\int \frac{1}{\sqrt{(x+2)^2-16}}\,\mathrm{d}x=\int \frac{1}{\sqrt{16((1/16)(x+2)^2-1)}}\,\mathrm{d}x,$$

majd alkalmazzuk a $t=(x+2)/4$ helyettesítést ($x=4t-2$, $(4t-2)'=4$).

$$\frac{1}{4}\int \frac{1}{\sqrt{t^2-1}}\cdot 4\,\mathrm{d}t=\mathrm{arch}\left(\frac{x+2}{4}\right)+c.$$

□

Ha az integrálandó függvény egy elsőfokú polinomból vont gyöknek és a független változó egy polinomjának a szorzata, akkor szintén a helyettesítés módszere célravezető. A gyökös kifejezést kell t-vel helyettesíteni.

1.4.14. Példa. Határozzuk meg az

$$\int x\sqrt{x+1}\,\mathrm{d}x$$

integrált! Ebben az esetben a $t=\sqrt{x+1}$ helyettesítés alkalmazható. Innét

$$x=t^2-1,\quad \left(t^2-1\right)'=2t.$$

A következő módon hajthatjuk végre az integrálást.

$$\int x\sqrt{x+1}\,\mathrm{d}x=\int(t^2-1)t\cdot 2t\,\mathrm{d}t=2\int(t^4-t^2)\,\mathrm{d}t=$$

$$=\frac{2}{5}t^5-\frac{2}{3}t^3=\frac{2}{5}(\sqrt{x+1})^5-\frac{2}{3}(\sqrt{x+1})^3+c.$$

□

Ellenőrző kérdések

1. Hogyan integráljuk a polinomfüggvényeket?
2. Mely függvényeket nevezzük racionális törtfüggvénynek?
3. Milyen típusú racionális törtfüggvények integrálását tanultuk?
4. Miért hasznos a parciális törtekre bontás?
5. Ismertessük a tanult racionális törtfüggvények integrálásának főbb lépéseit!
6. Soroljuk fel a legfontosabb trigonometrikus azonosságokat!
7. Hogyan integrálhatjuk a $\sin^2 x$ ill. $\cos^2 x$ függvényeket?
8. Mely függvényeket nevezzük irracionális függvényeknek?
9. Hogyan kell elvégezni egy másodfokú kifejezés teljes négyzetté alakítását?
10. Milyen helyettesítéssel kell integrálni egy másodfokú függvény gyökét ill. ennek reciprokát?

A feladat megoldása

1.4.1.
Mivel $x^2-x-6=(x+2)(x-3)$, ezért a következő módon számolhatunk.

$$\int\frac{x}{(x+2)(x-3)}\,\mathrm{d}x=\int\frac{2}{5(x+2)}\,\mathrm{d}x+\int\frac{3}{5(x-3)}\,\mathrm{d}x=\frac{2}{5}\ln|x+2|+\frac{3}{5}\ln|x-3|+c.$$

A második függvény nevezőjének nincs valós gyöke.

$$\int\frac{2x-3}{x^2+2x+2}\,\mathrm{d}x=\int\frac{(2x+2)-5}{x^2+2x+2}\,\mathrm{d}x=\ln|x^2+2x+2|-5\int\frac{1}{(x+1)^2+1}\,\mathrm{d}x=$$

$$=\ln|x^2+2x+2|-5\,\mathrm{arctg}(x+1)+c.$$

Önálló munkára kijelölt feladatok: DGY 68. oldal 106, 107, 110, 113, 115, 120, 122, 123, 125. feladat, 69. oldal 143-146, 148-150. feladat.

Határozatlan integrállal kapcsolatos vegyes feladatok: DGY 69. oldal 128, 131-139, 154. feladat.

2. fejezet

A határozott integrál

2.1. Egyváltozós függvények integrálhatósága és a határozott integrál

A határozott integrált az az igény hívta életre, hogy ki tudjuk számolni a függvények grafikonja és az x-tengely egy intervalluma közé eső síkidomok területét. Ebben a fejezetben értelmezzük a határozott integrál fogalmát.

Tekintsünk egy $[a, b]$ $(a < b)$ intervallumot, és osszuk fel n részre az

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$$

osztópontokkal. Ekkor az $\{x_0, \ldots, x_n\}$ halmazt az $[a, b]$ intervallum egy felosztásának nevezzük és jelölésére a $P(x_0, \ldots, x_n)$ szimbólumot, vagy ha félreértést nem okoz egyszerűen csak P-t használunk. Az $[x_{i-1}, x_i]$ intervallumot $(i = 1, \ldots, n)$ i-edik osztóintervallumnak nevezzük. Jelölje Δx_i az i-edik osztóintervallum hosszát, azaz az $x_i - x_{i-1}$ számot. Az osztóintervallumok hosszának maximumát Δ-val jelöljük és a felosztás finomságának hívjuk. Azt mondjuk, hogy egy felosztás finomabb egy $\delta > 0$ számnál, ha $\Delta < \delta$, azaz ha a felosztás finomsága δ-nál kisebb.

2.1.1. Feladat. Határozzuk meg a $P = \{0, 1.2, 2.15, 3, 4.35, 5\}$ felosztás finomságát!

Tekintsünk most egy $f(x)$ függvényt, amely értelmezett az $[a, b]$ intervallumon, és vegyük $[a, b]$-nek egy n részre való P felosztását. Válasszunk ki mindegyik osztóintervallumból egy-egy $x_i^\star$ értéket $(i = 1, \ldots, n)$ és képezzük az

$$I(f, P, P^\star) = \sum_{i=1}^{n} f(x_i^\star) \cdot \Delta x_i$$

ún. integrálközelítő összeget. A jelölésben az I az integrál szó kezdőbetűjéből származik, a zárójelben lévő mennyiségek pedig azt mutatják, hogy ez az f függvénytől, az $[a, b]$ intervallum felosztásától és a felosztás osztóintervallumaiból kiválasztott tetszőleges $x_i^\star$ számoktól függ, melyek halmazát $P^\star$-gal jelöltük.

2.1.2. Feladat. Tekintsük az $f(x) = x^2$ függvényt a $[0, 1]$ intervallumon. Határozzuk meg a függvény integrálközelítő összegét $P = \{0, 0.5, 1.2, 1.7, 2\}$ és $P^\star = \{0.2, 1, 1.3, 1.9\}$ esetén.

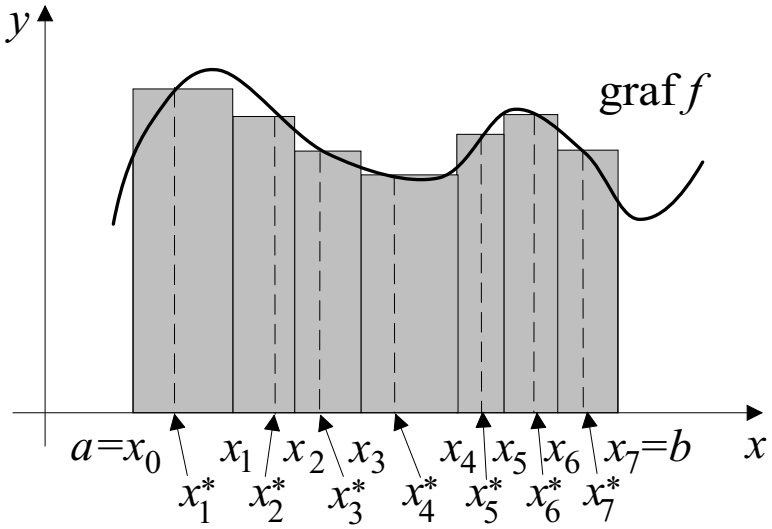
2.1.1. Definíció. Azt mondjuk, hogy az $I(f, P, P^\star)$ **integrálközelítő összegek tartanak az** $I \in \mathbb{R}$ **számhoz** $\Delta \to 0$ esetén, ha minden $\varepsilon > 0$ számhoz létezik olyan $\delta > 0$ szám, hogy bármely δ-nál finomabb P felosztásra, és bármely $x_i^\star \in [x_{i-1}, x_i]$ választás esetén $|I(f, P, P^\star) - I| < \varepsilon$. Jelölésben: $\lim_{\Delta \to 0} I(f, P, P^\star) = I$. $\diamondsuit$

2.1.2. Definíció. Ha az $[a, b]$ zárt intervallumon értelmezett $f(x)$ függvény integrálközelítő összegeinek van határértéke $\Delta \to 0$ esetén, és a határérték az $I \in \mathbb{R}$ szám, akkor azt mondjuk, hogy az $f(x)$ **függvény integrálható** az $[a, b]$ intervallumon és a **(határozott) integrálja** I. Jelölésben:

$$I = \int_a^b f(x)\,\mathrm{d}x.$$

$\diamondsuit$

Vegyük észre a jelölésbeli különbséget: határozatlan integrál esetén az $\int f(x)\,\mathrm{d}x$ az $f(x)$ függvény *primitív függvényeit* jelenti, míg az $\int_a^b f(x)\,\mathrm{d}x$ *szám* az $f(x)$ függvény határozott integrálját adja meg az $[a, b]$ intervallumon. Az a és b számokat integrálási határoknak is nevezzük.



2.1.1. ábra. *Az integrálközelítő összeg szemléltetése.*

Az integrálközelítő összeg és a határozott integrál geometriai jelentését a következőképpen adhatjuk meg. Tekintsünk egy nemnegatív függvényt egy $[a, b]$ intervallumon (lásd 2.1.1. ábra), majd osszuk fel az $[a, b]$ intervallumot osztóintervallumokra (az ábrán hat osztóponttal hét osztóintervallumra osztottuk az eredeti intervallumot). Az integrálközelítő összeg $f(x_i^\star) \cdot \Delta x_i$ tagja egy Δx_i alapú és $f(x_i^\star)$ magasságú téglalap területét adja meg, így az integrálközelítő összeg az ábrán besötétített lépcsőzetes alakzat területét jelenti. Emiatt, ha $\Delta \to 0$ esetén az integrálközelítő összegeknek van határértékük, azaz a függvény integrálható az $[a, b]$ intervallumon, akkor a határozott integrál értékével definiálhatjuk a függvénygrafikon és az x-tengely $[a, b]$ intervalluma közti tartomány területét (azaz az $x = a$, $x = b$, $y = 0$ és $y = f(x)$ görbék által határolt alakzat területét). Hasonló interpretációhoz jutunk ha nem csak nemnegatív függvényeket vizsgálunk, azzal a kiegészítéssel, hogy az integrál a függvénygrafikon és az x-tengely közé eső tartomány

területét negatív előjellel veszi figyelembe, ha a tartomány az x-tengely alatt helyezkedik el.

2.1.3. Tétel. *Ha egy $f(x)$ függvény integrálható egy $[a,b]$ intervallumon, akkor ott korlátosnak kell lennie.*

Az előző tétel szerint, csak korlátos függvény lehet integrálható egy $[a,b]$ intervallumon, azaz a korlátossság szükséges feltétele az integrálhatóságnak. A feltétel nem elégséges. Jól mutatja ezt az

$$f(x) = \begin{cases} 1, & \text{ha } x \text{ racionális,} \\ 0, & \text{ha } x \text{ irracionális} \end{cases}$$

ún. Dirichlet függvény pl. a $[0,1]$ intervallumon, amely nyilvánvalóan egy korlátos függvény, de mégsem integrálható. Vegyük ugyanis a $[0,1]$ intervallum egy tetszőleges felosztását és képezzük úgy az integrálközelítő összeget, hogy mindegyik intervallumból racionális számot választunk $x_i^\star$-nak. Ekkor

$$I(f,P,P^\star) = \sum_{i=1}^{n} f(x_i^\star)\cdot\Delta x_i = \sum_{i=1}^{n} 1\cdot\Delta x_i = \sum_{i=1}^{n}\Delta x_i = 1.$$

Ha viszont mindegyik intervallumból irracionális számot választunk, akkor

$$I(f,P,P^\star) = \sum_{i=1}^{n} f(x_i^\star)\cdot\Delta x_i = \sum_{i=1}^{n} 0\cdot\Delta x_i = 0.$$

A fentiek miatt $\Delta\to 0$ esetén az integrálközelítő összegeknek nincs határértéke, azaz a függvény nem integrálható a $[0,1]$ intervallumon.

Most elégséges feltételeket sorolunk fel egy függvény integrálhatóságára.

2.1.4. Tétel. *Ha az $f(x)$ függvény folytonos az $[a,b]$ intervallumon, akkor azon integrálható is.*

A tétel következménye, hogy minden elemi függvény (hatványfüggvények, polinomfüggvények, exponenciális függvények, logaritmus függvények, trigonometrikus függvények, stb.), ezek összege, különbsége, szorzata és hányadosa is integrálható minden olyan zárt intervallumon, amely teljesen a függvény értelmezési tartományához tartozik.

2.1.5. Definíció. Egy $[a,b]$ intervallumon értelmezett függvényt **szakaszonként folytonos függvénynek** hívunk, ha a függvény csak véges sok pontban nem folytonos, de ezekben a pontokban is van jobb- és baloldali határértéke. $\diamondsuit$

2.1.6. Tétel. *Ha az $f(x)$ függvény szakaszonként folytonos az $[a,b]$ intervallumon, akkor azon integrálható is.*

2.1.7. Tétel. *Ha az $f(x)$ függvény monoton az $[a,b]$ intervallumon, akkor azon integrálható is.*

2.1.3. Feladat. Határozzuk meg, hogy integrálhatóak-e a megadott függvények az adott intervallumon.

$$f(x) = \begin{cases} 0, & \text{ha } x = 0, \\ 1/x, & \text{ha } x \neq 0 \end{cases}, \ [-1,1]; \quad g(x) = 1/x, [1,2]; \quad h(x) = \operatorname{tg} x, \ \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$$

2.1.8. Példa. Határozzuk meg az $f(x) = 1 + x$ függvény integrálját a $[-1, 4]$ intervallumon! Először is a 2.1.4. tétel vagy a 2.1.7. tétel miatt a függvény biztosan integrálható. A függvény grafikonja egy egyenes, amelynek 1 a meredeksége és az y-tengelyt az $y = 1$ pontban metszi. Mivel az integrál jelen esetben az $f(x) = 1 + x$ függvény grafikonja és a $[-1, 4]$ intervallum közé eső területet adja meg, ami egy háromszög, emiatt az integrál értéke 25/2.

Határozzuk meg az integrált a definíció, azaz az integrálközelítő összegek segítségével. Osszuk fel a $[-1, 4]$ intervallumot n egyenlő részre. Ekkor az i-edik osztópont az

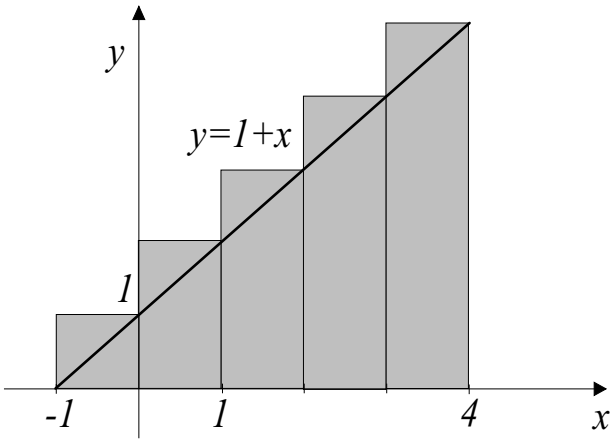
$$x_i = -1 + \frac{5i}{n}, \quad i = 0, \ldots, n$$

alakban írható és $\Delta x_i = 5/n$. Válasszuk $x_i^\star$-nak minden osztóintervallum esetén az intervallum jobb végpontját, és számítsuk ki az integrálközelítő összeget.

$$I(f, P, P^\star) = \sum_{i=1}^{n} f(x_i^\star) \cdot \Delta x_i = \sum_{i=1}^{n} \left(1 + \left(-1 + \frac{5i}{n}\right)\right) \cdot \frac{5}{n} =$$

$$\frac{25}{n^2} \sum_{i=1}^{n} i = \frac{25}{n^2} \cdot \frac{n(n+1)}{2} = \frac{25n(n+1)}{2n^2} = \frac{25(n+1)}{2n}.$$

A 2.1.2. ábrán a $[-1, 4]$ intervallumot 5 egyenlő részre osztottuk, azaz $n = 5$. A fenti képlet alapján az ehhez a felosztáshoz tartozó integrálközelítő összeg 15, amely az ábrán besötétített téglalapok területösszegét adja meg.



2.1.2. ábra. *Az $f(x) = 1 + x$ függvény egy integrálközelítő összege.*

Ha $\Delta \to 0$ (azaz $n \to \infty$), akkor az integrálközelítő összeg 25/2-hez tart. Mivel a függvény integrálható, így az integrálközelítő összegeknek van határértéke és az nem lehet más, mint 25/2. Eredményül tehát az

$$\int_{-1}^{4} (1 + x)\, \mathrm{d}x = \frac{25}{2}$$

egyenlőséget kaptuk. □

2.1.9. Példa. Határozzuk meg az $\int_0^1 x^2\, \mathrm{d}x$ határozott integrál értékét! A 2.1.4. tétel vagy a 2.1.7. tétel miatt ez az integrál valóban létezik, mert az $f(x) = x^2$ függvény folytonos a $[0, 1]$ intervallumon. Ebben a példában nem tudjuk megmondani a függvény

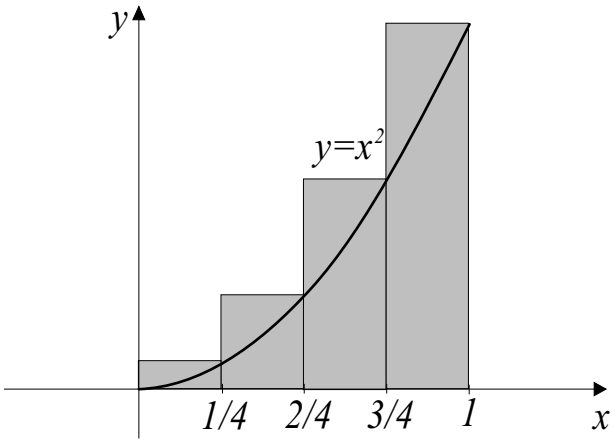
grafikonja alatti tartomány területét, ezt a területet éppen az integrál értéke fogja definiálni.

Osszuk fel a $[0,1]$ intervallumot n egyenlő részre. Ekkor az i-edik osztópont az $x_i = i/n,\ i = 0,\dots,n$ alakban írható és $\Delta x_i = 1/n$. Válasszuk $x_i^\star$-nak minden osztóintervallum esetén az intervallum jobb végpontját, és számítsuk ki az integrálközelítő összeget.

$$I(f,P,P^\star) = \sum_{i=1}^{n} f(x_i^\star)\cdot\Delta x_i = \sum_{i=1}^{n}\left(\frac{i}{n}\right)^2\cdot\frac{1}{n} =$$

$$\frac{1}{n^3}\sum_{i=1}^{n} i^2 = \frac{1}{n^3}\cdot\frac{n(n+1)(2n+1)}{6}.$$

Itt felhasználtuk, hogy az első n természetes szám négyzetösszege $n(n+1)(2n+1)/6$ alakban írható. A 2.1.3. ábrán a $[0,1]$ intervallumot 4 egyenlő részre osztottuk, azaz $n = 4$. Az ehhez a felosztáshoz tartozó integrálközelítő összeg $30/4^3 = 0.469$, amely az ábrán besötétített téglalapok területösszegét adja meg.



2.1.3. ábra. *Az $f(x) = x^2$ függvény egy integrálközelítő összege.*

Ha $\Delta \to 0$ (azaz $n \to \infty$), akkor az integrálközelítő összeg 1/3-hoz tart. Mivel a függvény integrálható, így az integrálközelítő összegeknek van határértéke és az nem lehet más, mint 1/3. Eredményül tehát az

$$\int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}$$

egyenlőséget kaptuk. □

Ellenőrző kérdések

1. Mit nevezünk egy $[a,b]$ intervallum felosztásának? Hogyan jelöljük?
2. Hogyan értelmezzük egy függvény integrálközelítő összegeit egy $[a,b]$ intervallumon?
3. Hogyan értelmezzük az integrálközelítő összegek határértékét, ha a felosztás finomsága nullához tart?
4. Hogyan értelmezzük a határozott integrált?
5. Mondjunk szükséges feltételt az integrálhatóságra!

6. Mondjunk elégséges feltételeket az integrálhatóságra!

7. Mondjunk legalább két olyan függvényt, amely nem integrálható a $[0, 1]$ intervallumon!

A feladatok megoldásai

2.1.1.
Az adott felosztás öt osztóintervallumra osztja a $[0, 5]$ intervallumot. Ezek hossza rendre $1.2, 0.95, 0.85, 1.35$ és 0.62. Tehát a finomság 1.35.

2.1.2.
Az adott felosztás négy osztóintervallumra osztja a $[0, 2]$ intervallumot.

$$\begin{array}{lll} \Delta x_1 = 0.5, & f(x_1^\star) = 0.2^2 = 0.04, & f(x_1^\star) \cdot \Delta x_1 = 0.02 \\ \Delta x_2 = 0.7, & f(x_2^\star) = 1^2 = 1, & f(x_2^\star) \cdot \Delta x_2 = 0.7 \\ \Delta x_3 = 0.5, & f(x_3^\star) = 1.3^2 = 1.69, & f(x_3^\star) \cdot \Delta x_3 = 0.845 \\ \Delta x_4 = 0.3, & f(x_4^\star) = 1.9^2 = 3.61, & f(x_4^\star) \cdot \Delta x_4 = 1.083, \end{array}$$

így integrálközelítő összegnek $0.02 + 0.7 + 0.845 + 1.083 = 2.648$ adódik.

2.1.3.
Az $f(x)$ függvény nem integrálható, mert nem korlátos az adott intervallumon. $g(x)$ integrálható, mert folytonos sőt monoton (csökkenő) is. Végül $h(x)$ integrálható, mert folytonos és monoton (növő) is az adott intervallumon.

2.2. A határozott integrál tulajdonságai

Ebben a fejezetben a határozott integrál legfontosabb tulajdonságait foglaljuk össze.

2.2.1. Tétel. *Ha az $f(x)$ és $g(x)$ függvények integrálhatóak az $[a,b]$ intervallumon és α tetszőleges valós szám, akkor $f(x)+g(x)$ és $\alpha\cdot f(x)$ is integrálható az $[a,b]$ intervallumon és*

$$\int_a^b (f(x)+g(x))\ \mathrm{d}x = \int_a^b f(x)\ \mathrm{d}x + \int_a^b g(x)\ \mathrm{d}x, \qquad \int_a^b (\alpha\cdot f(x))\ \mathrm{d}x = \alpha\int_a^b f(x)\ \mathrm{d}x.$$

Az integrál fenti tulajdonságait röviden úgy fejezzük ki, hogy azt mondjuk, hogy az integrál *lineáris*.

2.2.2. Tétel. *Ha az $f(x)$ és $g(x)$ függvények integrálhatóak az $[a,b]$ intervallumon, akkor $f(x)\cdot g(x)$ is integrálható az $[a,b]$ intervallumon.*

A tétel csak arról szól, hogy két integrálható függvény szorzata integrálható lesz, de nem mond semmit arról, hogy hogyan számítható ki ez az integrál. Könnyen látható, hogy nincs is olyan képlet, amely két függvény integráljának ismeretében megmondaná a szorzat integrálját. Tekintsük ugyanis az $f(x)=1$, $g(x)=2x$ és $h(x)=3x^2$ függvényeket. Mindhárom függvény integrálja 1 a $[0,1]$ intervallumon, azonban

$$\int_0^1 (1\cdot 2x)\ \mathrm{d}x = 1 \neq \int_0^1 (2x\cdot 3x^2)\ \mathrm{d}x = \frac{3}{2}.$$

2.2.3. Tétel. *Ha az $f(x)$ függvény integrálható az $[a,b]$ intervallumon, akkor bármely $[c,d]\subset[a,b]$ intervallumon is integrálható.*

2.2.4. Tétel. *Legyen az $f(x)$ függvény integrálható az $[a,c]$ intervallumon és a $[c,b]$ intervallumon is. Ekkor $f(x)$ integrálható az $[a,b]$ intervallumon is és*

$$\int_a^b f(x)\ \mathrm{d}x = \int_a^c f(x)\ \mathrm{d}x + \int_c^b f(x)\ \mathrm{d}x.$$

A tételben nem kötöttük ki, hogy a c pontnak a és b közé kell esnie, vagy azt, hogy nem eshet egybe a-val, vagy b-vel. Emiatt szükség van a következő definíciókra

$$\int_a^a f(x)\ \mathrm{d}x = 0, \qquad \int_b^a f(x)\ \mathrm{d}x = -\int_a^b f(x)\ \mathrm{d}x.$$

2.2.5. Tétel. *Legyen az $f(x)$ és $g(x)$ függvény integrálható az $[a,b]$ intervallumon és legyen $f(x)\le g(x)$ minden $x\in[a,b]$ esetén. Ekkor*

$$\int_a^b f(x)\ \mathrm{d}x \le \int_a^b g(x)\ \mathrm{d}x.$$

Ellenőrző kérdések

1. Soroljuk fel a határozott integrál tulajdonságait!
2. Melyik tulajdonság jelenti a határozott integrál linearitását?
3. Igaz-e, hogy ha két függvény integrálható, akkor a szorzatuk is?
4. Igaz-e, hogy ha két függvény összege integrálható, akkor a két függvény külön-külön is integrálható? Állításunkat példával igazoljuk!
5. Hogy szól a 2.2.5. tétel abban az esetben, ha $f(x)=0$ az $[a,b]$ intervallumon?

2.3. A Newton-Leibniz formula

Ebben a fejezetben azzal foglalkozunk, hogy a határozatlan integrál segítségével hogyan számíthatjuk ki a határozott integrált ill. hogy a határozott integrál segítségével hogyan tudjuk függvények határozatlan integrálját előállítani. Igazoljuk a Newton-Leibniz formulát, mely szerint egy függvény határozott integrálja kiszámítható a függvény egy primitív függvényének megváltozásaként.

A határozott és határozatlan integrál közti kapcsolatot a Newton-Leibniz formula teremti meg.

2.3.1. Tétel. *(Newton-Leibniz formula) Legyen az $f(x)$ függvény integrálható az $[a,b]$ intervallumon és tegyük fel, hogy van egy olyan $F(x)$ függvény, amely folytonos az $[a,b]$ intervallumon és az $]a,b[$ intervallumon $F(x)$ primitív függvénye $f(x)$-nek. Ekkor*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a).$$

Bizonyítás. Osszuk fel az $[a,b]$ intervallumot n intervallumra az

$$a = x_0 < x_1 < \ldots < x_n = b$$

osztópontok segítségével. Egy tetszőleges $[x_{i-1}, x_i]$ $(i = 1, \ldots, n)$ osztóintervallumon az $F(x)$ függvény folytonos és az intervallum belsejében differenciálható, ezért a Lagrange-tétel miatt létezik olyan $x_i^\star$ az $[x_{i-1}, x_i]$ intervallum belsejében, melyre

$$F'(x_i^\star) = f(x_i^\star) = \frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}}.$$

A fenti $x_i^\star$ értékekkel képezzük az $I(f, P, P^\star)$ integrálközelítő összeget:

$$I(f,P,P^\star) = \sum_{i=1}^n f(x_i^\star)\cdot\Delta x_i = \sum_{i=1}^n f(x_i^\star)(x_i - x_{i-1}) = \sum_{i=1}^n \frac{F(x_i)-F(x_{i-1})}{x_i - x_{i-1}}(x_i - x_{i-1}) =$$

$$= \sum_{i=1}^n (F(x_i) - F(x_{i-1})) = F(b) - F(a).$$

Mivel az $f(x)$ függvény integrálható $[a,b]$-n, ezért az integrálközelítő összegeknek van határértéke és az nem lehet más, mint $F(b) - F(a)$. ■

A Newton-Leibniz formula segítségével tehát határozatlan integrál segítségével lehet határozott integrálokat számítani. Az integrál értéke nem lesz más mint a függvény primitív függvényének megváltozása az adott intervallumon. Az $F(b) - F(a)$ megváltozásra szokás az $[F(x)]_a^b$ jelölést használni.

2.3.2. Példa. Számítsuk ki az $\int_{-1}^4 (1+x)\,\mathrm{d}x$ integrált! Ezt az integrált már kiszámítottuk a 2.1.8. példában. Most a Newton-Leibniz formulával oldjuk meg a feladatot. Az $f(x) = 1 + x$ függvény egy primitív függvénye az $F(x) = x + x^2/2$, azaz a tétel feltételei nyilvánvalóan teljesülnek.

$$\int_{-1}^4 (1+x)\,\mathrm{d}x = \left[x + \frac{x^2}{2}\right]_{-1}^4 = \left(4 + \frac{4^2}{2}\right) - \left(-1 + \frac{(-1)^2}{2}\right) = \frac{25}{2}.$$

□

2.3.3. Példa. Számítsuk ki az $\int_0^1 x^2 \, \mathrm{d}x$ integrált! Ezt az integrált már kiszámítottuk a 2.1.9. példában. Most a Newton-Leibniz formulával oldjuk meg a feladatot. Az $f(x) = x^2$ függvény egy primitív függvénye az $F(x) = x^3/3$, azaz a tétel feltételei nyilvánvalóan teljesülnek.

$$\int_0^1 x^2 \, \mathrm{d}x = \left[\frac{x^3}{3}\right]_0^1 = \left(\frac{1^3}{3}\right) - \left(\frac{0^3}{3}\right) = \frac{1}{3}.$$

□

A következőkben azt vizsgáljuk, hogy a határozatlan integrál hogyan állítható elő a függvények határozott integrálja segítségével. Tegyük fel, hogy az $f(x)$ függvény integrálható az $[a, b]$ intervallumon. Válasszunk ki egy tetszőleges x_0 pontot ebből az intervallumból és értelmezzük az $F(x)$ függvényt a következő módon

$$F(x) = \int_{x_0}^{x} f(t) \, \mathrm{d}t, \qquad (D_F = [a, b]).$$

Az $F(x)$ függvényt tehát egy integrál definiálja, melynek felső integrálási határa a függvény változója. Emiatt az $f(x)$ függvény változóját x helyett már egy másik változóval, t-vel jelöltük. Szokás $F(x)$-et $f(x)$ integrálfüggvényének is hívni. Nem véletlen, hogy az integrálfüggvényt is $F(x)$-szel jelöltük, hasonlóan a primitív függvényekhez, mert látni fogjuk, hogy az integrálfüggvény speciális esetekben primitív függvény is lesz egyben. A szereplő integrál valóban létezik tetszőleges $x \in [a, b]$ esetén, hiszen $[x, x_0]$ vagy $[x_0, x]$ részhalmaza $[a, b]$-nek és $[a, b]$-n az $f(x)$ függvény integrálható.

2.3.4. Tétel. *Tegyük fel, hogy az $f(x)$ függvény integrálható az $[a, b]$ intervallumon és $x_0 \in [a, b]$ tetszőleges szám. Ekkor az $[a, b]$ zárt intervallumon értelmezett*

$$F(x) = \int_{x_0}^{x} f(t) \, \mathrm{d}t \tag{2.3.1}$$

függvény folytonos az $[a, b]$ intervallumon és $F'(x) = f(x)$ minden olyan pontban az $]a, b[$ nyílt intervallumon, ahol az $f(x)$ függvény folytonos.

A tétel közvetlen következménye, hogy folytonos függvényeknek mindig van primitív függvényük és a primitív függvények a (2.3.1) képlettel állíthatók elő (más-más x_0 esetén más-más primitív függvényhez jutunk).

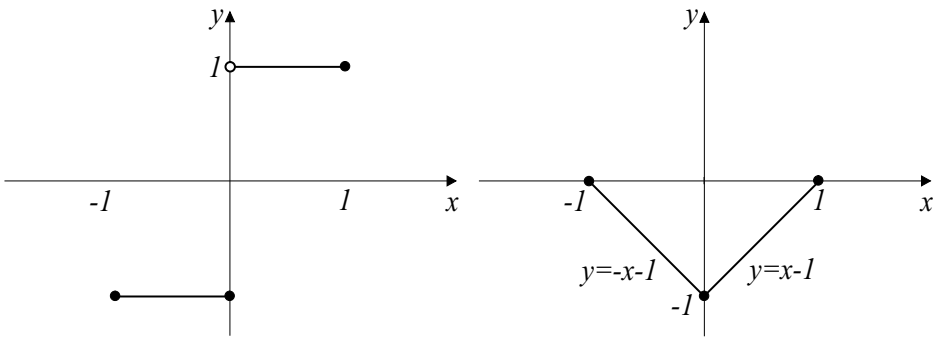
Az $F(x)$ függvény geometriai jelentése az, hogy tetszőleges x esetén megadja az x és x_0 által meghatározott intervallum és az $f(x)$ függvény grafikonja közti előjeles területet (x-tengely alatti tartományokat negatív, míg az x-tengely felettieket pozitív előjellel kell venni).

2.3.5. Példa. Határozzuk meg az

$$f(x) = \begin{cases} \text{-1}, & \text{ha } x \in [-1, 0], \\ 1, & \text{ha } x \in \,]0, 1] \end{cases}$$

függvény integrálfüggvényét $x_0 = -1$ választással! Az adott függvény valóban integrálható a $[-1, 1]$ intervallumon, hiszen szakaszonként folytonos. Emiatt az integrálfüggvény

$$F(x) = \int_{-1}^{x} f(t) \, \mathrm{d}t =$$



2.3.1. ábra. *Az $f(x)$ függvény és integrálfüggvénye.*

$$= \begin{cases} \int_{-1}^{x}(-1)\,\mathrm{d}t = [-t]_{-1}^{x} = -x-1, & \text{ha } x \in [-1,0] \\ \int_{-1}^{0} f(t)\,\mathrm{d}t + \int_{0}^{x} f(t)\,\mathrm{d}t = -1 + \int_{0}^{x} 1\,\mathrm{d}t = -1 + [t]_{0}^{x} = -1+x, & \text{ha } x \in [0,1]. \end{cases}$$

Az eredeti függvényt és integrálfüggvényét a 2.3.1. ábrán szemléltettük. Vegyük észre, hogy annak ellenére, hogy $f(x)$ nem folytonos, integrálfüggvénye már egy folytonos függvény. Az integrálfüggvény az $x = 0$ pont (itt $f(x)$ nem folytonos) kivételével differenciálható a $]-1,1[$ intervallum minden pontjában. □

A legtöbb esetben a Newton-Leibniz formula segítségével a határozott integrál meghatározható. Vannak azonban olyan esetek, amikor ez amiatt nem megy, mert a függvény primitív függvényét nem tudjuk zárt alakban felírni. A zárt alakban való felírás azt jelenti, hogy a függvény előállítható elemi függvények segítségével az alapműveletek (összeadás, kivonás, szorzás és osztás) véges sokszori alkalmazásával. Egy fontos példa az $f(x) = e^{-x^2}$ függvény, amely a standard normális eloszlás sűrűségfügvényében fog szerepelni. Ennek a függvénynek nyilvánvalóan van primitív függvénye, hiszen a függvény folytonos. Azonban a primitív függvény nem írható fel zárt alakban, csak hatványsor formájában, így a Newton-Leibniz formula alkalmazása közben végtelen sorok összegét kellene kiszámolnunk, ami a legtöbb esetben nagyon nehéz feladat. Emiatt a következő eljárást alkalmazzák: $f(x)$ primitív függvénye a 2.3.4. tétel miatt biztosan létezik és előállítható az

$$F(x) = \int_{x_0}^{x} e^{-t^2}\,\mathrm{d}t$$

integrállal. Az x_0 értéket lerögzítve az integrál tetszőleges x értékek esetén közelítő módszerekkel meghatározható, így a primitív függvény helyettesítési értékeit közelíteni tudjuk.

Ellenőrző kérdések

1. Mit mond ki a Newton-Leibniz formula és milyen feltételek mellett érvényes?
2. Mit nevezünk egy függvény integrálfüggvényének egy $[a,b]$ intervallumon?
3. Mik az integrálfüggvény tulajdonságai?
4. Mondjunk egy elégséges feltételt primitív függvény létezésére!

Önálló munkára kijelölt feladatok: DGY 60. oldal 5-16. feladat.

2.4. A helyettesítéses és parciális integrálás határozott integrálokra

Ebben a fejezetben azt vizsgáljuk meg, hogy hogyan lehet a helyettesítéses integrálást és a parciális integrálást egyszerűen alkalmazni a határozott integrálra.

A helyettesítéses integrálás módszerét a határozatlan integrál témakörében tanultuk. Idézzük fel a módszer: ha a $g(t)$ függvény invertálható és differenciálható egy $I \subset \mathbb{R}$ nyílt intervallumon és az $f(x)$ függvény értelmezve van $g(t)$ értékkészletén és ott van primitív függvénye, akkor az $f(g(t)) \cdot g'(t)$ függvénynek is van primitív függvénye I-n és

$$\int f(x)\,\mathrm{d}x = \int f(g(t)) \cdot g'(t)\,\mathrm{d}t|_{t=\tilde{g}(x)}.$$

Ha a fent szereplő függvényektől egy kicsit többet követelenünk meg, azaz $f(x)$-ről feltesszük, hogy folytonos (ebből következik, hogy van primitív függvénye) és $g(x)$-ről feltesszük, hogy még a deriváltja is folytonos, akkor a Newton-Leibniz formula közvetlenül alkalmazható az $\int_a^b f(x)\,\mathrm{d}x$ határozott integrálra.

$$\int_a^b f(x)\,\mathrm{d}x = \left[\int f(g(t)) \cdot g'(t)\,\mathrm{d}t|_{t=\tilde{g}(x)}\right]_a^b =$$

$$= \int f(g(t)) \cdot g'(t)\,\mathrm{d}t|_{t=\tilde{g}(b)} - \int f(g(t)) \cdot g'(t)\,\mathrm{d}t|_{t=\tilde{g}(a)} = \int_{\tilde{g}(a)}^{\tilde{g}(b)} f(g(t)) \cdot g'(t)\,\mathrm{d}t.$$

A fenti képlet azt mutatja, hogy nem szükséges az $\int_a^b f(x)\,\mathrm{d}x$ integrál kiszámításánál közvetlenül meghatározni az $f(x)$ függvény primitív függvényét, hogy arra alkalmazzuk a Newton-Leibniz formulát. Ehelyett elegendő az $f(g(t)) \cdot g'(t)$ függvényt integrálni a $[\tilde{g}(a), \tilde{g}(b)]$ intervallumon. Figyeljük meg, hogy a helyettesítéses integrálás során három lépést kell elvégezni. Először az $f(x)$ függvényben az x változó helyett $g(t)$-t kell írnunk. Másodszor az integrálandó függvényt be kell szoroznunk $g'(t)$-vel, azaz a helyettesítő függvény deriváltjával. Harmadszor pedig meg kell változtatnunk az integrálási határokat $[a, b]$-ről $[\tilde{g}(a), \tilde{g}(b)]$-re.

Egy másik fontos integrálási módszer a parciális integrálás volt: ha az $u(x)$ és $v(x)$ függvények differenciálhatóak egy $I \subset \mathbb{R}$ intervallumon, és az $u'(x) \cdot v(x)$ függvénynek van primitív függvénye, akkor az $u(x) \cdot v'(x)$ függvénynek is van primitív függvénye és

$$\int u(x) \cdot v'(x)\,\mathrm{d}x = u(x) \cdot v(x) - \int u'(x) \cdot v(x)\,\mathrm{d}x.$$

Ha még feltesszük azt is, hogy az $u(x)$ függvény deriváltja is folytonos, akkor a Newton-Leibniz formula szerint $u(x) \cdot v'(x)$ határozott integrálja a következő módon határozható meg.

$$\int_a^b u(x) \cdot v'(x)\,\mathrm{d}x = \left[u(x) \cdot v(x) - \int u'(x) \cdot v(x)\,\mathrm{d}x\right]_a^b =$$

$$= [u(x) \cdot v(x)]_a^b - \left[\int u'(x) \cdot v(x)\,\mathrm{d}x\right]_a^b = [u(x) \cdot v(x)]_a^b - \int_a^b u'(x) \cdot v(x)\,\mathrm{d}x,$$

hiszen egy függvény primitív függvényének a megváltozása az $[a, b]$ intervallumon éppen a függvény ezen intervallumon vett határozott integráljával egyezik meg.

Önálló munkára kijelölt feladatok: DGY 61. oldal 17-28. feladat.

2.5. Az improprius integrál

Ebben a fejezetben olyan függvények integráljával foglalkozunk, amelyek vagy nem korlátosak, vagy végtelen hosszúságú intervallumon kell őket integrálni.

A határozott integrált csak olyan függvények esetén definiáltuk, amelyek egy korlátos zárt intervallumon voltak értelmezve, valamint láttuk, hogy csak korlátos függvények lehetnek integrálhatóak. Ebben a fejezetben kiterjesztjük a korábban tanult integrálfogalmat nemkorlátos függvényekre ill. végtelen hosszúságú intervallumon való integrálás esetére. Az integrál ezen kiterjesztett változatát improprius (helytelen, nem megfelelő) integrálnak nevezik. Az improprius integrált a határozott integrálok határértékeként állítjuk elő.

2.5.1. Példa. Tekintsük pl. az

$$\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x$$

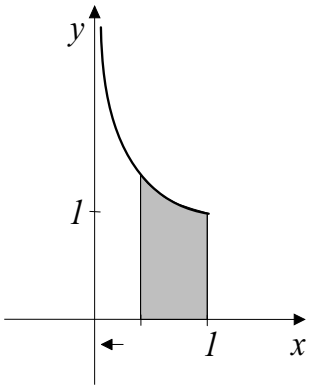
integrált. Itt egy olyan függvényt kell integrálnunk a $[0,1]$ intervallumon, amely csak a $]0,1]$ intervallumon értelmezett, ráadásul a függvény nem is korlátos ezen az intervallumon. Vegyük észre, hogy az $1/\sqrt{x}$ függvény tetszőleges $\omega > 0$ esetén integrálható az $[\omega, 1]$ intervallumon (ezen az intervallumon már korlátos a függvény). Az integrál értéke

$$\int_\omega^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = \Big[2\sqrt{x}\Big]_\omega^1 = 2 - 2\sqrt{\omega}.$$

Ha most ω értékével tartunk nullához, akkor $2-2\sqrt{\omega}$ tart kettőhöz, így a keresett integrál értékét célszerű 2-nek tekinteni, azaz

$$\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = 2.$$

Azt mondjuk ilyenkor, hogy a fenti improprius integrál létezik, és értéke 2. Tekintsük



2.5.1. ábra. *Az $f(x) = 1/\sqrt{x}$ függvény a $[0,1]$ intervallumon.*

a 2.5.1. ábrát, amely az $1/\sqrt{x}$ függvényt mutatja a $[0,1]$ intervallumon. Tudjuk, hogy az ω-tól 1-ig való integrálás a besötétített tartomány területét adja meg. Ha most ω nullához tart, akkor ez a terület kettőhöz tart, így az $1/\sqrt{x}$ függvény grafikonja és a $[0,1]$ intervallum közti tartomány területét is célszerű 2-nek tekinteni. Vegyük észre, hogy ez a tartomány kiterjedését tekintve végtelen nagy (hiszen az y-tengely irányában tetszőlegesen nagy értéket felvehet a függvény), mégis véges nagyságú területe van. □

2.5.2. Példa. Másik példaként tekintsük az

$$\int_1^\infty \frac{1}{x^2}\,\mathrm{d}x$$

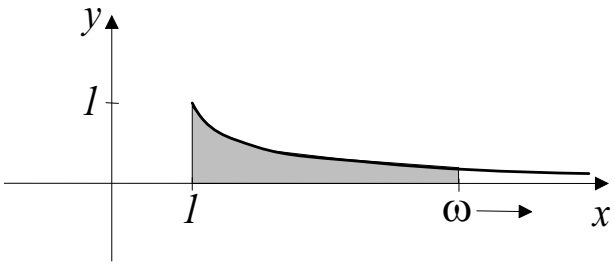
integrált. Először is azt vesszük észre, hogy ez az integrál nem a hagyományos határozott integrál, hiszen itt az $[1,\infty[$ intervallumon integráljuk a függvényt. Az $1/x^2$ függvény tetszőleges $\omega > 1$ esetén integrálható az $[1,\omega]$ intervallumon (ez már véges hosszúságú, zárt intervallum). Az integrál értéke

$$\int_1^\omega \frac{1}{x^2}\,\mathrm{d}x = \left[\frac{-1}{x}\right]_1^\omega = \frac{-1}{\omega} + 1 = 1 - \frac{1}{\omega}.$$

Ha most ω értékével tartunk végtelenhez, akkor $1-1/\omega$ tart egyhez, így a keresett integrál értékét célszerű 1-nek tekinteni, azaz

$$\int_1^\infty \frac{1}{x^2}\,\mathrm{d}x = 1.$$

Tehát a fenti improprius integrál létezik, és értéke 1. Az 1-től ω-ig való integrálás



2.5.2. ábra. *Az $f(x) = 1/x^2$ függvény az $[1,\infty[$ intervallumon.*

a 2.5.2. ábrán a besötétített tartomány területét adja meg. Ha most ω végtelenhez tart, akkor ez a terület egyhez tart, így az $1/x^2$ függvény grafikonja és az $[1,\infty[$ intervallum közti tartomány területét is célszerű 1-nek tekinteni. Ez a tartomány is végtelen kiterjedésű, de területe véges érték. □

2.5.3. Definíció. Tekintsünk egy $f(x)$ függvényt, mely integrálható tetszőleges $\omega < b$ ($a < b \in \mathbb{R}$ vagy $b = \infty$) esetén az $[a,\omega]$ intervallumon. Amenyiben az $\int_a^\omega f(x)\,\mathrm{d}x$ integrálnak van véges határértéke, ha ω tart b-hez, akkor azt mondjuk, hogy az $f(x)$ függvény **improprius integrálja** létezik az $[a,b]$ intervallumon és az integrál értékét $\int_a^b f(x)\,\mathrm{d}x$ fogja jelölni. Azaz

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{\omega\to b}\int_a^\omega f(x)\,\mathrm{d}x.$$

♢

2.5.4. Definíció. Tekintsünk egy $f(x)$ függvényt, mely integrálható tetszőleges $\omega > a$ ($b > a \in \mathbb{R}$ vagy $a = -\infty$) esetén az $[\omega,b]$ intervallumon. Amenyiben az $\int_\omega^b f(x)\,\mathrm{d}x$ integrálnak van véges határértéke, ha ω tart a-hoz, akkor azt mondjuk, hogy az $f(x)$

függvény **improprius integrálja** létezik az $[a, b]$ intervallumon és az integrál értékét $\int_a^b f(x)\,\mathrm{d}x$ fogja jelölni. Azaz

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{\omega \to a} \int_\omega^b f(x)\,\mathrm{d}x.$$

$\diamondsuit$

A 2.5.2 példánál az előbbi, míg a 2.5.1 példánál az utóbbi definíció alapján számítottuk ki az improprius integrált. Felhívjuk a figyelmet arra, hogy az improprius integrál jelöléséből nem látszik azonnal, hogy improprius integrálról van szó. Azt nekünk kell észrevennünk a megadott intervallum és az integrálandó függvényből.

2.5.1. Feladat. Határozzuk meg, hogy az alábbi integrálok közül melyik improprius integrál és melyik nem.

$$a)\ \int_0^1 \frac{1}{x^2}\,\mathrm{d}x,\ b)\ \int_1^2 \frac{1}{x^2}\,\mathrm{d}x,\ c)\ \int_0^\infty \sin x\,\mathrm{d}x,\ d)\ \int_0^1 \ln x\,\mathrm{d}x$$

2.5.2. Feladat. Határozzuk meg, az előző feladat c) és d) pontjában szereplő integrálokat!

Sokszor, főleg a valószínűségszámítás tantárgy keretein belül, szerepel az a fontos eset, amikor az egész valós számegyenesen integrálunk egy függvényt. Ennek jelölése $\int_{-\infty}^\infty f(x)\,\mathrm{d}x$. Ez az improprius integrál akkor létezik, ha külön-külön a $\int_{-\infty}^0 f(x)\,\mathrm{d}x$ és $\int_0^\infty f(x)\,\mathrm{d}x$ improprius integrálok is léteznek és ekkor

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x = \int_{-\infty}^0 f(x)\,\mathrm{d}x + \int_0^\infty f(x)\,\mathrm{d}x.$$

2.5.5. Példa. Tekintsük az

$$\int_{-\infty}^\infty \sqrt[3]{x}\,\mathrm{d}x$$

integrált. Mivel

$$\int_0^\infty \sqrt[3]{x}\,\mathrm{d}x = \lim_{\omega \to \infty} \left[\frac{3}{4}x^{4/3}\right]_0^\omega = \lim_{\omega \to \infty} \frac{3}{4}\omega^{4/3} = \infty,$$

azaz ez az improprius integrál nem létezik, így az $\int_{-\infty}^\infty \sqrt[3]{x}\,\mathrm{d}x$ improprius integrál sem létezik. Figyeljük meg, hogy a $\sqrt[3]{x}$ függvény grafikonja szimmetrikus az origóra, azaz ugyanakkora része van a grafikonnak az x-tengely alatt, mint felett. Mégsem mondjuk azt, hogy a függvény integrálja (előjeles terület) a $]-\infty, \infty[$ intervallumon nulla, hiszen ez a nulla két végtelen terület különbségeként állna elő. $\square$

Ha már gyakorlatot szereztünk az improprius integrálok számításában, akkor a következő rövidebb kiszámítási módot is alkalmazhatjuk. Pl.

$$\int_1^\infty \frac{1}{x^3}\,\mathrm{d}x = \left[\frac{1}{-2x^2}\right]_1^\infty = 0 - \left(-\frac{1}{2}\right) = \frac{1}{2}.$$

Itt ugyanúgy jártunk el, mintha a Newton-Leibniz formulát alkalmaztuk volna, csak behelyettesítéskor a végtelenben vett határértéket határoztuk meg a helyettesítési érték helyett.

Ellenőrző kérdések

1. Miért van szükség az improprius integrál bevezetésére?

2. Mi az improprius integrál definíciója? Mely esetekben alkalmazzuk?

3. Hogyan értelmezzük egy függvény 1-től ∞-ig vett integrálját?

4. Hogyan értelmezzük egy függvény $-\infty$-től ∞-ig vett integrálját? Mit ad ez meg szemléletesen, ha a függvény nemnegatív?

A feladatok megoldásai

2.5.1.
Az a) esetben a függvény nem korlátos a $[0,1]$ intervallumon, így improprius integrálról van szó. A b) eset nem improprius integrál. A c) esetben improprius integrálról van szó, mert az integrálási tartomány végtelen. A d) eset is improprius integrál, mert az $\ln x$ függvény nem korlátos a $[0,1]$ intervallumon.

2.5.2.

$$\int_0^\infty \sin x \,\mathrm{d}x = \lim_{\omega\to\infty}\int_0^\omega \sin x \,\mathrm{d}x = \lim_{\omega\to\infty}[-\cos x]_0^\omega = \lim_{\omega\to\infty}(-\cos\omega + 1).$$

Ez a határérték nem létezik, mert a koszinusz függvénynek nincs határértéke a végtelenben, így az improprius integrál sem létezik.

$$\int_0^1 \ln x \,\mathrm{d}x = \lim_{\omega\to 0}\int_\omega^1 \ln x \,\mathrm{d}x = \lim_{\omega\to 0}\int_\omega^1 [x(\ln x - 1)]_\omega^1 = \lim_{\omega\to 0}(-1 - \omega(\ln\omega - 1)) = -1.$$

Az $\ln x$ függvényt parciálisan integráltuk, míg a $\lim_{\omega\to 0}(-1-\omega(\ln\omega-1))$ határérték meghatározásához a L'Hospital-szabályt használtuk.

Önálló munkára kijelölt feladatok: DGY 76. oldal 1-6, 8-14, 17, 18. feladat, DGY 77. oldal 21-32. feladat.

2.6. A határozott integrál alkalmazásai

Ebben a fejezetben a határozott integrál néhány alkalmazását mutatjuk be. Az integrált szöveges feladatok megoldására használjuk, újra szólunk a területszámításról, numerikus sorok konvergenciájára adunk meg újabb kritériumot, hatványsorok integrálásáról mondunk ki egy fontos tételt.

Az integrálszámításnak nincs közvetlen közgazdaságtani alkalmazása. Hasznos eszköz azonban differenciálegyenletek megoldása esetén, vagy a valószínűségi változók tárgyalásakor a valószínűségszámítás tantárgy keretein belül. Ebben a fejezetben néhány más matematikai alkalmazást mutatunk be.

Területszámítás

A határozott integrál egyik legfontosabb alkalmazásáról, a területszámításról, már az előző fejezetekben többször szóltunk. Az

$$\int_a^b f(x)\,\mathrm{d}x$$

formulával értelmezzük egy integrálható $f(x)$ függvény grafikonja és az $[a,b]$ szakasz által meghatározott tartomány előjeles területét. Az előjeles kifejezés itt azt jelenti, hogy a tartomány x-tengely feletti részét pozitív, míg az x-tengely alatti részt negatív előjellel vesszük figyelembe.

2.6.1. Példa. Határozzuk meg a $\sin x$ függvény grafikonja és az x-tengely által közrezárt tartomány területét a $[0,2\pi]$ intervallumon. Ha a $\sin x$ függvényt integrálnánk 0-tól 2π-ig, akkor nullát kapnánk, hiszen az integrál előjeles területet ad meg és a $\sin x$ függvénynek ugyanakkora területű része van az x-tengely alatt mint felett a $[0,2\pi]$ intervallumon. Ehelyett a következő módon járhatunk el.

$$T = 2\int_0^{\pi} \sin x\,\mathrm{d}x = 2[-\cos x]_0^{\pi} = 4.$$

□

2.6.2. Példa. Határozzuk meg az $f(x) = 4 - x^2$ függvény grafikonja és az x-tengely által közrezárt véges tartomány területét! Az $f(x)$ függvény grafikonja az x-tengelyt a -2 és 2 pontokban metszi, ezen az intervallumon $f(x)$ grafikonja az x-tengely felett van.

$$T = \int_{-2}^{2}(4-x^2)\,\mathrm{d}x = \left[4x - \frac{x^3}{3}\right]_{-2}^{2} = \frac{32}{3}.$$

□

Szöveges feladatok megoldása

Szöveges feladatok esetén az integrálást a

$$\sum_{i=1}^{n} f(x_i^\star)\cdot\Delta x_i \tag{2.6.1}$$

összeg határértékének ($\Delta \to 0$) kiszámítására használhatjuk, ha az x_i értékek egy meghatározott $[a,b]$ intervallumból kerülnek ki, $a = x_0 < x_1 < \ldots < x_n = b$ és $x_i^\star \in [x_{i-1}, x_i]$.

Mivel (2.6.1) az $f(x)$ függvény egy integrálközelítő összege, ezért ha $f(x)$ integrálható, akkor

$$\lim_{\Delta\to 0}\left(\sum_{i=1}^{n} f(x_i^\star)\cdot\Delta x_i\right)=\int_a^b f(x)\,\mathrm{d}x.$$

2.6.3. Példa. Egy jármű sebességét az idő függvényében a $v(t)=t^2$ függvény írja le (m/s mértékegységben). Mekkora utat tesz meg ez a jármű az indulásától számított első tíz másodpercben?

Osszuk fe a $[0,10]$ időintervallumot n osztóintervallumra ($0=t_0<t_1<\ldots<t_n=10$) és mindegyik osztóintervallumból jelöljünk ki egy-egy tetszőleges időpontot ($t_i^\star$). Mivel a megtett út az idő és a sebesség szorzata, ezért az első osztóintervallum alatt kb. $(t_1^\star)^2\Delta t_1$ utat tesz meg a jármű, a második alatt kb. $(t_2^\star)^2\Delta t_2$, stb. Emiatt az első tíz másodpercben megtett út kb.

$$\sum_{i=1}^{n}(t_i^\star)^2\Delta t_i.$$

Ez, mivel a $v(t)=t^2$ függvény integrálható a $[0,10]$ intervallumon, ha $\Delta\to 0$, akkor az $\int_0^{10} t^2\,\mathrm{d}t$ értékhez tart. Így kapjuk, hogy a megtett út

$$\int_0^{10} t^2\,\mathrm{d}t=\left[\frac{t^3}{3}\right]_0^{10}=\frac{1000}{3}=333.3\ (m).$$

□

2.6.4. Példa. Egy üzemben vizsgálatot végeztek a dolgozók teljesítményének időbeli változásáról. Azt kapták, hogy egy dolgozó által előállított termékek számát a munkakezdéstől számítva a t. percben jó közelítéssel az

$$f(t)=\frac{10t^2}{10+t^2}$$

függvény írja le. (Pl. az 5. percben 7.14 terméket állít elő.) Hány terméket állít elő egy dolgozó egy nyolcórás műszak alatt?

Bár az $f(t)$ függvény csak pozitív egész számokra értelmezett, mi most tekintsük mégis folytonos függvénynek. Az előállított termékek száma 8 óra (azaz 480 perc alatt) kb.

$$\sum_{i=1}^{n}\frac{10(t_i^\star)^2}{10+(t_i^\star)^2}\cdot\Delta t_i,$$

ez pedig az

$$\int_0^{480}\frac{10t^2}{10+t^2}\,\mathrm{d}t=\int_0^{480}\left(10-\frac{100}{10+t^2}\right)\,\mathrm{d}t=\int_0^{480}\left(10-10\frac{1}{1+(t/\sqrt{10})^2}\right)\,\mathrm{d}t=$$

$$=\left[10t-10\sqrt{10}\mathrm{arctg}(t/\sqrt{10})\right]_0^{480}\approx 4750.$$

□

Integrálkritérium numerikus sorokra

Numerikus sorokra többféle konvergenciakritériumot tanultunk az első félévben (majoráns és minoráns kritérium, gyök- ill. hányadoskritérium). Az improprius integrálok jó lehetőséget adnak egy újabb kritériumra, az integrálkritériumra. Ennek működését két példán mutatjuk be.

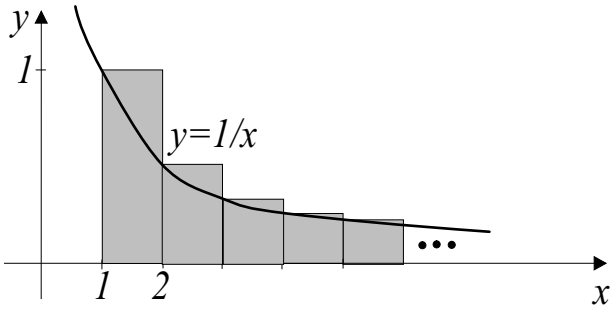
2.6.5. Példa. Tekintsük a

$$\sum_{k=1}^{\infty} \frac{1}{k}$$

sort azaz az úgynevezett harmonikus sort. Ezt a sort úgy is felfoghatjuk, mint a 2.6.1. ábrán szemléltetett végtelen sok téglalap területének összegét, hiszen ezen téglalapok alapja egy egység, míg magasságuk az k. pontban éppen $1/k$. Látható továbbá, hogy ezen téglalapok területösszege biztosan nagyobb, mint az $1/x$ függvény integrálja 1-től végtelenig (improprius integrál). Tehát

$$\sum_{k=1}^{\infty} \frac{1}{k} \geq \int_1^{\infty} \frac{1}{x}\,\mathrm{d}x = [\ln x]_1^{\infty} = \infty,$$

azaz $\sum_{k=1}^{\infty} \frac{1}{k}$ nem konvergens sor. □



2.6.1. ábra. *A harmonikus sor tagjait adó téglalapterületek.*

2.6.6. Példa. Tekintsük a

$$\sum_{k=1}^{\infty} \frac{1}{k^2}$$

sort. Erről azt szeretnénk belátni, hogy konvergens. A sort úgy is felfoghatjuk, mint a 2.6.2. ábrán szemléltetett végtelen sok téglalap területének összege, hiszen ezen téglalapok alapja egy egység, míg magasságuk a k. pontban éppen $1/k^2$. Látható továbbá, hogy ezen téglalapok területösszege biztosan kisebb, mint 1 plusz az $1/x^2$ függvény integrálja 1-től végtelenig (improprius integrál). Tehát

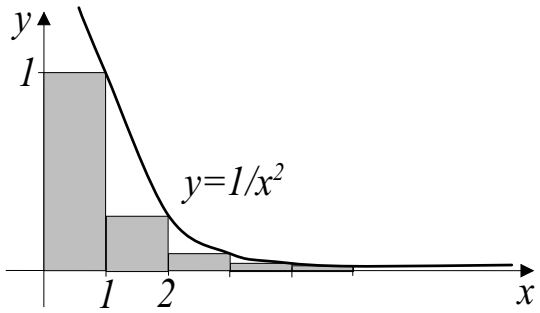
$$\sum_{k=1}^{\infty} \frac{1}{k^2} \leq 1 + \int_1^{\infty} \frac{1}{x^2}\,\mathrm{d}x = 1 + \left[\frac{-1}{x}\right]_1^{\infty} = 2,$$

azaz $\sum_{k=1}^{\infty} \frac{1}{k^2}$ konvergens, sőt azt is látjuk, hogy összege kisebb 2-nél. (A pontos összeg $\pi^2/6 \approx 1.64$.) □

Hatványsorok integrálása

Hatványsorokkal foglalkoztunk az első félévben. Most az ott tanultakat kiegészítjük egy újabb tétellel, ami a hatványsorok integrálására vonatkozik.

2.6.7. Tétel. *A $\sum_{k=0}^{\infty} a_k x^k$ és $\sum_{k=0}^{\infty} a_k x^{k+1}/(k+1)$ hatványsorok konvergenciasugarai megegyeznek. Ha $f(x) = \sum_{k=0}^{\infty} a_k x^k$ és $g(x) = \sum_{k=0}^{\infty} a_k x^{k+1}/(k+1)$, akkor $g'(x) = f(x)$ a konvergenciatartományok belsejében.*



2.6.2. ábra. *A* $\sum_{k=1}^{\infty} \frac{1}{k^2}$ *sor tagjait adó téglalapterületek.*

A tétel szerint ha egy hatványsort tagonként integrálunk, akkor közben a konvergenciasugár nem változik, és az integrálással az eredeti hatványsor által előállított függvény egy primitív függvényét kapjuk.

2.6.8. Példa. Határozzuk meg az

$$1 + 2x + 3x^2 + \ldots = \sum_{k=0}^{\infty} (k+1)x^k$$

hatványsor összegét.

Integráljuk a hatványsort tagonként, majd alkalmazzuk a végtelen mértani sor összegképletét

$$\int \left(\sum_{k=0}^{\infty} (k+1)x^k \right) \, \mathrm{d}x = \sum_{k=0}^{\infty} x^{k+1} + c = \sum_{k=1}^{\infty} x^k + c = \frac{1}{1-x} - 1 + c.$$

Ebből azonnal látható, hogy

$$\sum_{k=0}^{\infty} (k+1)x^k = \left(\int \left(\sum_{k=0}^{\infty} (k+1)x^k \right) \, \mathrm{d}x \right)' = \frac{1}{(1-x)^2}.$$

A konvergenciatartomány a $]-1,1[$ nyílt intervallum. □

2.6.9. Példa. Újabb példaként induljunk ki abból, hogy tudjuk, hogy

$$1 - x + x^2 - x^3 \pm \ldots = \sum_{k=0}^{\infty} (-1)^k x^k = \frac{1}{1+x}, \qquad x \in]-1,1[.$$

Ezt a végtelen mértani sor összegképletéből nyerhetjük, ha x helyébe $-x$-et helyettesítünk. Integráljuk az egyenlet mindkét oldalát

$$\sum_{k=0}^{\infty} (-1)^k \frac{x^{k+1}}{k+1} + c = \ln(1+x)$$

(a c konstanst elegendő csak az egyenlet egyik oldalán kiírni). Mivel $\ln 1 = 0$, ezért c értékének nullának kell lennie. Azaz így megkaptuk a $\sum_{k=0}^{\infty} (-1)^k \frac{x^{k+1}}{k+1}$ hatványsor összegét

$$\sum_{k=0}^{\infty} (-1)^k \frac{x^{k+1}}{k+1} = \ln(1+x).$$

□

Ellenőrző kérdések

1. Hogyan alkalmazható az integrál területszámításra?
2. Mi egy függvény határozott integráljának szemléletes jelentése?
3. Mit hívunk egy függvény integrálközelítő összegének, és az mihez tart, ha a felosztás finomsága nullához tart.
4. Hogyan lehet függvények integrálásával eldönteni a numerikus sorok konvergenciáját?
5. Mondjuk ki a hatványsorok integrálásáról szóló tételt!

Önálló munkára kijelölt feladatok: DGY 71. oldal 5-8. feladat, DGY 72. oldal 10-17. feladat, DGY 86. oldal 42-44. feladat.

3. fejezet

Kétváltozós függvények integrálása

3.1. A kétváltozós integrál értelmezése

Az egyváltozós függvények integrálásának általánosításaként értelmezzük a kétváltozós függvények integrálját és megvizsgáljuk annak tulajdonságait.

Ha egy $f(x) \geq 0$ egyváltozós függvény integrálható az $[a, b]$ intervallumon, akkor az $[a, b]$ intervallum és az $f(x)$ függvény grafikonja közti tartomány területét az $\int_a^b f(x)\,\mathrm{d}x$ érték definiálja. Az egyváltozós integrál segítségével tehát területet tudunk számolni. Már most sejthetjük, hogy egy kétváltozós függvény integrálja térfogatszámításra lesz használható.

A kétváltozós függvények integráljának definíciójánál az egyváltozós függvényeknél megismert utat fogjuk követni. Legyenek adottak az $a_1 < b_1$ és $a_2 < b_2$ valós számok. Ha az a_1 és b_1 számokat bejelöljük egy derékszögű koordinátarendszer x-tengelyén, és ezen pontokon keresztül párhuzamosokat húzunk az y-tengellyel, majd az a_2 és b_2 pontokat jelöljük be az y-tengelyen, és a pontokon keresztül párhuzamosokat húzunk az x-tengellyel, akkor a megrajzolt egyenesek egy téglalapot határoznak meg, melyet $T(a_1, b_1, a_2, b_2)$-vel, vagy egyszerűen T-vel fogunk jelölni. Tehát

$$T(a_1, b_1, a_2, b_2) = \{(x, y) \in \mathbb{R}^2 \mid a_1 \leq x \leq b_1, a_2 \leq y \leq b_2\}.$$

Figyeljük meg, hogy a T téglalap oldalai párhuzamosak a koordináta-tengelyekkel! A kétváltozós függvények integrálját először ilyen téglalapokon értelmezzük.

Tekintsünk egy $T(a_1, b_1, a_2, b_2)$ téglalapot. Osszuk fel az $[a_1, b_1]$ intervallumot n részre az

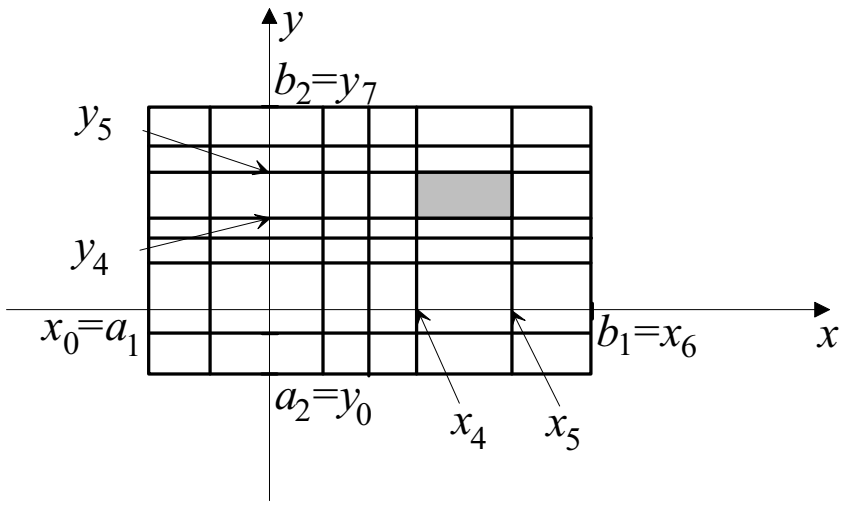
$$a_1 = x_0 < x_1 < \ldots < x_{n-1} < x_n = b_1$$

osztópontokkal, ill. az $[a_2, b_2]$ intervallumot m részre az

$$a_2 = y_0 < y_1 < \ldots < y_{m-1} < y_m = b_2$$

osztópontokkal. Ezen osztópontok segítségével az eredeti téglalapot kisebb, ún. osztótéglalapokra bonthatjuk fel, amit a téglalap egy felosztásának nevezünk. A felosztást $P(x_0, x_1, \ldots, x_n; y_0, y_1, \ldots, y_m)$ vagy csak egyszerűen P jelöli. Az $[x_{i-1}, x_i]$ és $[y_{j-1}, y_j]$ intervallumok által meghatározott osztótéglalapot (i, j)-edik osztótéglalapnak nevezzük.

3.1.1. Példa. A 3.1.1. ábra a $T(a_1, b_1, a_2, b_2)$ téglalap egy felosztását mutatja. Itt először az $[a_1, b_1]$ intervallumot osztottuk fel az $a_1 = x_0 < x_1 < \ldots < x_6 = b_1$ osztópontokkal hat intervallumra, majd az $[a_2, b_2]$ intervallumot az $a_2 = y_0 < y_1 < \ldots < y_7 = b_2$



3.1.1. ábra. *Téglalap egy felosztása.*

osztópontokkal hét részre. Ezen osztópontok segítségével az eredeti téglalapot 42 osztótéglalapra bontottuk fel. Ezek közül az $(5,5)$-ödik osztótéglalapot besötétítettük. □

Az egyváltozós függvények esetéhez hasonlóan definiáljuk a következő értékeket:

$$\Delta x_i = x_i - x_{i-1}, \quad \Delta y_j = y_j - y_{j-1} \quad (i = 1, \dots, n;\ j = 1, \dots, m),$$

és a felosztás Δ finomsága legyen az osztótéglalapok átlóinak maximuma, azaz

$$\Delta = \max_{i=1,\dots,n;\ j=1,\dots,m} \left\{ \sqrt{(\Delta x_i)^2 + (\Delta y_j)^2} \right\}.$$

Azt mondjuk, hogy egy felosztás finomabb egy $\delta > 0$ számnál, ha $\Delta < \delta$, azaz ha a felosztás finomsága δ-nál kisebb.

Tekintsünk most egy $f(x,y)$ kétváltozós függvényt, amely értelmezett a T téglalapon, és vegyük T-nek egy P felosztását. Válasszunk ki mindegyik osztótéglalapból egy-egy $(x_i^\star, y_j^\star)$ pontot $(i = 1, \dots, n;\ j = 1, \dots, m)$ és képezzük az

$$I(f, P, P^\star) = \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i^\star, y_j^\star) \cdot \Delta x_i \cdot \Delta y_j$$

ún. integrálközelítő összeget. A jelölésben az I az integrál szó kezdőbetűjéből származik, a zárójelben lévő mennyiségek pedig azt mutatják, hogy ez az f függvénytől, a T téglalap felosztásától és a felosztás osztótéglalapjaiból kiválasztott tetszőleges $(x_i^\star, y_j^\star)$ pontoktól függ, melyek halmazát $P^\star$-gal jelöltük.

3.1.2. Definíció. Azt mondjuk, hogy az $I(f, P, P^\star)$ **integrálközelítő összegek tartanak az** $I \in \mathbb{R}$ **számhoz** $\Delta \to 0$ esetén, ha minden $\varepsilon > 0$ számhoz létezik olyan $\delta > 0$ szám, hogy bármely δ-nál finomabb P felosztásra, és bármely $(x_i^\star, y_j^\star)$ pontválasztás esetén az (i,j)-edik osztótéglalapból, teljesül, hogy $|I(f, P, P^\star) - I| < \varepsilon$. Jelölésben: $\lim_{\Delta \to 0} I(f, P, P^\star) = I$. ♢

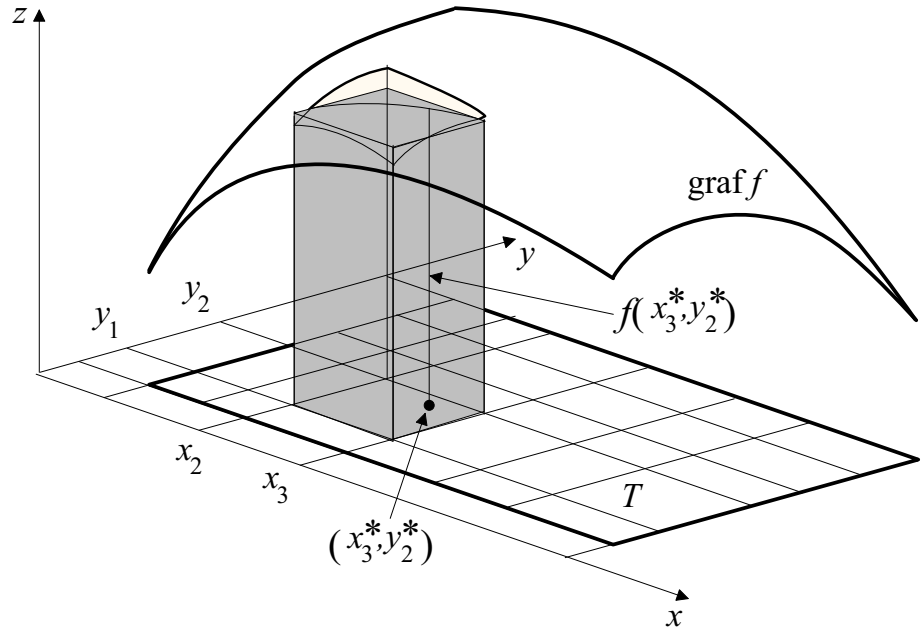
3.1.3. Definíció. Ha a T téglalapon értelmezett $f(x,y)$ függvény integrálközelítő összegeinek van határértéke $\Delta \to 0$ esetén, és a határérték az $I \in \mathbb{R}$ szám, akkor azt

mondjuk, hogy az $f(x,y)$ **kétváltozós függvény integrálható** a T téglalapon és az **integrálja** I. Jelölésben:

$$I = \int_T f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

$\diamondsuit$

Az integrálközelítő összeg és a kétváltozós integrál geometriai jelentését a következőképpen adhatjuk meg. Tekintsünk egy nemnegatív függvényt egy T téglalapon (lásd



3.1.2. ábra. *Az integrálközelítő összegek szemléltetése.*

3.1.2. ábra), majd osszuk fel a T téglalapot osztótéglalapokra (az ábrán 30 osztórész látható). Az integrálközelítő összeg $f(x_i^\star, y_j^\star)\cdot\Delta x_i\cdot\Delta y_j$ tagja egy Δx_i és Δy_j alapélű és $f(x_i^\star, y_j^\star)$ magasságú téglatest térfogatát adja meg. (A $(3,2)$-edik osztótéglalaphoz tartozó téglatestet besötétítettük.) Emiatt, ha $\Delta\to 0$ esetén az integrálközelítő összegeknek van határértékük, azaz a függvény integrálható a T téglalapon, akkor az integrál értékével definiálhatjuk a függvénygrafikon és az xy-sík T téglalapja közti tartomány térfogatát. Hasonló interpretációhoz jutunk ha nem csak nemnegatív függvényeket vizsgálunk, azzal a kiegészítéssel, hogy az integrál a függvénygrafikon és az xy-sík közé eső tartomány térfogatát negatív előjellel veszi figyelembe, ha a tartomány az xy-sík alatt helyezkedik el.

Kétváltozós függvények közül is csak a korlátos függvények lehetnek integrálhatók. Ezt fejezi ki a következő tétel.

3.1.4. Tétel. *Ha egy $f(x,y)$ függvény integrálható egy T téglalapon, akkor ott korlátosnak kell lennie.*

A korlátosság tehát szükséges feltétele az integrálhatóságnak. Egy elégséges feltételt a következő tétel ad.

3.1.5. Tétel. *Ha az $f(x,y)$ függvény folytonos a T téglalapon, akkor azon integrálható is.*

A tétel következménye, hogy minden elemi függvényekből összetett kétváltozós függvény integrálható minden olyan téglalapon, amely teljesen a függvény értelmezési tartományához tartozik.

3.1.1. Feladat. Legyen T egy téglalap. Igazoljuk, hogy a konstans nulla függvény integrálható T-n!

3.1.2. Feladat. Tekintsük a $T(0,1,0,1)$ téglalapot. Igazoljuk, hogy az

$$f(x,y) = \begin{cases} 0, & x \text{ vagy } y \text{ irracionális,} \\ 1, & \text{különben} \end{cases}$$

függvény nem integrálható a T téglalapon.

A következő tételek a kétváltozós függvények integráljának néhány fontos tulajdonságát fogalmazzák meg.

3.1.6. Tétel. *Ha az $f(x,y)$ és $g(x,y)$ függvények integrálhatóak a T téglalapon, akkor $f(x,y)+g(x,y)$ és $\alpha \cdot f(x,y)$ (α tetszőleges valós szám) is integrálható, valamint*

$$\int_T (f(x,y)+g(x,y))\,\mathrm{d}x\,\mathrm{d}y = \int_T f(x,y)\,\mathrm{d}x\,\mathrm{d}y + \int_T g(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

$$\int_T (\alpha \cdot f(x,y))\,\mathrm{d}x\,\mathrm{d}y = \alpha \cdot \int_T f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

3.1.7. Tétel. *Tekintsük a $T(a_1,b_1,a_2,b_2)$ téglalapot, melyet egy tetszőleges c valós számmal ($a_1 < c < b_1$) a $T_1(a_1,c,a_2,b_2)$ és $T_2(c,b_1,a_2,b_2)$ osztótéglalapokra bontunk fel. Az $f(x,y)$ függvény akkor és csak akkor integrálható a T téglalapon, ha a T_1 és T_2 téglalapokon is integrálható, valamint*

$$\int_T f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{T_1} f(x,y)\,\mathrm{d}x\,\mathrm{d}y + \int_{T_2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Az utóbbi tétel állítása szerint, ha a $T(a_1,b_1,a_2,b_2)$ téglalapot az $[a_1,b_1]$ intervallumban felvett osztópont segítségével két részre osztjuk, akkor egy függvény, amely integrálható volt a T téglalapon, integrálható lesz a két osztótéglalapon is, ill. ha a függvény a két osztótéglalapon integrálható, akkor a T téglalapon is. A két osztótéglalapon vett integrálok összege megegyezik a T téglalapon vett integrállal. Természetesen ez az állítás akkor is igaz, ha az $[a_1,b_1]$ intervallum helyett az $[a_2,b_2]$ intervallumban felvett ponttal osztjuk két részre a téglalapot. A tételt többször egymás után alkalmazva kapjuk, hogy ha a T téglalapnak vesszük egy P felosztását, akkor az $f(x,y)$ függvény akkor és csak akkor integrálható minden osztótéglalapon, ha a T téglalapon integrálható, és az osztótéglalapokon vett integrálok összege megegyezik a T téglalapon vett integrállal.

Eddig a kétváltozós integrált csak téglalapokon értelmeztük. Most kiterjesztjük az integrálhatóság fogalmát tetszőleges korlátos halmazra. Tegyük fel, hogy az $f(x,y)$ függvény értelmezési tartománya a síkbeli, korlátos H halmaz. Mivel H korlátos, biztosan létezik egy olyan T téglalap, amelyre $H \subset T$. Az $f(x,y)$ függvény nincs értelmezve T azon pontjaiban, melyek nem tartoznak hozzá a H halmazhoz. Definiáljuk az $\tilde{f}(x,y)$ függvényt a következő módon:

$$\tilde{f}(x,y) = \begin{cases} f(x,y), & \text{ha } (x,y) \in H, \\ 0, & \text{ha } (x,y) \in T \setminus H. \end{cases}$$

Az $\tilde{f}(x,y)$ függvény már T minden pontjában értelmezett. A függvény a H halmazon kívüli pontokban nullát vesz fel, a H halmazon pedig megegyezik az $f(x,y)$ függvénnyel. Az $\tilde{f}(x,y)$ függvényt az $f(x,y)$ függvény T téglalapra való kiterjesztésének nevezzük.

3.1.8. Definíció. Tekintsünk egy $f(x,y)$ függvényt, amelynek értelmezési tartománya a H korlátos halmaz. Legyen T egy olyan téglalap, melyre $H \subset T$ és $\tilde{f}(x,y)$ az $f(x,y)$ függvény kiterjesztése a T téglalapra. Az mondjuk, hogy az $f(x,y)$ függvény **integrálható** a H halmazon, ha $\tilde{f}(x,y)$ integrálható a T téglalapon. Az $\int_T \tilde{f}(x,y)\,\mathrm{d}x\,\mathrm{d}y$ értéket ekkor az $f(x,y)$ függvény H halmazon vett **integráljának** nevezzük és az

$$\int_H f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

módon jelöljük. $\diamondsuit$

Egy függvény tetszőleges halmazon vett integrálját a téglalapokon vett integrál segítségével definiáltuk. Igazolható, hogy az $\int_H f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ érték független attól, hogy melyik T téglalapra terjesztettük ki a $H \subset T$ halmazon értelmezett $f(x,y)$ függvényt.

A téglalapokon vett integrálok tulajdonságait felhasználva bizonyíthatjuk a következő két tételt.

3.1.9. Tétel. *Ha az $f(x,y)$ és $g(x,y)$ függvény is integrálható a H korlátos halmazon, akkor $f(x,y)+g(x,y)$ és $\alpha\cdot f(x,y)$ (α tetszőleges valós szám) is integrálható H-n, valamint*

$$\int_H (f(x,y)+g(x,y))\,\mathrm{d}x\,\mathrm{d}y = \int_H f(x,y)\,\mathrm{d}x\,\mathrm{d}y + \int_H g(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

$$\int_H (\alpha\cdot f(x,y))\,\mathrm{d}x\,\mathrm{d}y = \alpha\cdot\int_H f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

3.1.10. Tétel. *Ha az $f(x,y)$ függvény integrálható a H_1 és H_2 diszjunkt halmazokon, akkor $f(x,y)$ integrálható a $H_1\cup H_2$ halmazon is és*

$$\int_{H_1\cup H_2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{H_1} f(x,y)\,\mathrm{d}x\,\mathrm{d}y + \int_{H_2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Ellenőrző kérdések

1. Ismertessük röviden az egyváltozós függvények integrálhatóságának definíciójához vezető út fontosabb állomásait!
2. Definiáljuk a téglalap és felosztás fogalmakat!
3. Hogyan definiáljuk az integrálközelítő összeget kétváltozós függvények esetén?
4. Mondjunk elégséges feltételt egy függvény integrálhatóságára!
5. Hogyan értelmeztük egy kétváltozós függvény integrálját egy tetszőleges korlátos halmazon?
6. Milyen műveleti szabályok igazak a kétváltozós függvények integráljára?
7. Mi a fő alkalmazási területe a kétváltozós függvények integráljának?

A feladatok megoldásai

3.1.1.
A konstans nulla függvényre a T téglalap bármely P felosztása esetén $I(f, P, P^\star) = 0$. Így nyilvánvalóan $I(f, P, P^\star) \to 0$, ha $\Delta \to 0$. Ez azt jelenti, hogy a konstans nulla függvény integrálható a T téglalapon és az integrálja nulla.

3.1.2.
Az adott f függvényre a T téglalap bármely P felosztása esetén megválaszthatjuk úgy az $(x_i^\star, y_j^\star)$ pontokat az egyes osztótéglalapokban, hogy $I(f, P, P^\star) = 0$ teljesüljön, és úgy is, hogy $I(f, P, P^\star) = 1$ legyen, hiszen bármely osztótéglalapban található olyan pont, melynek minkét koordinátája racionális és olyan is amelyiknek valamelyik koordinátája irracionális. Emiatt az $I(f, P, P^\star)$ integrálközelítő összegeknek nem lehet határértéke, ha $\Delta \to 0$, ami azt jelenti, hogy az adott függvény nem integrálható.

3.2. A kétváltozós integrál kiszámítása

Az előző fejezetben definiáltuk egy kétváltozós függvény integrálhatóságát. Felsoroltuk az integrál néhány tulajdonságát is. Ebben a fejezetben azzal foglalkozunk, hogy hogyan lehet a kétváltozós integrálokat kiszámítani. Látni fogjuk, hogy egy kétváltozós függvény integrálja két egyváltozós függvény integrálására vezethető vissza.

Ebben a fejezetben azt vizsgáljuk meg, hogy hogyan lehet a kétváltozós függvények integrálját kiszámítani. A téglalapon, az ún. normáltartományokon való integrálással és a helyettesítéses integrállal ismerkedünk meg.

3.2.1. Integrálás téglalapon

A legegyszerűbb eset az, amikor egy kétváltozós függvény integrálját szeretnénk kiszámítani egy T téglalapon.

3.2.1. Tétel. *Tegyük fel, hogy az $f(x,y)$ függvény integrálható a $T(a_1,b_1,a_2,b_2)$ téglalapon. Ekkor (ha a jobb oldalon szereplő függvények integrálhatóak) az integrál értéke az*

$$\int_T f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{a_2}^{b_2}\left(\int_{a_1}^{b_1} f(x,y)\,\mathrm{d}x\right)\,\mathrm{d}y = \int_{a_1}^{b_1}\left(\int_{a_2}^{b_2} f(x,y)\,\mathrm{d}y\right)\,\mathrm{d}x$$

képlettel számítható ki.

Ha az $f(x,y)$ függvény folytonos, akkor a tétel feltételei teljesülnek. Egy kétváltozós függvény integrálása visszavezethető két egyváltozós függvény egymás utáni integrálására, ahol vagy x és y szerint, vagy fordított sorrendben integrálunk. Emiatt a kétváltozós függvények integrálját szokás kettősintegrálnak is nevezni. Figyeljük meg, hogy az első integráljelhez az utolsó változó tartozik, míg a másodikhoz az első. Ha megállapodunk abban, hogy a továbbiakban mindig betartjuk ezt a megfeleltetést, akkor – ha félreértést nem okoz – a zárójeleket el is hagyhatjuk.

Az eddigiek alkalmazására mutatunk két példát.

3.2.2. Példa. Az $f(x,y) = 3x+y$ kétváltozós függvény integrálját fogjuk meghatározni a $T(0,1,3,5)$ téglalapon. Mivel az $f(x,y)$ függvény folytonos az adott halmazon, ezért a 3.2.1. tételben szereplő összes integrál létezik, és az integrál értéke

$$\int_T (3x+y)\,\mathrm{d}x\,\mathrm{d}y = \int_3^5\int_0^1 (3x+y)\,\mathrm{d}x\,\mathrm{d}y = \int_3^5 \left[3\frac{x^2}{2}+xy\right]_0^1\,\mathrm{d}y =$$

$$= \int_3^5 \left(\frac{3}{2}+y\right)\,\mathrm{d}y = \left[\frac{3}{2}y+\frac{y^2}{2}\right]_3^5 = 11.$$

Mivel $f(x,y) \geq 0$ a T téglalapon, ezért az integrál értéke a T téglalap és a függvény grafikonja közti térbeli tartomány térfogatát adja meg. □

3.2.3. Példa. Meghatározzuk az $f(x,y) = xy^2$ függvény integrálját a $T(0,1,0,1)$ téglalapon (melyet szokás egységnégyzetnek is nevezni). A 3.2.1. tételt alkalmazva kapjuk, hogy

$$\int_T xy^2\,\mathrm{d}x\,\mathrm{d}y = \int_0^1\int_0^1 xy^2\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \left[\frac{x^2y^2}{2}\right]_0^1\,\mathrm{d}y =$$

$$= \int_0^1 \frac{y^2}{2}\,\mathrm{d}y = \left[\frac{y^3}{6}\right]_0^1 = \frac{1}{6}.$$

□

Az utóbbi példában szereplő integrál értékét egyszerűbben is kiszámíthatjuk a következő észrevételek segítségével. Induljunk ki újra az

$$\int_T xy^2\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_0^1 xy^2\,\mathrm{d}x\,\mathrm{d}y$$

kifejezésből. A belső integrálban az y^2 tényező csak egy konstans szorzóként viselkedik, hiszen az x változó szerint integrálunk. Emiatt az y^2 szorzó kiemelhető az integráljel elé

$$\int_0^1 \left(y^2 \int_0^1 x\,\mathrm{d}x\right)\,\mathrm{d}y.$$

Most pedig $\int_0^1 x\,\mathrm{d}x$ egy konstans szorzó a külső integrálban, így kiemelhető az integrálból.

$$\left(\int_0^1 x\,\mathrm{d}x\right)\left(\int_0^1 y^2\,\mathrm{d}y\right).$$

Az integrál tehát két egyváltozós határozott integrál szorzataként számítható ki

$$\int_T xy^2\,\mathrm{d}x\,\mathrm{d}y = \left[\frac{x^2}{2}\right]_0^1 \cdot \left[\frac{y^3}{3}\right]_0^1 = \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{6}.$$

Ezt a megfigyelésünket tétel formájában is megfogalmazzuk.

3.2.4. Tétel. *Tegyük fel, hogy a $T(a_1, b_1, a_2, b_2)$ téglalapon értelmezett $f(x, y)$ kétváltozós függvény valamilyen $f_1(x)$ és $f_2(y)$ függvények segítségével felírható $f(x, y) = f_1(x) \cdot f_2(y)$ alakban ($(x, y) \in T$). Ekkor*

$$\int_T f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \left(\int_{a_1}^{b_1} f_1(x)\,\mathrm{d}x\right) \cdot \left(\int_{a_2}^{b_2} f_2(y)\,\mathrm{d}y\right),$$

amennyiben a képletben szereplő integrálok léteznek.

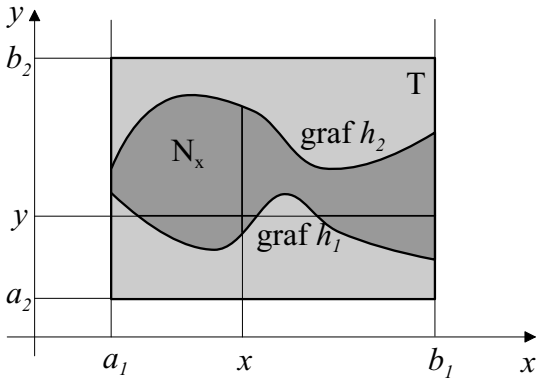
3.2.5. Példa. Meghatározzuk az $f(x, y) = e^{x+y}$ függvény integrálját a $T(0, \ln 3, 0, \ln 4)$ téglalapon. Mivel $f(x, y) = e^{x+y} = e^x \cdot e^y$, ezért a 3.2.4. tételt alkalmazva kapjuk, hogy

$$\int_T e^{x+y}\,\mathrm{d}x\,\mathrm{d}y = \left(\int_0^{\ln 3} e^x\,\mathrm{d}x\right) \cdot \left(\int_0^{\ln 4} e^y\,\mathrm{d}y\right) = [e^x]_0^{\ln 3} \cdot [e^y]_0^{\ln 4} = 2 \cdot 3 = 6.$$

□

3.2.1. Feladat. Határozzuk meg az $f(x, y) = \sin x \cdot \cos y$ függvény integrálját a $T(0, \pi/2, 0, \pi/2)$ téglalapon!

3.2.2. Feladat. Határozzuk meg az $f(x, y) = x + y$ függvény integrálját a $T(0, \pi/2, 0, \pi/2)$ téglalapon!



3.2.1. ábra. *Egy x-szerinti normáltartomány.*

3.2.2. Integrálás normáltartományon

A téglalapok után az ún. normáltartományokon való integrálás módszerével ismerkedünk meg. Tekintsük a 3.2.1. ábrát. Tegyük fel, hogy egy $f(x,y)$ kétváltozós függvény integrálható az ábrán látható N_x tartományon. Az N_x jelölést majd később indokoljuk. Most csak azt figyeljük meg, hogy az N_x tartományt két, az y-tengellyel párhuzamos, egyenes és a h_1 és h_2 függvények grafikonjai határolják. Az $f(x,y)$ függvény tetszőleges N_x-et tartalmazó T téglalapra való $\tilde{f}(x,y)$ kiterjesztése is integrálható és $\int_{N_x} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_T \tilde{f}(x,y)\,\mathrm{d}x\,\mathrm{d}y$. Tehát az $f(x,y)$ függvény N_x-en vett integrálja megegyezik pl. az $f(x,y)$ függvény kiterjesztésének az ábrán látható téglalapon vett integráljával. Téglalapon vett integrálokkal az előző fejezetben foglalkoztunk, a 3.2.1. tétel segítségével tudtuk őket kiszámítani. A 3.2.1. tételben két kiszámítási módot is megadtunk, de a mostani esetben csak a másodikat érdemes használni.

$$\int_{N_x} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_T \tilde{f}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{a_1}^{b_1}\int_{a_2}^{b_2} \tilde{f}(x,y)\,\mathrm{d}y\,\mathrm{d}x =$$

$$= \int_{a_1}^{b_1}\int_{h_1(x)}^{h_2(x)} \tilde{f}(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_{a_1}^{b_1}\int_{h_1(x)}^{h_2(x)} f(x,y)\,\mathrm{d}y\,\mathrm{d}x.$$

Rögzített x érték esetén az y változó szerint a_2-től b_2-ig kell integrálni az $\tilde{f}$ függvényt. Mivel ez a függvény nulla $h_1(x)$-nél kisebb ill. $h_2(x)$-nél nagyobb y értékekre, ezért elegendő csak ezen két határ között integrálni. Kihasználtuk továbbá, hogy $f(x,y)$ és $\tilde{f}(x,y)$ megegyeznek az N_x halmazon. Amennyiben először rögzített y mellett x szerint akartunk volna integrálni, akkor meg kellett volna határoznunk az integrálási határokat (lásd 3.2.1. ábra). Ez nagyon bonyolult feladat lenne, sokkal egyszerűbb az integrálást az y változóval kezdeni.

3.2.6. Definíció. Legyen $h_1(x)$ és $h_2(x)$ két, az $[a_1,b_1]$ intervallumon értelmezett folytonos függvény, melyekre $h_1(x) \le h_2(x)$ minden $a_1 \le x \le b_1$ esetén. Az

$$\{(x,y) \in \mathbb{R}^2 \mid x \in [a_1,b_1], h_1(x) \le y \le h_2(x)\}$$

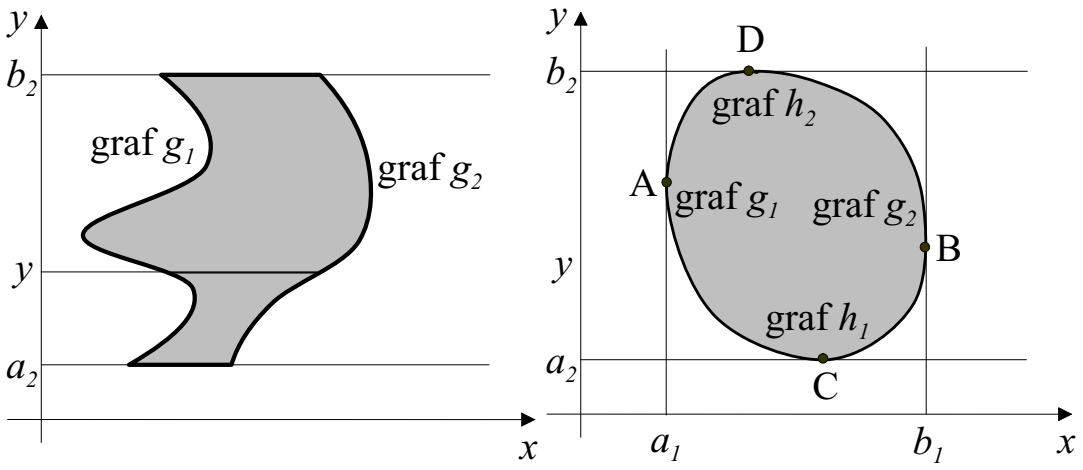
síkbeli halmazt ***x*-szerinti normáltartománynak** hívjuk. Jelölés: $N_x(a_1,b_1,h_1,h_2)$, vagy röviden csak N_x. $\diamondsuit$

A 3.2.1. ábrán tehát amiatt használtuk az N_x jelölést az integrálási tartományra, mert az egy x-szerinti normáltartomány volt. Hasonlóan értelmezzük az y-szerinti normáltartományokat is. A 3.2.2. ábra bal oldalán egy y-szerinti normáltartományt szemléltettünk.

3.2.7. Definíció. Legyen $g_1(y)$ és $g_2(y)$ két az $[a_2, b_2]$ intervallumon értelmezett folytonos függvény, melyekre $g_1(y) \le g_2(y)$ minden $a_2 \le y \le b_2$ esetén. Az

$$\{(x, y) \in \mathbb{R}^2 \mid y \in [a_2, b_2], g_1(y) \le x \le g_2(y)\}$$

síkbeli halmazt ***y*-szerinti normáltartománynak** hívjuk. Jelölés: $N_y(a_2, b_2, g_1, g_2)$, vagy röviden csak N_y. $\diamondsuit$



3.2.2. ábra. *Bal oldal: egy y-szerinti normáltartomány. Jobb oldal: x- és y-szerinti normáltartomány.*

3.2.8. Tétel. *Tegyük fel, hogy az $f(x, y)$ kétváltozós függvény integrálható az $N_x(a_1, b_1, h_1, h_2)$ normáltartományon. Ekkor*

$$\int_{N_x} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_{a_1}^{b_1} \int_{h_1(x)}^{h_2(x)} f(x, y)\,\mathrm{d}y\,\mathrm{d}x,$$

amennyiben a jobb oldalon szereplő integrálok léteznek.

3.2.9. Tétel. *Tegyük fel, hogy az $f(x, y)$ kétváltozós függvény integrálható az $N_y(a_2, b_2, g_1, g_2)$ normáltartományon. Ekkor*

$$\int_{N_y} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_{a_2}^{b_2} \int_{g_1(y)}^{g_2(y)} f(x, y)\,\mathrm{d}x\,\mathrm{d}y,$$

amennyiben a jobb oldalon szereplő integrálok léteznek.

Természetesen vannak olyan síkbeli tartományok is, melyek tekinthetőek x-szerinti és y-szerinti normáltartománynak is. Ilyenkor a fenti tételek közül a számunkra legjobban megfelelőt használjuk. A 3.2.2. ábra jobb oldalán látható tartomány esetén pl. az A és B pontok közti két görbe egy x-szerinti, míg a C és D pontok köztiek egy y-szerinti normáltartományt határoznak meg.

3.2.10. Példa. Kiszámítjuk az $f(x, y) = x + y$ függvény integrálját a $(0, 0)$, $(1, 0)$ és $(1, 1)$ pontok által meghatározott háromszögtartományon. A megadott háromszögtartomány tekinthető x- és y-szerinti normáltartománynak is. Az x-szerinti normáltartományként való felírás az egyszerűbb, hiszen ekkor a határoló görbéket az x változó függvényében

kell felírnunk. A háromszögtartomány tehát az $N_x(0,1,0,x)$ x-szerinti normáltartomány. (A tartományt az $x=0$, $x=1$ egyenesek ill. a $h_1(x)=0$ és $h_2(x)=x$ függvények grafikonjai határolják.) Alkalmazzuk a 3.2.8. tételt az integrál kiszámítására.

$$\int_{N_x}(x+y)\,\mathrm{d}x\,\mathrm{d}y=\int_0^1\int_0^x(x+y)\,\mathrm{d}y\,\mathrm{d}x=\int_0^1\left[xy+\frac{y^2}{2}\right]_0^x\,\mathrm{d}x=\int_0^1\left(\frac{3x^2}{2}\right)\,\mathrm{d}x=\frac{1}{2}.$$

□

3.2.3. Feladat. Írjuk át az

$$\int_0^1\int_y^{\sqrt{y}}1\,\mathrm{d}x\,\mathrm{d}y$$

integrált x-szerinti normáltartományra vonatkozó integrállá, és számítsuk ki az integrál értékét! Mit ad meg ez az integrál?

3.2.3. Integrálás helyettesítéssel

Az előző két fejezetben azzal foglalkoztunk, hogy hogyan lehet kétváltozós függvények integrálját kiszámítani téglalapokon és normáltartományokon. Gyakran előfordulhat azonban, hogy új ismeretlenek bevezetésével (ún. helyettesítéses integrálással) az integrálandó függvényt sokkal egyszerűbb kezelni, vagy könnyebb az integrálási tartományt megadni. Ezzel az integrál kiszámítása lényegesen leegyszerűsíthető. Most azt vizsgáljuk meg, hogy az új változók bevezetése hogyan változtatja meg az integrál kiszámítási módját.

Emlékezzünk vissza az egyváltozós esetre, ahol szintén tanultunk helyettesítéses integrálásról. Pl. az $f(x)=\sqrt{x}e^{\sqrt{x}}$ függvény $[1,4]$ intervallumon vett integrálját az $x=t^2$ helyettesítéssel számíthatjuk ki, azaz

$$\int_1^4\sqrt{x}e^{\sqrt{x}}\,\mathrm{d}x=\int_1^2 te^t\cdot 2t\,\mathrm{d}t=\ldots=4e^2-2e.$$

Az új változókra való áttérés során három dolog változott meg. Egyrészt az x változó helyébe beírtuk az új változóval kifejezett értékét ($x=t^2$), másrészt megváltoztak az integrálási határok ($x\in[1,4]$ esetén $t\in[1,2]$) és harmadrészt az integrálandó függvényt be kellett szorozni a régi változót az új változóval kifejező függvény deriváltjával ($x'(t)=2t$). Látni fogjuk, hogy kétváltozós függvények esetén is hasonló következményei lesznek az új változókra való áttérésnek.

Először bevezetjük a Jacobi-determináns fogalmát, amely fontos szerepet fog játszani a helyettesítéses integrálásnál: az abszolút értéke fogja átvenni az egyváltózóban megszokott helyettesítő függvény deriváltjának szerepét.

Legyenek adottak az $x(u,v)$ és $y(u,v)$ kétváltozós függvények (később látni fogjuk, hogy miért jelöltük x-szel és y-nal őket). Ekkor az $(u,v)\mapsto(x(u,v),y(u,v))$ hozzárendelés egy síkbeli transzformációt fog megvalósítani a sík (u,v) és $(x(u,v),y(u,v))$ pontjai között.

3.2.11. Definíció. Tegyük fel, hogy az $x(u,v)$ és $y(u,v)$ függvények parciálisan differenciálhatók és deriváltjaik folytonosak egy (u,v) pontban. Ekkor a

$$J(u,v)=\begin{vmatrix}x'_u(u,v) & x'_v(u,v)\\ y'_u(u,v) & y'_v(u,v)\end{vmatrix}=x'_u(u,v)\cdot y'_v(u,v)-x'_v(u,v)\cdot y'_u(u,v)$$

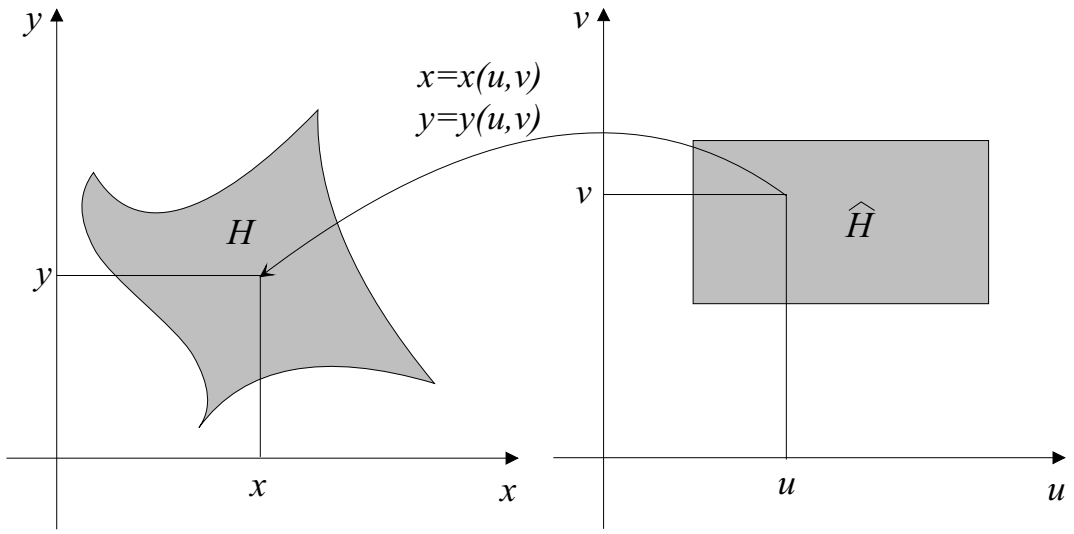
kifejezést az $(u,v) \mapsto (x(u,v), y(u,v))$ transzformáció (u,v) pontbeli **Jacobi-determinánsának** nevezzük. $\diamondsuit$

3.2.12. Tétel. *Tegyük fel, hogy az $f(x,y)$ kétváltozós függvény integrálható a H halmazon, az $x(u,v)$ és $y(u,v)$ függvények kölcsönösen egyértelmű leképezést definiálnak a $\widehat{H} = \{(u,v) \in \mathbb{R}^2 \mid (x(u,v), y(u,v)) \in H\}$ és H halmazok között, és hogy ennek a leképezésnek a $J(u,v)$ Jacobi-determinánsa nem nulla $\widehat{H}$ egyik pontjában sem. Ekkor*

$$\int_H f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\widehat{H}} f(x(u,v), y(u,v)) \cdot |J(u,v)|\,\mathrm{d}u\,\mathrm{d}v.$$

Figyeljük meg, hogy a kétváltozós függvények helyettesítéses integrálásánál is három dolog változik az új változókra való áttérés során. Az x és y változók helyébe beírjuk a helyettesítő függvényeket. Az integrálandó függvényt be kell szorozni a transzformáció Jacobi-determinánsának abszolút értékével. Megváltoznak az integrálási határok, azaz H helyett a $\widehat{H}$ halmazon kell integrálni.

A fenti tételben leírt megfeleltetést könnyebben elképzelhetjük, ha a H halmazt az xy-sík, míg a $\widehat{H}$ halmazt az uv-sík részhalmazának tekintjük. Ezt mutatja a 3.2.3. ábra. Természetesen a helyettesítés célja, hogy a helyettesítés elvégzése után az integrál sokkal



3.2.3. ábra. *Az integrálási tartomány megváltozása helyettesítéses integrálás során.*

egyszerűbben kiszámítható legyen. Ebben segít, ha olyan transzformációt vezetünk be, amely a H halmazt téglalapra vagy normáltartományra képezi le.

Megjegyezzük, hogy a helyettesítésre vonatkozó tétel akkor is igaz marad, ha megengedjük, hogy az $(u,v) \mapsto (x(u,v), y(u,v))$ transzformáció a $\widehat{H}$ halmaz határán ne legyen kölcsönösen egyértelmű ill. azt, hogy a határon a Jacobi-determináns nulla legyen.

Korábban már bevezettük a síkbeli polárkoordináta-rendszert. Ez az $x = x(r,\varphi) = r\cos\varphi$, $y = y(r,\varphi) = r\sin\varphi$ síkbeli transzformációt adta meg. A szereplő függvények nyilvánvalóan parciálisan differenciálhatók és a deriváltak folytonosak minden $(r,\varphi) \in \mathbb{R}^2$ pontban. Eszerint a transzformációnak minden pontban létezik a Jacobi-determinánsa, amely

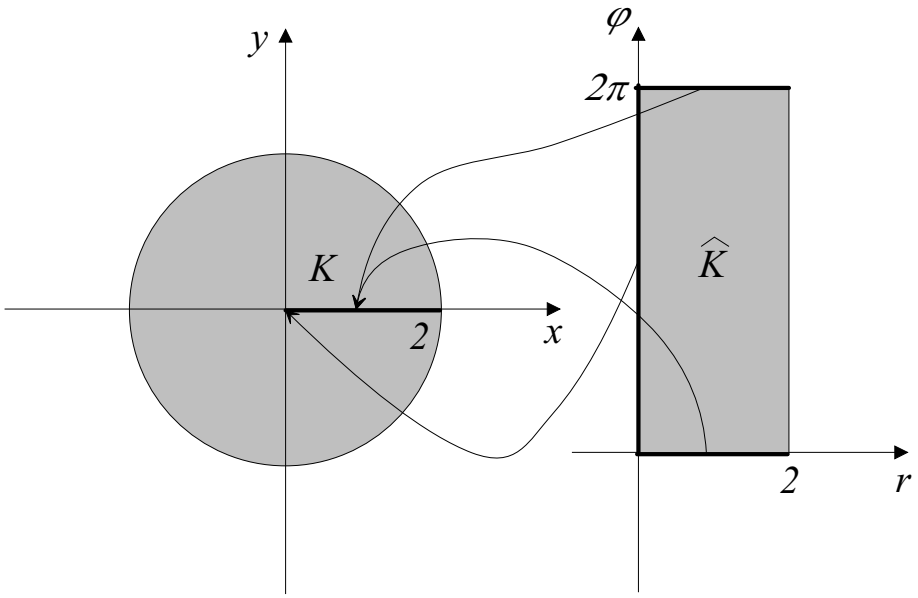
$$J(r,\varphi) = \begin{vmatrix} x'_r(r,\varphi) & x'_\varphi(r,\varphi) \\ y'_r(r,\varphi) & y'_\varphi(r,\varphi) \end{vmatrix} = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r.$$

Mivel $r \geq 0$, ezért $|J(r,\varphi)| = r$.

3.2.13. Példa. Példaképpen kiszámítjuk az $f(x,y) = x^2+y^2$ (forgásparaboloid) grafikonja és az xy-síkbeli, origó középpontú, 2 sugarú körlap által meghatározott tartomány térfogatát. A térfogat (mivel $x^2+y^2 \geq 0$) az

$$\int_K (x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y$$

integrállal számítható ki, ahol K az origó közepű, 2 sugarú körlapot jelenti. Használjuk a síkbeli polárkoordinátás helyettesítést (lásd a 3.2.4. ábrát). Ekkor kölcsönösen egyértelmű



3.2.4. ábra. *Síkbeli polárkoordinátás helyettesítés.*

megfeleltetés létesíthető a $\widehat{K}$ téglalap belseje és a K körlap megfelelő pontjai között, továbbá a Jacobi-determináns nem nulla a téglalap belsejében. Az integrál kiszámításához alkalmazhatjuk tehát a 3.2.12. tételt.

$$\int_K (x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y = \int_{\widehat{K}} \Big(((r\cos\varphi)^2+(r\sin\varphi)^2)\cdot r\Big)\,\mathrm{d}r\,\mathrm{d}\varphi = \int_{\widehat{K}} r^3\,\mathrm{d}r\,\mathrm{d}\varphi.$$

A jobb oldalon álló integrál téglalapra vonatkozó integrál, amit a 3.2.4. tétel segítségével számíthatunk ki.

$$\int_{\widehat{K}} r^3\,\mathrm{d}r\,\mathrm{d}\varphi = \int_0^{2\pi}\int_0^2 r^3\,\mathrm{d}r\,\mathrm{d}\varphi = \left(\int_0^2 r^3\,\mathrm{d}r\right)\left(\int_0^{2\pi} 1\,\mathrm{d}\varphi\right) = \left[\frac{r^4}{4}\right]_0^2 [\varphi]_0^{2\pi} = 8\pi.$$

A keresett térfogat tehát 8π térfogategység. □

3.2.14. Példa. A most következő példában olyan integrált fogunk kiszámolni, melynek meghatározása helyettesítés nélkül nagyon hosszadalmas feladat lenne. Most is az $f(x,y) = x^2+y^2$ függvényt fogjuk integrálni, de most a $H = \{(x,y) \in \mathbb{R}^2 \mid 1 \leq x^2+y^2 \leq 4\}$ halmazon. A H halmaz az origó középpontú egy ill. kettő sugarú körök által határolt gyűrű, amelyen ha helyettesítés nélkül szeretnénk integrálni, akkor azt négy normáltartományra kellene felbontani, és azokon külön-külön elvégezni az integrálást. A

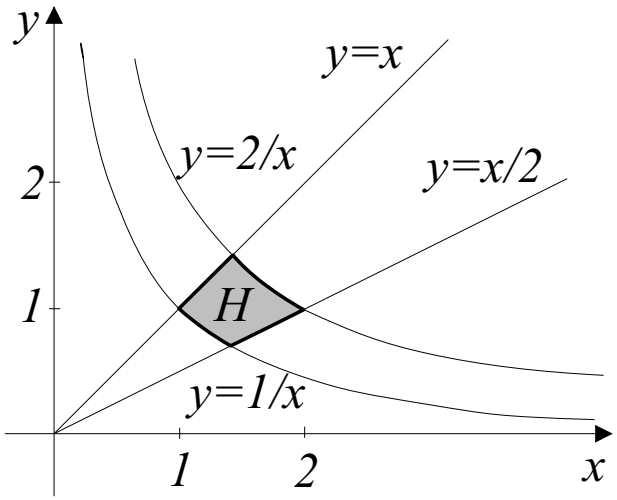
korábbi síkbeli polárkoordinátás helyettesítést használva az integrálást a $\widehat{H} = \{(r,\varphi) \in \mathbb{R}^2 \mid 1 \le r \le 2, 0 \le \varphi \le 2\pi\}$ téglalapon kell elvégezni.

$$\int_H (x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y = \int_{\widehat{H}} \Big(((r\cos\varphi)^2+(r\sin\varphi)^2)\cdot r\Big)\,\mathrm{d}r\,\mathrm{d}\varphi = \int_{\widehat{H}} r^3\,\mathrm{d}r\,\mathrm{d}\varphi =$$

$$= \int_0^{2\pi}\int_1^2 r^3\,\mathrm{d}r\,\mathrm{d}\varphi = \int_0^{2\pi}\left[\frac{r^4}{4}\right]_1^2 \mathrm{d}\varphi = \int_0^{2\pi}\frac{15}{4}\,\mathrm{d}\varphi = \left[\frac{15}{4}\varphi\right]_0^{2\pi} = \frac{15}{2}\pi.$$

A H gyűrű és az $f(x,y)$ forgásparaboloid grafikonja közti tartomány térfogata tehát $15\pi/2$. □

3.2.15. Példa. Határozzuk meg a 3.2.5. ábrán látható H tartomány területét! A



3.2.5. ábra. *A H tartomány szemléltetése.*

terület meghatározásához az $\int_H 1\,\mathrm{d}x\,\mathrm{d}y$ integrált kell kiszámítani. Az első lehetőség, hogy a H tartományt felbontjuk két x-szerinti normáltartomány uniójára, azokon integráljuk a konstans egy függvényt, majd az integrálokat összeadjuk. (Az terület meghatározható egyváltozós függvények integráljának segítségével is, de ebben az esetben is két tagra kell bontani az integrált.) A második lehetőség a helyettesítéses integrálás, hiszen vegyük észre, hogy ha az $u = x/y$ és $v = xy$ megfeleltetéseket használjuk, akkor a H halmaz oldalain az u és v értékek konstans egyet vagy kettőt fognak felvenni. Fejezzük ki az $u = x/y$ és $v = xy$ összefüggésekből x-et és y-t az u és v paraméterek segítségével. Azt kapjuk, hogy $x = \sqrt{vu}$ és $y = \sqrt{v/u}$. Látható, hogy az $x(u,v) = \sqrt{vu}$, $y(u,v) = \sqrt{v/u}$ transzformáció kölcsönösen egyértelmű módon képezi le a

$$\widehat{H} = \{(u,v) \in \mathbb{R}^2 \mid 1 \le u \le 2, 1 \le v \le 2\}$$

téglalapot a H halmazra. Számítsuk ki az $(u,v) \mapsto (\sqrt{vu}, \sqrt{v/u})$ transzformáció Jacobi-determinánsát.

$$J(u,v) = \begin{vmatrix} x'_u(u,v) & x'_v(u,v) \\ y'_u(u,v) & y'_v(u,v) \end{vmatrix} = \begin{vmatrix} \frac{v}{2\sqrt{vu}} & \frac{u}{2\sqrt{vu}} \\ \frac{-v}{2u^2\sqrt{v/u}} & \frac{1}{2u\sqrt{v/u}} \end{vmatrix} = \frac{1}{2u}.$$

Most áttérhetünk az integrál kiszámítására, ahol a helyettesítés után a $\widehat{H}$ téglalapon kell integrálnunk.

$$\int_H 1\,\mathrm{d}x\,\mathrm{d}y = \int_{\widehat{H}} 1\cdot\left|\frac{1}{2u}\right|\,\mathrm{d}u\,\mathrm{d}v = \int_1^2\int_1^2 \frac{1}{2u}\,\mathrm{d}u\,\mathrm{d}v =$$

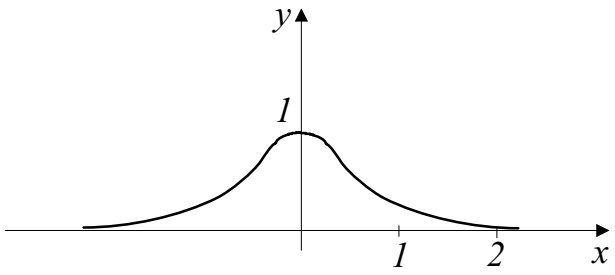
$$= \int_1^2 \frac{1}{2} \left[\ln u\right]_1^2 \, \mathrm{d}v = \int_1^2 \frac{1}{2} \ln 2 \, \mathrm{d}v = \frac{1}{2} \ln 2 \approx 0.34.$$

Láthatjuk, hogy a megoldás legbonyolultabb része igazából a Jacobi-determináns kiszámítása volt. Ha most megjegyezzük, hogy az $(u, v) \mapsto (\sqrt{vu}, \sqrt{v/u})$ transzformációhoz a $J(u, v) = 1/(2u)$ Jacobi-determináns tartozik, akkor ha legközelebb ugyanilyen transzformációval kell integrálnunk, akkor már nem kell kiszámolnunk a determináns értékét. Hasonlóan jártunk el a síkbeli polárkoordinátás transzformáció során is, ahol előre kiszámoltuk, hogy a Jacobi-determináns r, így a feladatok megoldása során már nem kellett meghatároznunk. □

3.2.16. Példa. A valószínűségszámítás egyik nevezetes eloszlása az ún. normális eloszlás. Erről a következő félévben tanulunk részletesen. A normális eloszlás vizsgálatában kulcsszerepe lesz az

$$\int_{-\infty}^{\infty} e^{-x^2} \, \mathrm{d}x$$

improprius integrálnak. Ez az integrál – amennyiben létezik – a $g(x) = e^{-x^2} \geq 0$ függvény grafikonja és az x-tengely közti tartomány területét adja meg (lásd 3.2.6. ábra). Annak ellenére, hogy g egyváltozós, a fenti improprius integrál egy kétváltozós függvény



3.2.6. ábra. *A $g(x) = e^{-x^2}$ függvény grafikonja.*

integráljának vizsgálatával számítható ki egyszerűen. Tekintsük ugyanis az $f(x, y) = e^{-(x^2+y^2)}$ függvényt, amely nyilvánvalóan folytonos, így tetszőleges $R > 0$ mellett integrálható a $T_R = T(-R, R, -R, R)$ téglalapon (négyzeten), valamint a K_R ill. $K_{\sqrt{2}R}$ origó közepű R ill. $\sqrt{2}R$ sugarú körlapokon. Mivel $K_R \subseteq T_R \subseteq K_{\sqrt{2}R}$,

$$\int_{K_R} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y \leq \int_{T_R} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y \leq \int_{K_{\sqrt{2}R}} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y. \tag{3.2.1}$$

A bal és a jobb oldalon álló két integrált síkbeli polárkoordinátákra áttérve könnyen kiszámíthatjuk.

$$\int_{K_R} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y = \int_0^{2\pi} \int_0^R e^{-r^2} r \, \mathrm{d}r \, \mathrm{d}\varphi = \left(\int_0^{2\pi} 1 \, \mathrm{d}\varphi \right) \cdot \left(\int_0^R e^{-r^2} r \, \mathrm{d}r \right) =$$

$$= 2\pi \cdot \left[\frac{-e^{-r^2}}{2} \right]_0^R = 2\pi \cdot \left(\frac{-e^{-R^2}}{2} + \frac{1}{2} \right) = \pi - \pi e^{-R^2},$$

hasonlóan

$$\int_{K_{\sqrt{2}R}} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y = 2\pi \cdot \left(\frac{-e^{-2R^2}}{2} + \frac{1}{2} \right) = \pi - \pi e^{-2R^2}.$$

A középső kifejezés a 3.2.4. tétel alkalmazásával az

$$\int_{T_R} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = \int_{-R}^{R}\int_{-R}^{R} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y =$$

$$= \left(\int_{-R}^{R} e^{-x^2}\,\mathrm{d}x\right)\cdot\left(\int_{-R}^{R} e^{-y^2}\,\mathrm{d}y\right) = \left(\int_{-R}^{R} e^{-x^2}\,\mathrm{d}x\right)^2$$

alakban írható. A nyert eredményeket a (3.2.1) képletbe visszahelyettesítve

$$\pi - \pi e^{-R^2} \le \left(\int_{-R}^{R} e^{-x^2}\,\mathrm{d}x\right)^2 \le \pi - \pi e^{-2R^2}$$

adódik. Az $R \to \infty$ határátmenetet elvégezve az egyenlőtlenség minkét oldala π-hez tart. Így a rendőr-elvből következik, hogy az $\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x$ improprius integrál létezik és

$$\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}.$$

□

3.2.4. Feladat. Igazoljuk, hogy tetszőleges $\mu \in \mathbb{R}$ és $\sigma > 0$ számok esetén

$$\int_{-\infty}^{\infty} \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right)\,\mathrm{d}x = 1.$$

Ellenőrző kérdések

1. Hogyan számítjuk ki kétváltozós függvények téglalapon vett integrálját?
2. Milyen esetekben számítható ki a kétváltozós integrál egyváltozós függvények integráljainak szorzataként?
3. Milyen halmazokat hívunk normáltartománynak?
4. Hogyan integrálunk normáltartományon?
5. Mit értünk egy transzformáció Jacobi-determinánsán? Hol játszik fontos szerepet a Jacobi-determináns?
6. Hogyan integrálunk kétváltozós függvényeket a helyettesítés módszerével?
7. Mekkora a síkbeli polárkoordináta transzformáció Jacobi-determinánsa!
8. Mekkora az $\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x$ improprius integrál és hogyan határoztuk meg?

A feladatok megoldásai

3.2.1.
Alkalmazzuk a 3.2.4. tételt az integrál kiszámítására.

$$\int_T \sin x\cdot\cos y\,\mathrm{d}x\,\mathrm{d}y = \left(\int_0^{\pi/2}\sin x\,\mathrm{d}x\right)\cdot\left(\int_0^{\pi/2}\cos y\,\mathrm{d}y\right) = [-\cos x]_0^{\pi/2}\cdot[\sin y]_0^{\pi/2} = 1.$$

3.2.2.
Most a 3.2.4. tétel nem alkalmazható, mert az integrálandó függvény nem áll elő egy x-től

és egy y-tól függő függvény szorzataként. Az integrál a 3.2.1. tétel segítségével számítható ki.

$$\int_T (x+y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^{\pi/2} \left[\frac{x^2}{2} + xy\right]_0^{\pi/2} \mathrm{d}y = \int_0^{\pi/2} \left(\frac{\pi^2}{8} + \frac{\pi y}{2}\right) \mathrm{d}y = \left[\frac{\pi^2 y}{8} + \frac{\pi y^2}{4}\right]_0^{\pi/2} = \frac{\pi^3}{8}.$$

3.2.3.
Az adott integrál egy y-szerinti normáltartományra vonatkozó integrált ad meg. Ha a tartományt x-szerinti normáltartománynak tekintjük, akkor a tartományt lentről a $h_1(x) = x^2$, fentről pedig a $h_2(x) = x$ függvények grafikonjai határolják, azaz az $N_x(0, 1, x^2, x)$ normáltartományon kell integrálni.

$$\int_0^1 \int_y^{\sqrt{y}} 1\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_{x^2}^{x} 1\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 [y]_{x^2}^{x}\,\mathrm{d}x = \int_0^1 (x - x^2)\,\mathrm{d}x = \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = \frac{1}{6}.$$

Az integrál annak az egységnyi magas egyenes hasábnak a térfogatát adja meg, amelynek az alapja az $N_x(0, 1, x^2, x)$ normáltartomány. Ez számértékben megegyezik a normáltartomány területével.

3.2.4.
Alkalmazzuk a $t = (x - \mu)/(\sqrt{2}\sigma)$ helyettesítést, majd a 3.2.16. példa eredményét.

$$\int_{-\infty}^{\infty} \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right) \mathrm{d}x = \int_{-\infty}^{\infty} \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-t^2} \cdot \sqrt{2}\sigma\right) \mathrm{d}t = \frac{1}{\sigma\sqrt{2\pi}} \cdot \sqrt{2}\sigma \cdot \int_{-\infty}^{\infty} e^{-t^2}\,\mathrm{d}t =$$

$$= \frac{1}{\sqrt{\pi}} \cdot \int_{-\infty}^{\infty} e^{-t^2}\,\mathrm{d}t = \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} = 1.$$

Ezt akartuk megmutatni.

Önálló munkára kijelölt feladatok: DGY 110. oldal 3-14. feladat, 110. oldal 15-25. feladat, 111. oldal 26-40. feladat, 113. oldal 43-48. feladat, 115. oldal 54-56. feladat.

4. fejezet

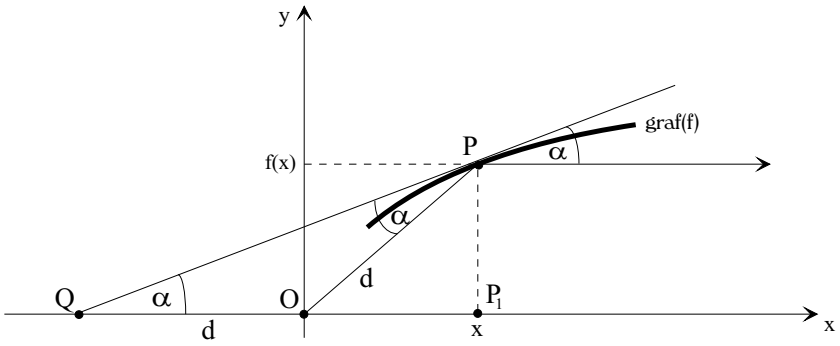
Differenciálegyenletek

4.1. Alapfogalmak

Jelen fejezetben – egy példát hívva segítségül – megismerkedünk a differenciálegyenletek osztályozásával és a megoldásukhoz kapcsolódó fogalmakkal.

Legyen $\mathcal{F}$ egy forgásfelület alakú tükör, amely a forgástengely egy rögzített pontjából érkező fénysugarakat a forgástengellyel párhuzamosan veri vissza. Mivel a fénysugár útja megfordítható, ez egyben azt is jelenti, hogy a szimmetriatengellyel párhuzamosan érkező sugarak egy pontban gyűlnek össze. Milyen lehet $\mathcal{F}$ alakja?

Helyezzük el a térbeli derékszögű koordináta-rendszerben a tükröt úgy, hogy a fénysugarak kiinduló pontja az O origó legyen, a tükör forgástengelye pedig az x-tengely. A forgásszimmetria miatt elegendő a feladatot az xy-síkban megoldani, melynek $\mathcal{F}$-fel való metszésvonalát tekintsük valamely $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x)$ függvény (a továbbiakban röviden csak $f(x)$ vagy f) grafikonjának. Indítsunk el az origóból egy tetszőleges fénysugarat, amely d távolság megtétele után a P pontban verődik vissza.



4.1.1. ábra. *A fénysugár útja.*

A fényvisszaverődés törvénye szerint a beesési és visszaverődési szögek (α) megegyeznek, ezek berajzolásához szükség van még az f függvény grafikonjához a P pontban húzott érintőegyenesre. Messe az érintőegyenes az x-tengelyt a Q pontban. Az általuk közbezárt szög szintén α, hiszen egyállású a visszaverődés szögével. A POQ háromszög egyenlő szárú, emiatt Q távolsága az origótól szintén d. Legyen a $P_1(x, 0)$ pont a $P(x, f(x))$ pont x-tengelyre eső merőleges vetülete. Alkalmazzuk Pithagorasz tételét az OP_1P derékszögű

háromszögre: $d^2 = x^2 + f^2(x)$, ahonnan

$$d = \sqrt{x^2 + f^2(x)}. \tag{4.1.1}$$

A QP_1P derékszögű háromszögből

$$\mathrm{tg}\alpha = \frac{f(x)}{d + x} \tag{4.1.2}$$

következik. A korábban tanultakból ismert, hogy a pontbeli differenciálhányados éppen a függvény grafikonjának adott pontjához húzott érintőegyenes iránytangensét adja meg. Tehát $f'(x) = \mathrm{tg}\alpha$. Figyelembe véve (4.1.1) és (4.1.2) eredményeit, az

$$f'(x) = \frac{f(x)}{\sqrt{x^2 + f^2(x)} + x} \tag{4.1.3}$$

egyenlőséget kapjuk. Mivel ez bármely origóból induló fénysugárra teljesül, pontosabban f grafikonjának minden $(x, f(x))$ pontjára, (4.1.3) úgy is felfogható, mint egy olyan egyenlet, amely az ismeretlen f függvényre vonatkozik. Ebből szeretnénk meghatározni f-et.

4.1.1. Definíció. **Differenciálegyenletnek** nevezzük az olyan függvényegyenletet, amelyben egy ismeretlen függvény és annak deriváltja(i), független változó(k), továbbá ismert függvények és konstansok fordulnak elő. ♢

Egy differenciálegyenlet lehet közönséges vagy parciális aszerint, hogy a benne szereplő ismeretlen függvény egy- vagy többváltozós. További osztályozásra nyújt lehetőséget az ismeretlen függvény legmagasabb rendű deriváltja. Így beszélhetünk elsőrendű, másodrendű, harmadrendű, stb. differenciálegyenletről. A fentiek alapján (4.1.3) egy elsőrendű közönséges differenciálegyenlete az ismeretlen f függvénynek, míg pl.

$$\frac{\partial^2 y(x,t)}{\partial t^2} = c^2 \frac{\partial^2 y(x,t)}{\partial x^2}$$

egy másodrendű parciális differenciálegyenlet (ún. hullámegyenlet), melyben $y(x,t)$ az ismeretlen (kétváltozós) függvénye az x és t független változóknak (c állandó). A jegyzetben csak első- és másodrendű közönséges differenciálegyenletekkel foglalkozunk, ezek rendre az $F(y', y, x) = 0$ ill. $F(y'', y', y, x) = 0$ formulákkal írhatók le általánosan, ahol F egy tetszőleges három- ill. négyváltozós függvény. Itt, és a későbbiekben is az ismeretlen függvényt – a hagyományos írásmódnak megfelelően – y jelöli, melynek független változója x.

4.1.1. Feladat. Írjunk példát elsőrendű parciális differenciálegyenletre, majd másod- ill. harmadrendű közönséges differenciálegyenletre!

Térjünk most vissza a fejezet elején felvetett problémához. A (4.1.3) differenciálegyenlet

$$f'(x) = \frac{f(x) \cdot \left(\sqrt{x^2 + f^2(x)} - x\right)}{(x^2 + f^2(x)) - x^2} = \frac{\sqrt{x^2 + f^2(x)} - x}{f(x)}$$

alakra hozható, ahonnan

$$\frac{2 \cdot (f'(x) \cdot f(x) + x)}{\sqrt{x^2 + f^2(x)}} = 2$$

következik. Mindkét oldalt integrálva az x változó szerint a kapott eredmények csak egy c_1 konstansban különbözhetnek, amelyet a jobb oldalon fogunk feltüntetni: $\int 2\,dx = 2x + c_1$, $(c_1 \in \mathbb{R})$. A bal oldal integrálásához vegyük észre, hogy az

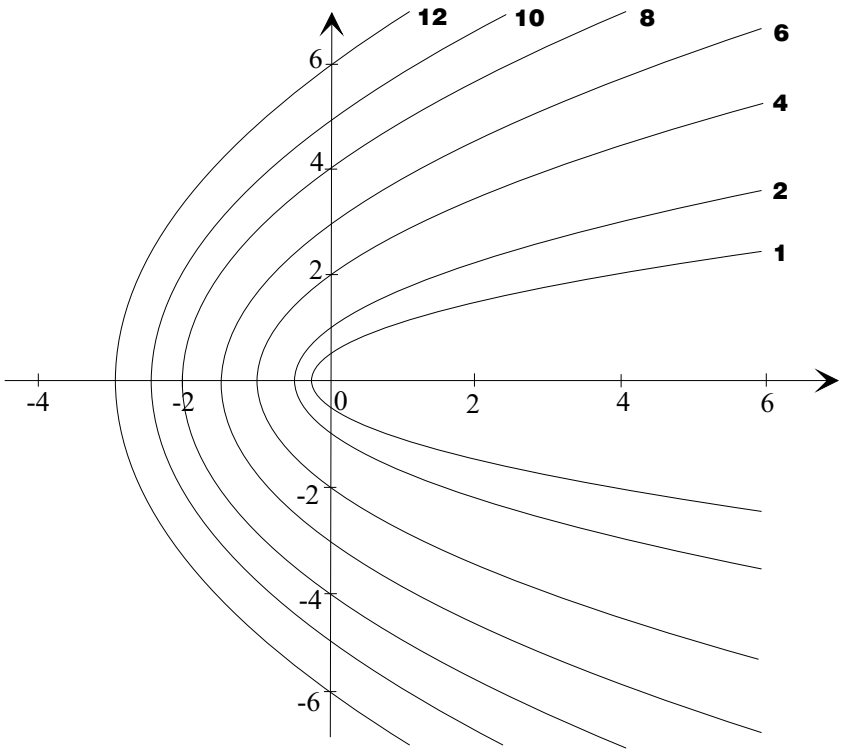
$$(2x + 2f'(x)f(x)) \cdot (x^2 + f^2(x))^{-1/2}$$

alakban is írható. Ezt elemző szemmel nézve, nem nehéz rájönni, hogy egy olyan szorzatról van szó, ahol az egyik tényező egy hatvány ($(x^2 + f^2(x))^{-1/2}$), a másik pedig az előző belső függvényének a deriváltja $(2x + 2f'(x)f(x))$. Alkalmazva az $\int g'g^n = g^{n+1}/(n+1)$ összefüggést, a bal oldal primitív függvénye $2\sqrt{x^2 + f^2(x)}$, tehát

$$2\sqrt{x^2 + f^2(x)} = 2x + c_1 \quad (c_1 \in \mathbb{R}).$$

Innen négyzetre emeléssel és rendezéssel

$$f(x) = \pm\sqrt{c_1 x + \left(\frac{c_1}{2}\right)^2}. \tag{4.1.4}$$



4.1.2. ábra. *A megoldások görbeserege.*

A kapott eredményt megvizsgálva megállapíthatjuk, hogy $c_1 = 0$ esetén f a konstans nulla függvény lenne, amely nem ad megoldást az eredeti optikai problémára, továbbá, hogy a $c_1 = c > 0$ és a $c_1 = -c < 0$ esetek grafikonjai szimmetrikusak az y-tengelyre. Elég tehát a $c_1 > 0$ esettel foglalkozni a kapott megoldások áttekintésekor. Ha a $\sqrt{c_1 x + (c_1/2)^2}$ és $-\sqrt{c_1 x + (c_1/2)^2}$ függvények grafikonját egy ábrán tüntetjük fel valamely pozitív c_1 esetén, akkor az $y^2 = c_1 x + (c_1/2)^2$ egyenletű parabola grafikonjához jutunk. Világos, hogy

ilyen parabola végtelen sok van. Összességében azt mondhatjuk, hogy a (4.1.3) differenciálegyenlet összes megoldását (4.1.4) írja le, amelyeket ábrázolva egy görbesereget kapunk. (A 4.1.2. ábrán látható paraboláknál feltüntettük a hozzájuk tartozó c_1 értékeket.)

A keresett $\mathcal{F}$ forgásfelület tehát forgásparaboloid alakú, amely a tengelyével párhuzamosan érkező fénysugarakat a fókuszpontban (origó) gyűjti össze. Ez a tulajdonság jelenik meg parabolaantennáknál elektromágneses hullámokra.

Egy differenciálegyenletet, mint függvényegyenletet kielégítő (kellően sokszor deriválható) függvényt a differenciálegyenlet egy megoldásának nevezzük. A cél az, hogy megadjuk a differenciálegyenlet összes megoldását. A megoldás szó használatos még a megoldás menetének rövidítésére is, tehát arra az algoritmusra amellyel előállítjuk az összes megoldást vagy annak egy részhalmazát. Most az összes megoldás néhány speciális részhalmazát értelmezzük.

4.1.2. Definíció. Egy n-edrendű közönséges differenciálegyenlet **általános megoldása** az a függvény, amely kielégíti a differenciálegyenletet (tehát megoldás) és pontosan n darab tetszőleges, egymástól független állandót (paramétert) tartalmaz. $\diamondsuit$

4.1.3. Definíció. Egy n-edrendű közönséges differenciálegyenlet **partikuláris megoldásán** azt a függvényt értjük, amely kielégíti a differenciálegyenletet (tehát megoldás) és legfeljebb $n-1$ darab tetszőleges, egymástól független állandóval írható fel. $\diamondsuit$

4.1.4. Példa. Az $y''(x) = 24x^2$ másodrendű differenciálegyenletben (röviden $y'' = 24x^2$) $y = y(x)$ az ismeretlen függvény. Kétszeri integrálással

$$y(x) = 2x^4 + c_1 x + c_2 \quad (c_1, c_2 \in \mathbb{R}) \tag{4.1.5}$$

következik. (4.1.5) általános megoldás, mert két, egymástól függetlenül tetszőlegesen megválasztható valós számot tartalmaz. Az általános megoldás grafikonja egy kétparaméteres görbesereg lesz. Egy partikuláris megoldáshoz jutunk, ha az általános megoldásból valamilyen ún. kezdeti feltételekkel kiválasztunk egy konkrét függvényt.

Legyenek pl. az $y(1) = -1$ és az $y'(0) = 2$ egyenlőségek a kezdeti feltételek. Ekkor c_1 és c_2 konkretizálásához a

$$\left.\begin{array}{rcrcr} -1 & = & y(1) & = & 2 + c_1 + c_2 \\ 2 & = & y'(0) & = & c_1 \end{array}\right\}$$

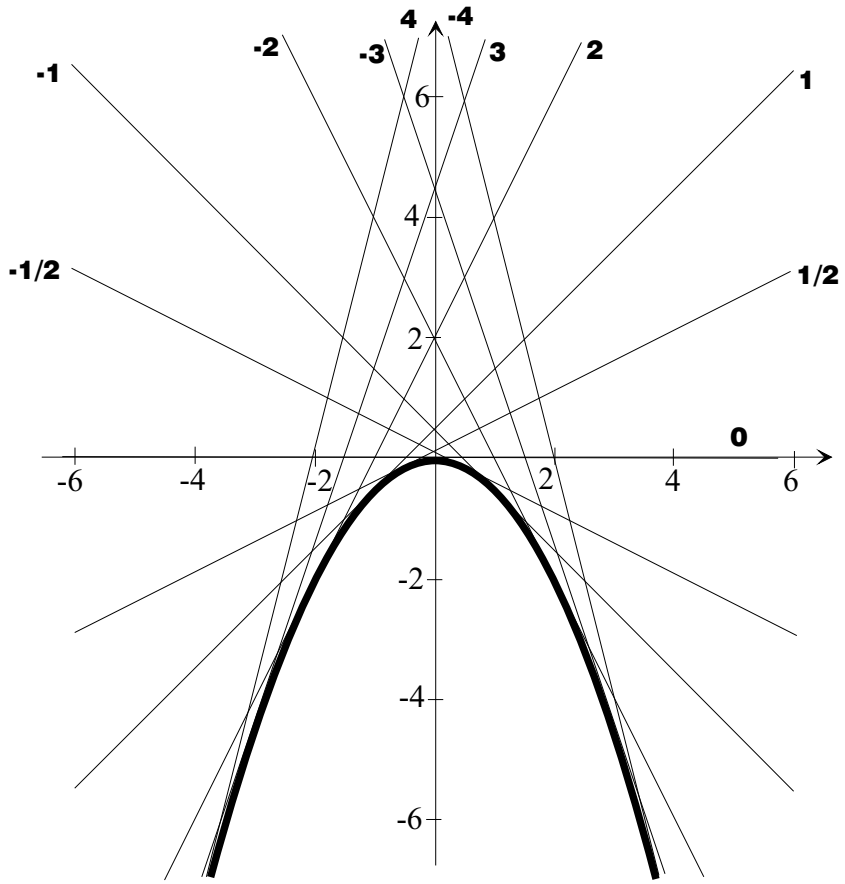
egyenletrendszert kell megoldani. Mivel a rendszer megoldása $c_1 = 2$ és $c_2 = -5$, így a keresett partikuláris megoldás $y(x) = 2x^4 + 2x - 5$. Ez a függvény tehát kielégíti a differenciálegyenletet, továbbá az $y(1) = -1$ és $y'(0) = 2$ feltételeket is. $\square$

Ha egy differenciálegyenlet mellé még a kezdeti feltétel(ek) is adott(ak), akkor kezdetiérték problémáról beszélünk.

Az előző példában egy partikuláris megoldást állítottunk elő az általános megoldásból kiindulva. Egyes differenciálegyenleteknél azonban ilyen módon nem kapható meg az összes partikuláris megoldás. Ha az $(x_0, y(x_0))$ ponton több különböző megoldás grafikonja is átmegy, akkor azt mondjuk, hogy a kezdetiérték probléma nem oldható meg egyértelműen. Azt a megoldást, melynek egyik pontjában sem egyértelmű a kezdetiérték probléma szinguláris megoldásnak nevezzük. A szinguláris megoldások nem származtathatók a differenciálegyenlet általános megoldásában szereplő konstans(ok) speciális megválasztásával,

vizsgálatuk több nehézséget is rejt. Mi egy példa erejéig foglalkozunk velük a könnyebb érthetőség kedvéért, egyébként az általános megoldások meghatározását tekintjük feladatnak.

4.1.5. Példa. Vizsgáljuk az $(y'(x))^2+2xy'(x)-2y(x)=0$ (röviden és áttekinthetőbben $(y')^2+2xy'-2y=0$) elsőrendű egyenletet. Könnyű ellenőrizni, hogy az általános megoldása $y(x)=cx+c^2/2$ ($c\in\mathbb{R}$). Másrészt az $y_s(x)=-x^2/2$ függvény kielégíti az adott differenciálegyenletet, és y_s nem tartozik bele az általános megoldásba, azaz y_s szinguláris megoldás. Megjegyezzük, hogy az $y_s(x)=-x^2/2$ függvény az $y(x)=cx+c^2/2$ egyenletű egyparaméteres egyenesseregnek az ún. burkolója. (Lásd 4.1.3. ábra, amelyen a megfelelő c értékeket is feltüntettük.) □



4.1.3. ábra. *Szinguláris megoldás és az általános megoldás.*

Később bizonyos típusú első és másodrendű közönséges differenciálegyenletek megoldási módszereit tárgyaljuk anélkül hogy magával a megoldhatóság kérdésével foglalkoznánk. Megemlítjük, hogy az ismeretlen függvényt legtöbbször $y(x)$-szel jelöljük, vagy a differenciálegyenlet áttekinthetősége miatt inkább csak röviden y-nal, míg a független változót x-szel. Gyakran kell olyan függvényeket használnunk – elsősorban a fent említett típusok általános leírásánál –, amelyek argumentuma maga az ismeretlen függvény, vagy a független változó, esetleg mindkettő. Ezeket a latin ábécé – y-tól és x-től különböző kis- vagy nagybetűivel azonosítjuk (lásd pl. a 4.1.1. feladat felett az első ill. másodrendű közönséges differenciálegyenletek általános alakjánál.)

Az egyes részfejezetek felépítése egyszerű: elsőként mindig az adott differenciálegyen-

let-típus ismérveit, majd megoldási módszerét ismertetjük. Ezután egy példával szemléltetjük az eljárást.

Ellenőrző kérdések

1. Mit értünk egy differenciálegyenlet általános, partikuláris ill. szinguláris megoldásán?
2. Mi a különbség a közönséges és parciális differenciálegyenletek között?
3. Mit jelent az, hogy egy differenciálegyenlet n-edrendű?
4. Mit értünk kezdetiérték problémán?

A feladat megoldása

4.1.1.
Egy-egy példa:

$$\frac{\partial f(x_1,x_2)}{\partial x_1}+\frac{\partial f(x_1,x_2)}{\partial x_2}=f(x_1,x_2),\quad y''-(\sin x)y+1=3e^x,\quad \operatorname{tg}(y'''-y)+2y'=-7.$$

4.2. Elsőrendű differenciálegyenletek

A most következő részben néhány elsőrendű közönséges differenciálegyenlet-típussal és az általános megoldásukat szolgáltató módszerekkel ismerkedünk meg. A tárgyalt típusokat egy-egy mintapélda megoldásával szemléltetjük.

4.2.1. Szétválasztható változójú differenciálegyenletek

Szétválasztható változójúnak nevezünk egy elsőrendű közönséges differenciálegyenletet, ha

$$y'(x) = f(x) \cdot g(y(x)) \quad , \tag{4.2.1}$$

vagy röviden

$$y' = f(x) \cdot g(y) \tag{4.2.2}$$

alakra hozható. A (4.2.2) formulában $f(x)$ adott függvénye az x független változónak, $g(y)$ az ismeretlen $y = y(x)$ függvény függvénye, továbbá az ismeretlen függvény y' deriváltjában lineáris az egyenlet.

Az általános megoldás meghatározásához rendezzük át az (4.2.2) egyenletet úgy, hogy az egyik oldalon az $f(x)$ függvény legyen, a másik oldalon pedig az ismeretlen y függvényt tartalmazó kifejezések:

$$\frac{y'}{g(y)} = f(x) \quad (g(y) \neq 0).$$

Ezután integráljuk mindkét oldalt az x változó szerint, az így kapott eredmények egy konstans összeadandóban különböznek egymástól, amelyet majd a jobb oldalon fogunk feltüntetni.

$$\int \frac{y'}{g(y)}\, dx = \int f(x)\, dx \tag{4.2.3}$$

jobb oldalának kiszámítása (és a kiszámítás nehézsége is) nyilvánvalóan az $f(x)$ függvényen múlik. A bal oldalon hajtsuk végre a $t = y(x)$ helyettesítést. Ekkor a helyettesítéses integrálás szabálya szerint a

$$\int \frac{1}{g(t)}\, dt$$

integrált kell kiszámítani. Ennek megtörténte után visszaírjuk t helyére $y(x)$-et (röviden y-t), és (ha lehetséges) kifejezzük az ismeretlen függvényt, hiszen (4.2.3) mindkét oldala ismert.

4.2.1. Példa. Egy szövetkezeti halastóban a halak össztömegét (valamely alkalmasan választott mértékegységben) az $y = y(x)$ függvény írja le az $x \geq 0$ időváltozó függvényében, amely kielégíti az

$$y' = (1 - y)y - 1/4$$

differenciálegyenletet. Kezdetben a betelepített halak tömege $y(0) = m_0 > 0$ volt. Hogyan változik hosszú távon a tóban élő halak tömege?

Mielőtt megoldjuk a differenciálegyenletet, ellenőrizhetjük, hogy valóban szétválasztható változójú: látható, hogy – használva a (4.2.2)-beli jelöléseket –, $f(x)$ a konstans 1 függvény, valamint $g(y) = (1 - y)y - 1/4 = -(y - 1/2)^2$. A megoldási eljárásban javasolt átosztás után $y \neq 1/2$-et feltételezve

$$\int \frac{y'}{(1-y)y - \frac{1}{4}} = \int -\left(y - \frac{1}{2}\right)^{-2} y'\, dx = \int 1\, dx,$$

melynek bal oldalán elvégezve az $y(x) = t$ helyettesítést

$$\int -\left(y - \frac{1}{2}\right)^{-2} y' \, dx = \int -\left(t - \frac{1}{2}\right)^{-2} \, dt = \left(t - \frac{1}{2}\right)^{-1}.$$

A jobb oldalon egyszerű a helyzet: $\int 1 \, dx = x + c$ ($c \in \mathbb{R}$). Visszahelyettesítve t helyére $y(x)$-et az

$$\frac{1}{y(x) - \frac{1}{2}} = x + c$$

egyenlőséghez jutunk, ahonnan megkapjuk a halak tömegét leíró differenciálegyenlet általános megoldását:

$$y(x) = \frac{1}{2} + \frac{1}{x + c} \quad , \quad (c \in \mathbb{R}).$$

Ez valóban általános megoldás, mert a differenciálegyenlet elsőrendű, és az előző formula tartalmaz egy szabadon választható c paramétert.

Ezután a kezdeti feltétellel meghatározott partikuláris megoldás kiválasztására kerül sor.

$$m_0 = y(0) = \frac{1}{2} + \frac{1}{c} > 0$$

automatikusan teljesül, ha $c > 0$. Ebben az esetben $m_0 > 1/2$. Ha c negatív, akkor az m_0-ra vonatkozó pozitivitási feltételből $c < -2$ következik, ekkor $0 < m_0 < 1/2$. Ha eltekintünk attól, hogy a megoldott differenciálegyenlet milyen gyakorlati problémát ír le, akkor az általános megoldásban $c = 0$ is megengedhető. Most azonban ezt az esetet ki kell zárni, mivel $m_0 = y(0)$-nak nem lenne értelme. Tehát $c \neq 0$, és a korábbiakból az is kiderült, hogy $m_0 \neq 1/2$. A kezdeti feltételből c-re

$$c = \frac{1}{m_0 - \frac{1}{2}} = \frac{2}{2m_0 - 1}$$

adódik, azaz a kívánt partikuláris megoldás

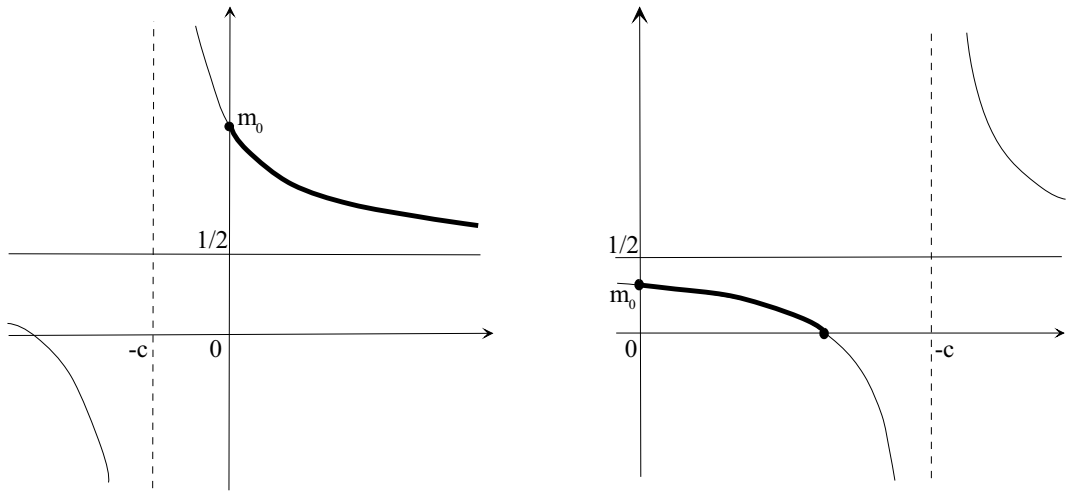
$$y_p(x) = \frac{1}{2} + \frac{1}{x + \frac{2}{2m_0 - 1}}.$$

Végül vizsgáljuk meg, hogy mit mondhatunk hosszú távon a tóban levő halak tömegéről. Ha $x \to \infty$ akkor $y_p \to 1/2$, tehát első látásra úgy tűnik, hogy a halak tömege függetlenül az $x = 0$ időpontban betelepített tömegtől 1/2-hez tart, más szóval egy egyensúlyi állapot felé közelít. Ez a benyomás módosul, ha pontosabb vizsgálatot végzünk, ehhez ábrázoljuk a két fő eset ($c \geq 0, c < -2$) grafikonját $x \geq 0$ esetén.

Az $y(x)$ függvény értelmezési tartománya $\mathbb{R} \setminus \{-c\}$, a függvénynek szakadása van az $x = -c$ helyen. Ha $c > 0$, akkor a szakadási hely az origótól balra helyezkedik el, és a halak tömege szigorúan monoton csökkenve tart az egyensúlyi állapot felé, mivel $x \geq 0$ (lásd 4.2.1. ábra, bal oldali grafikon). Más a helyzet $c < -2$ esetén, mert a szakadási hely az origótól jobbra található, és az 4.2.1. ábra jobb oldalán látható módon a halak tömege m_0-ról szigorúan monoton csökken, majd az

$$x_0 = -c - 2 = -\frac{2}{2m_0 - 1} - 2$$

időpontban kihal az állomány. Így hosszú távon eltűnnek a halak a tóból. Ha pl. a kezdeti tömeg $m_0 = 0.3$ volt, akkor $x_0 = 3$ időegység telik el a megsemmisülésig.



4.2.1. ábra. $y(x) = 1/2 + 1/(x + c)$ *grafikonja* $c > 0$ *ill.* $c < -2$ *esetén.*

Végül tekintsük az $y(x) = 1/2$ esetet. Az általános megoldás meghatározása közben ezt a lehetőséget kizártuk, mert a nevezőben $(y-1/2)^2$ szerepelt. Tehát csak a módszer miatt történt $y(x) \neq 1/2$ feltételezése, és behelyettesítéssel meggyőződhetünk, hogy $y(x) = 1/2$ is megoldása a differenciálegyenletnek. Az is nyilvánvaló, hogy bárhogyan is választjuk meg a c konstansot az általános megoldásban, sosem jutunk az $y(x) = 1/2$-hez, azaz $y(x) = 1/2$ szinguláris megoldás. Ez a külön álló eset azt fejezi ki, hogy $m_0 = 1/2$ tömegű halpopuláció betelepítése esetén a halak tömege az idők végezetéig tartja az egyensúlyi állapotot. □

Vegyük most újból szemügyre a 4.2.1. példa megoldási módszerét. Ebben minden lépést külön indokoltunk a jobb érthetőség kedvéért. Most leírunk egy egyszerűsítési lehetőséget, amelyet nem lett volna célszerű mindjárt bevetni, mert zavarhatta volna az eljárás megértését.

A bal oldal integrálásánál $t = y(x)$ helyettesítést hajtottuk végre, utána a t változó szerint integráltunk, majd az integrálás után t helyére visszahelyettesítettük az $y(x)$ függvényt. Ezen elv birtokában megspórolhatjuk a t változót (és ezzel a visszahelyettesítést), ha az integrálásban az $y(x)$ függvény helyett y-t változónak tekintjük, és az y' szorzót egyszerűen elhagyjuk. Az y változó szerinti integrálás végeztével y-t ismét egy x-től függő $y(x)$ függvénynek fogjuk fel. Tehát

$$\int -\left(y - \frac{1}{2}\right)^{-2} y' \, dx = \int -\left(y - \frac{1}{2}\right)^{-2} \, dy = \left(y - \frac{1}{2}\right)^{-1},$$

majd innen az $\int 1 \, dx = x + c$ összefüggést figyelembe véve

$$\frac{1}{y(x) - \frac{1}{2}} = x + c \,,$$

és visszatértünk a rendes kerékvágásba.

Röviden szólunk a szétválasztható változójú differenciálegyenletek megoldási algoritmusa leírásának egy – elsősorban régebbi tankönyvekben használatos – módjáról. Ehhez elevenítsük fel a differenciálhányadosra használt $y'(x) = \frac{dy}{dx}$ jelölést. A 4.2.1. példa egyenlete

$$\frac{dy}{dx} = -\left(y - \frac{1}{2}\right)^2$$

formájúvá alakul. $\frac{dy}{dx}$ egy jelölés, tehát nem egy tört, emiatt matematikailag nem korrekt, de mégis használható (mert ugyanazt a helyes eredményt szolgáltatja) ha azt mondjuk, hogy "szorozzunk át dx-szel", majd rendezzük úgy az egyenletet, hogy dy oldalán az y-os tagok, dx oldalán az x-es tagok legyenek szorzótényezők (innen származik a szétválasztható változójú elnevezés), majd integráljuk mindkét oldalt a megfelelő változók szerint:

$$-\left(y-\frac{1}{2}\right)^{-2}\,dy = dx \quad \Longrightarrow \quad \int -\left(y-\frac{1}{2}\right)^{-2}\,dy = \int dx.$$

Az integrálásokat elvégezve az

$$\left(y-\frac{1}{2}\right)^{-1} = x + c, \quad (c \in \mathbb{R})$$

egyenlőséghez jutunk, ahonnan y-t ismét $y(x)$ függvénynek tekintve kifejezhetjük a differenciálegyenlet általános megoldását.

4.2.1. Feladat. Igazoljuk, hogy az $y' - y = 2$ differenciálegyenlet szétválasztható változójú, majd adjuk meg az általános megoldását. Ezután határozzuk meg azt a partikuláris megoldást, melyre $y(0) = 2003$.

4.2.2. Szétválasztható változójúra visszavezethető differenciálegyenletek

A következőkben két olyan differenciálegyenlet-típussal ismerkedünk meg, amelyeket megfelelő transzformáció alkalmazásával szétválasztható változójúvá lehet alakítani.

Elsőként tekintsük az

$$y'(x) = f(ay(x) + bx + c) \quad , \tag{4.2.4}$$

vagy röviden

$$y' = f(ay + bx + c) \tag{4.2.5}$$

differenciálegyenletet, ahol $a \neq 0$, b és c valós számok, $y = y(x)$ az ismeretlen függvény, x a független változója, továbbá f valamely adott függvénye az $ay + bx + c$ kifejezésnek.

Legyen $z(x) = ay(x) + bx + c$ vagy rövidebben $z = ay + bx + c$. Ebből deriválással $z' = ay' + b$, majd az ismeretlen y függvény deriváltját kifejezve $y' = (z' - b)/a$-hoz jutunk. Az ismeretlen z függvényre a kiindulási differenciálegyenlettel ekvivalens

$$\frac{z' - b}{a} = f(z)$$

egyenlet teljesül, amely megfelelő rendezéssel

$$\frac{z'}{af(z) + b} = 1$$

szétválasztható változójúvá alakult. Ezt a korábban ismertetett módszerrel $z(x)$-re megoldjuk, és

$$y(x) = \frac{z(x) - bx - c}{a}$$

megadja a kezdeti differenciálegyenlet általános megoldását.

4.2.2. Példa. Keressük meg az $y' = 2y + x + 1$ differenciálegyenlet általános megoldását, majd azt a partikuláris megoldást, melyre $y(5) = 1$ teljesül!

A vizsgált differenciálegyenlet valóban (4.2.5) alakú, hiszen itt $f(ay+bx+c) = (2y+x+1)$. Legyen $z = 2y + x + 1$, így $z' = 2y' + 1$, és az eredeti differenciálegyenletbe való behelyettesítésből adódóan a z függvényre a

$$\frac{z' - 1}{2} = z \tag{4.2.6}$$

differenciálegyenlet vonatkozik. Ebből $z' = 2z + 1$, majd $z'/(2z + 1) = 1$ $(z \neq -1/2)$, és végül

$$\int \frac{1}{2z+1} z' \, dx = \int 1 \, dx.$$

Az integrálásokat elvégezve

$$\frac{1}{2} \ln |2z + 1| = x + c_1, \quad (c_1 \in \mathbb{R}),$$

majd

$$|2z + 1| = e^{2x+2c_1} = c_2 e^{2x}, \quad (c_2 = e^{2c_1} \in \mathbb{R}^+)$$

következik. Mivel e^{2x} pozitív, a bal oldalon az abszolút értéktől úgy lehet megszabadulni, hogy c_2-re megengedjük a negatív számokat is: $2z + 1 = c_2 e^{2x}$, $(c_2 \in \mathbb{R} \setminus \{0\})$. Tehát

$$z = \frac{c_2 e^{2x} - 1}{2}, \quad (c_2 \in \mathbb{R} \setminus \{0\}). \tag{4.2.7}$$

Vegyük észre, hogy a korábban – a megoldási módszer miatt – kizárt $z = -1/2$ konstans függvény is megoldása a (4.2.6) egyenletnek. Ez a különálló eset beolvasztható a (4.2.7) megoldásba $c_2 = 0$ megengedésével.

Az eredeti differenciálegyenlet általános megoldása

$$y(x) = \frac{z(x) - x - 1}{2} = \frac{1}{2}\left(\frac{c_2 e^{2x} - 1}{2} - x - 1\right) = c_3 e^{2x} - \frac{1}{2}x - \frac{3}{4}, \quad (c_3 = \frac{c_2}{4} \in \mathbb{R}).$$

A kívánt partikuláris megoldást megkapjuk, ha $c_3 e^{10} - 5/2 - 3/4 = 1$-ből $c_3 = 17e^{-10}/4$-et kifejezzük, és visszaírjuk az általános megoldásba:

$$y_p(x) = \frac{17}{4e^{10}} e^{2x} - \frac{1}{2}x - \frac{3}{4}.$$

□

4.2.2. Feladat. Oldjuk meg az $y' - y = 2x$ differenciálegyenletet.

A másik szétválasztható változójúra visszavezethető differenciálegyenlet-típus az ún. homogén fokszámú differenciálegyenlet.

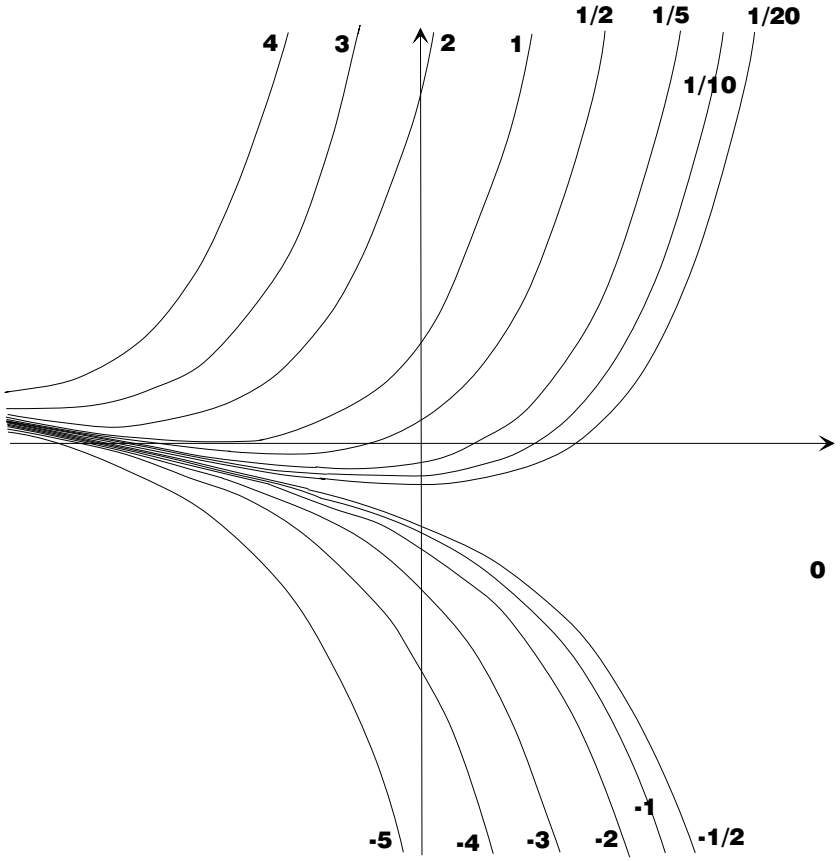
4.2.3. Definíció. Az $f(x, y)$ kétváltozós függvény **n-ed rendben homogén** ha tetszőleges $\lambda \in \mathbb{R}$ esetén

$$f(\lambda x, \lambda y) = \lambda^n f(x, y), \quad (n \in \mathbb{N}).$$

♢

4.2.4. Példa. Az $f(x, y) = x^4 + 5xy^3$ függvény 4-ed rendben homogén, mert

$$f(\lambda x, \lambda y) = (\lambda x)^4 + 5(\lambda x)(\lambda y)^3 = \lambda^4(x^4 + 5xy^3) = \lambda^4 f(x, y).$$



4.2.2. ábra. *Az* $y' = 2y + x + 1$ *differenciálegyenlet megoldásai.*

A $g(x, y) = x^2 + \sin xy$ függvény nem homogén, mert

$$g(\lambda x, \lambda y) = (\lambda x)^2 + \sin((\lambda x)(\lambda y)) = \lambda^2 x^2 + \sin \lambda^2 xy \neq \lambda^n (x^2 + \sin xy)$$

semilyen $n \in \mathbb{N}$ esetén. □

4.2.3. Feladat. Mutassuk meg, hogy a $h(x, y) = e^{x/y}$ függvény 0-ad rendben homogén!

Homogén fokszámúnak nevezünk egy differenciálegyenletet, ha

$$a(x, y(x)) + b(x, y(x))y'(x) = 0 \quad , \tag{4.2.8}$$

vagy röviden

$$a(x, y) + b(x, y)y' = 0 \tag{4.2.9}$$

alakú, és a kétváltozós $a(x, y)$, $b(x, y)$ függvények ugyanolyan rendben homogének ha y-t változónak tekintjük bennük. (A differenciálegyenletben $y = y(x)$ az ismeretlen függvény, x a független változója, $a(x, y)$ és $b(x, y)$ adott függvényei az x változónak és az y függvénynek.)

Homogén fokszámú differenciálegyenlet esetén a

$$z(x) = \frac{y(x)}{x}$$

helyettesítést kell alkalmazni (röviden $z = y/x$). Ez $y = z \cdot x$ formában is írható. Megjegyezzük, hogy bizonyos esetekben a $z = x/y$ helyettesítés előnyösebb lehet az előző verziónál. A módszer azért működik, mert (4.2.9) mindig

$$y' = f\left(\frac{y}{x}\right)$$

alakúra módosítható, majd a fenti helyettesítéssel $xz' + z = f(z)$ következik, ahonnan $z'x = f(z) - z$ alapján a szétválasztható változójú

$$\frac{1}{f(z) - z}\, z' = \frac{1}{x}$$

differenciálegyenletet kapjuk.

4.2.5. Példa. Oldjuk meg az $x^3 + y^3 - 3xy^2y' = 0$ differenciálegyenletet.

Könnyű ellenőrizni, hogy az $a(x, y) = x^3 + y^3$ és a $b(x, y) = -3xy^2$ függvények mindegyike 3-ad rendben homogén, azaz a differenciálegyenlet homogén fokszámú. x^3-nel végigosztva az egyenletet

$$1 + \left(\frac{y}{x}\right)^3 - 3\left(\frac{y}{x}\right)^2 y' = 0$$

adódik. Ebből $y = z \cdot x$ és $y' = z'x + z$ figyelembe vételével

$$1 + z^3 - 3z^2(z'x + z) = 1 - 2z^3 - 3xz^2z' = 0$$

következik. Rendezéssel jutunk a szétválasztható változójú

$$\frac{3z^2}{1 - 2z^3}\, z' = \frac{1}{x}\,, \quad (z \neq \frac{1}{\sqrt[3]{2}})$$

differenciálegyenlethez.

$$\int \frac{3z^2}{1 - 2z^3}\, dz = -\frac{1}{2}\int \frac{-6z^2}{1 - 2z^3}\, dz = -\frac{1}{2}\ln|1 - 2z^3| \quad , \quad \int \frac{1}{x}\, dx = \ln|x|.$$

A szokásos konstansot $c_1 = \ln c_2$ ($c_2 \in \mathbb{R}^+$) alakban fogjuk a jobb oldalon feltüntetni. Az intgrálások alapján az

$$\ln \frac{1}{\sqrt{|1 - 2z^3|}} = \ln|x| + \ln c_2 = \ln(c_2|x|)$$

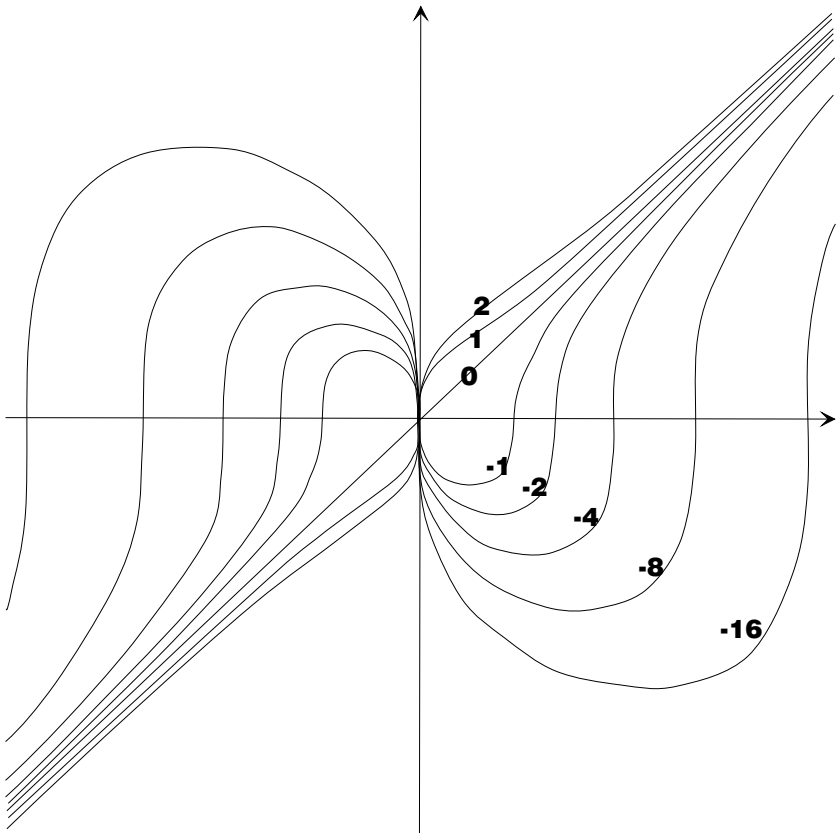
egyenlőséget kapjuk. A logaritmus függvény szigorúan monoton növekszik, tehát a két argumentum is megegyezik. Azokat négyzetre emelve

$$\frac{1}{|2z^3 - 1|} = c_2^2x^2$$

következik. Legyen $c_3 = c_2^2$. Ha c_3-ra negatív számokat is megengedünk, akkor az abszolút érték elhagyható, és z-t kifejezve, majd x-szel szorozva az

$$y(x) = x\sqrt[3]{\frac{1}{2c_3x^2} + \frac{1}{2}} = \sqrt[3]{c_4x + \frac{x^3}{2}} \tag{4.2.10}$$

általános megoldáshoz jutunk, ahol $c_4 = 1/(2c_3) \neq 0$. Vizsgáljuk meg, hogy a $z(x) = 1/\sqrt[3]{2}$ különleges esetből származó $y(x) = x/\sqrt[3]{2}$ függvény megoldása-e az eredeti differenciálegyenletnek. Behelyettesítéssel igenlő választ kapunk, és éppen ehhez a függvényhez



4.2.3. ábra. *Az* $x^3 + y^3 - 3xy^2y' = 0$ *differenciálegyenlet megoldásai.*

jutunk akkor is, ha a 0 értéket adnánk a (4.2.10) megoldásban c_4-nek. Tehát a differenciálegyenlet általános megoldása (szemléltetése a 4.2.3 ábrán látható)

$$y(x) = \sqrt[3]{c_4 x + \frac{x^3}{2}}\ , \quad (c_4 \in \mathbb{R}).$$

□

4.2.4. Feladat. Oldjuk meg a $2x + y + (x + y)y' = 0$ differenciálegyenletet.

4.2.3. Lineáris differenciálegyenletek

Egy elsőrendű közönséges differenciálegyenletet lineárisnak hívunk, ha

$$f(x)y'(x) + g(x)y(x) = h(x) \quad , \tag{4.2.11}$$

vagy röviden

$$f(x)y' + g(x)y = h(x) \tag{4.2.12}$$

alakú, ahol $y = y(x)$ az ismeretlen függvény, továbbá $f(x)$, $g(x)$ és $h(x)$ az x független változó adott függvényei, $f(x)$ nem a konstans nulla függvény. A lineáris differenciálegyenlet homogén ha $h(x)$ a konstans nulla függvény, ellenkező esetben inhomogén.

Kezdjük a vizsgálatokat a homogén esettel, amely szétválasztható. Ekkor $h(x) = 0$ miatt a differenciálegyenlet $f(x)y' = -g(x)y$ alakban írható, amelyből

$$\frac{y'}{y} = -\frac{g(x)}{f(x)}\,, \quad (y \neq 0)$$

következik. Innen

$$\int \frac{y'}{y}\,dx = -\int \frac{g(x)}{f(x)}\,dx\,,$$

ahol a bal oldal integrálja $\ln|y|$ lesz. A jobb oldalon levő integrál kiszámítása az $f(x)$ és $g(x)$ függvényektől függ. A logaritmust eltüntetve az egyenlet általános megoldása:

$$y(x) = e^{c_1 - \int \frac{g(x)}{f(x)}\,dx} = c_2 e^{-\int \frac{g(x)}{f(x)}\,dx}\,. \tag{4.2.13}$$

Az $f(x)y' + g(x)y = h(x)$ inhomogén esetben a differenciálegyenletből származtatható $f(x)y' + g(x)y = 0$ homogén egyenlet általános megoldásából indulunk ki. Jelöljük ezt y_H-val. Tegyük fel, hogy valamilyen módszerrel megtaláltuk az inhomogén egyenlet egy partikuláris megoldását, legyen ez y_p. Ekkor az eredeti, inhomogén egyenlet általános megoldását $y_I = y_H + y_p$ szolgáltatja. Ugyanis összeadva az

$$\left.\begin{array}{rcrcl} f(x)y_H' & + & g(x)y_H & = & 0 \\ f(x)y_p' & + & g(x)y_p & = & h(x) \end{array}\right\}$$

egyenlőségeket

$$h(x) = f(x)\left(y_H' + y_p'\right) + g(x)\left(y_H + y_p\right) = f(x)y_I' + g(x)y_I$$

következik, ami pontosan azt jelenti, hogy $y_I = y_H + y_p$ az inhomogén egyenlet általános megoldása, hiszen y_H kielégíti a differenciálegyenletet és tartalmaz egy szabadon megválasztható paramétert.

Felmerül a kérdés, hogyan lehetne megkapni egy y_p partikuláris megoldást. Lagrange francia matematikus a következő ötlettel állt elő. Mivel az inhomogén egyenlet és a belőle származtatott homogén egyenlet csak a $h(x)$ függvényben különböznek, ezért az inhomogén egyenlet egy partikuláris megoldása "nem sokban térhet el" a homogén egyenlet már meghatározott általános megoldásától. A módszer lényege abban áll, hogy a partikuláris megoldást

$$y_p(x) = c(x)e^{-\int \frac{g(x)}{f(x)}\,dx} \tag{4.2.14}$$

alakban keressük. Ha ezt összehasonlítjuk a (4.2.13) általános megoldással, akkor jól látható, hogy az ottani c_2 konstans szorzó helyett itt egy $c(x)$ egyelőre ismeretlen függvény áll. A módszert emiatt az állandók variálása módszerének hívják. A $c(x)$ függvényt az inhomogén egyenletbe való behelyettesítés után tudjuk kiszámolni. Konkrétan

$$y_p'(x) = e^{-\int \frac{g(x)}{f(x)}\,dx}\left(c'(x) - \frac{g(x)}{f(x)}\,c(x)\right)\,,$$

$$h(x) = f(x)y_p' + g(x)y_p = e^{-\int \frac{g(x)}{f(x)}\,dx}\left(f(x)c'(x) - g(x)c(x)\right) + e^{-\int \frac{g(x)}{f(x)}\,dx}g(x)c(x),$$

majd az egyszerűsítések után

$$h(x) = f(x)c'(x)e^{-\int \frac{g(x)}{f(x)}\,dx}$$

adódik. Innen $c'(x)$-et kifejezve, majd integrálva

$$c(x) = \int \frac{h(x)}{f(x)}\, e^{\int \frac{g(x)}{f(x)}\, dx}\, dx$$

következik. $c(x)$ meghatározása után az y_p partikuláris megoldást a (4.2.14) képlettel kapjuk meg.

A fentiek szerint egy inhomogén lineáris differenciálegyenlet általános megoldásának meghatározása az alábbi három lépésből áll.

1. Megkeressük az inhomogén differenciálegyenletből származtatott szétválasztható változójú homogén egyenlet általános megoldását.
2. Megkeressük az inhomogén egyenlet egy partikuláris megoldását a Lagrange-módszerrel.
3. Az előző két pontban meghatározott eredményeket összadjuk.

4.2.6. Példa. Számítsuk ki az

$$y' - \frac{y}{x} = x^2 + 3x - 2$$

inhomogén lineáris differenciálegyenlet általános megoldását. Elsőként az

$$y' - \frac{y}{x} = 0$$

homogén egyenlet általános megoldását határozzuk meg. Az egyenletet átrendezve $y'/y = 1/x$-et kapunk ($y \neq 0$) és mindkét oldalt integrálva az x változó szerint az $\ln|y| = \ln(c_1|x|)$ egyenlőséghez jutunk ($c_1 \in \mathbb{R}^+$). Innen a szokásos megfontolások után $y_H = c_2 x$ adódik.

Most az inhomogén egyenlet egy partikuláris megoldásának megkeresése következik, amelyet $y_p = c(x) \cdot x$ alakúnak tételezünk fel. $y_p' = xc'(x) + c(x)$ felhasználásával az inhomogén egyenletbe való behelyettesítés:

$$xc'(x) + c(x) - \frac{xc(x)}{x} = xc'(x) = x^2 + 3x - 2.$$

Tehát $c'(x) = x + 3 - 2/x$, ahonnan $c(x) = x^2/2 + 3x - 2\ln|x|$, majd

$$y_p = xc(x) = \frac{x^3}{2} + 3x^2 - 2x\ln|x|$$

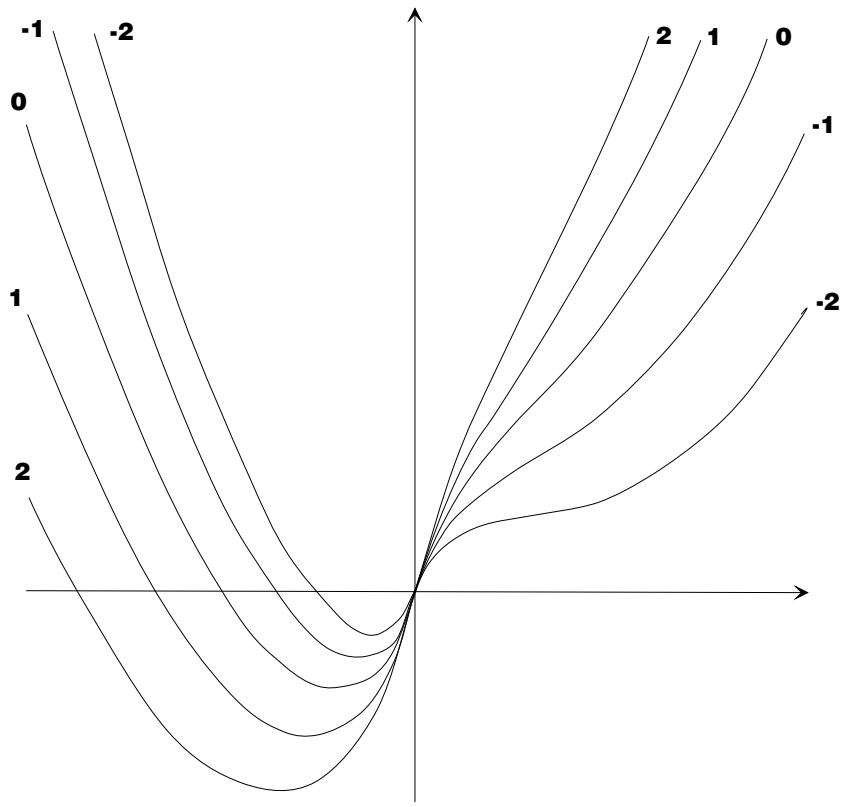
következik. $c(x)$-et integrálással kaptuk, de szándékosan nem írtunk hozzá konstans összeadandót, mert elegendő egyetlen $c(x)$ függvényt megadni.

A korábbi eredményeket felhasználva

$$y_I = c_2 x + \frac{x^3}{2} + 3x^2 - 2x\ln|x|\,, \quad (c_2 \in \mathbb{R})$$

adja az inhomogén egyenlet általános megoldását. □

4.2.5. Feladat. Oldjuk meg az $y'\sin x - y\cos x = 1$ differenciálegyenletet.



4.2.4. ábra. *Az* $y' - y/x = x^2 + 3x - 2$ *differenciálegyenlet megoldásai.*

Ellenőrző kérdések

1. Milyen elsőrendű differenciálegyenlet-típusok megoldási algoritmusát tanultuk?
2. Honnan lehet felismerni az egyes típusokat?
3. Mire szolgál a Lagrange-féle állandók variálásának módszere?

A feladatok megoldásai

4.2.1.
A szétválaszthatóság az eredeti alakból y hozzáadásával látható, vagy a rendezés utáni

$$\frac{y'}{y+2} = 1$$

alakból. Az általános megoldás $y(x) = ce^x - 2$, $(c \in \mathbb{R})$. A partikuláris megoldások közül a $2005e^x - 2$ felel meg a feltételnek.

4.2.2.
A probléma az előző feladatra vezethető vissza a $z = y + 2x$ helyettesítéssel. Az általános megoldás

$$y(x) = ce^x - 2x + 2\ .$$

4.2.3.
$h(\lambda x, \lambda y) = e^{\frac{\lambda x}{\lambda y}} = e^{\frac{x}{y}} = h(x, y) = \lambda^0 h(x, y)$.

4.2.4.
Az $y = z \cdot x$ helyettesítés az

$$\int \frac{z}{z^2 + 2z + 2}\, dz = \int -\frac{1}{x}\, dx$$

egyenlőségre vezet. Ebből

$$\frac{1}{2}\ln(z^2 + 2z + 2) - \operatorname{arctg}(z + 1) = c_1 - \ln|x|,$$

majd z helyére y/x-et írunk. Az így kapott egyenlőség végeredmény, mert az y függvényt nem tudjuk explicit módon kifejezni belőle. Időnként előfordulhat, hogy az általános megoldást csak ún. implicit módon tudjuk felírni.

4.2.5.
$y_I = c\sin x - \cos x$.

Önálló munkára kijelölt feladatok: DGY 74. oldal 38-43. feladat, 76. oldal 51-66. feladat.

4.3. Másodrendű differenciálegyenletek

E fejezetben izelítőt adunk a másodrendű közönséges differenciálegyenletek megoldási módszereiből. Részletesen a másodrendű lineáris differenciálegyenletekkel foglalkozunk. A tanultak szemléltetésére kidolgozott példákat közlünk.

A másodrendű közönséges differenciálegyenletek $F(y'', y', y, x) = 0$ általános alakja már szerepelt korábban. Tekinsük az alábbi nagyon egyszerű példát: $y'' = x$. Ennek általános megoldása két, egymástól függetlenül megválasztható paramétert tartalmaz. A megoldáshoz kétszeri integrálással juthatunk:

$$y(x) = \int \left(\int x \, dx \right) \, dx = \int \left(\frac{x^2}{2} + c_1 \right) \, dx = \frac{x^3}{6} + c_1 x + c_2, \quad (c_1, c_2 \in \mathbb{R}).$$

Ebből partikuláris megoldást úgy nyerhetünk, ha valamelyik, vagy akár mindkét paraméternek konkrét értéket adunk. Ezen értékek általában egy vagy két kezdeti feltétellel vannak meghatározva.

A másodrendű lineáris differenciálegyenlet általános alakja

$$f(x)y''(x) + g(x)y'(x) + h(x)y(x) = j(x), \tag{4.3.1}$$

vagy röviden

$$f(x)y'' + g(x)y' + h(x)y = j(x), \tag{4.3.2}$$

ahol $y = y(x)$ az ismeretlen függvény, $f(x) \neq 0$, $g(x)$, $h(x)$ és $j(x)$ az x független változó adott függvényei. (4.3.2) homogén ha $j(x)$ azonosan nulla, egyébként inhomogén. Amennyiben $f(x)$, $g(x)$ és $h(x)$ konstans függvények (és $f(x)$ nem a konstans nulla függvény), akkor állandó együtthatós differenciálegyenletről beszélünk. Vizsgálatunk tárgyát a másodrendű, állandó együtthatós lineáris differenciálegyenletek képezik. Mielőtt egy megoldási módszerük ismertetésére sor kerülne, röviden szólunk a megoldások szerkezetéről.

Először felelevenítjük a lineáris függetlenség definícióját. A véges dimenziós vektortereknél pontosan értelmeztük ezt a fogalmat, és beszéltünk arról is, hogy vannak végtelen dimenziós vektorterek. Ilyen például a valós függvények tere.

4.3.1. Definíció. Az $a(x)$ és $b(x)$ függvényeket **lineárisan függetlennek** nevezzük, ha az

$$\alpha a(x) + \beta b(x) = 0 \tag{4.3.3}$$

egyenlőségből $\alpha = \beta = 0 \in \mathbb{R}$ következik. ((4.3.3) jobb oldalán a konstans 0 függvény áll). Ellenkező esetben a két függvény **lineárisan összefüggő**. $\diamondsuit$

A fenti definíció két függvényre lett kimondva, mert ennyi elegendő a másodrendű differenciálegyenletek vizsgálatánál. Vegyük észre, hogy az összefüggőség azt jelenti, hogy az egyik függvény a másik konstansszorosa. A lineáris függetlenség vizsgálata a $W = W(x)$ ún. Wronski-féle determinánssal végezhető el könnyen.

4.3.2. Tétel. *A (4.3.2) differenciálegyenlet y_1 és y_2 megoldása akkor és csak akkor lineárisan független, ha a*

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_2 y_1'$$

Wronski-féle determináns nem az azonosan 0 függvény.

Most közöljük azt a tételt, amelyet felhasználunk a lineáris differenciálegyenletek megoldásánál.

4.3.3. Tétel.

1. *Ha y_1 és y_2 egy másodrendű homogén lineáris differenciálegyenlet két lineárisan független partikuláris megoldása, akkor az általános megoldás $y_H = c_1y_1+c_2y_2$ alakú, ahol $c_1, c_2 \in \mathbb{R}$.*

2. *Ha y_p egy másodrendű inhomogén lineáris differenciálegyenlet egy tetszőleges partikuláris megoldása, továbbá y_H az inhomogén egyenletből származtatott homogén egyenlet általános megoldása, akkor az inhomogén egyenlet általános megoldása $y_I = y_H + y_p$.*

A tétel lényegében azt mondja ki, hogy mindkét esetben partikuláris megoldások segítségével írható fel az általános megoldás. Vegyük észre, hogy mind y_1, y_2, mind y_p felírásában nem kell semmilyen konstansnak szerepelnie (elvileg lehetőség lenne egy-egy szabadon választható konstansra, mivel y_1, y_2 és y_p másodrendű közönséges differenciálegyenletek partikuláris megoldásai), mert y_H és így y_I is tartalmazza a megfelelő számú szabad paramétert. (Ha a partikuláris megoldások valamelyikében mégis lenne egy konstans, akkor a szóban forgó általános megoldás átalakítással két szabadon választható paraméterrel is felírható lenne.)

4.3.1. Állandó együtthatós lineáris differenciálegyenletek

(4.3.2) megadja a másodrendű lineáris differenciálegyenlet általános alakját. Az együtthatófüggvények rendre legyenek egyenlőek az $a \neq 0$, b és c konstans függvényekkel ($a, b, c \in \mathbb{R}$). Ilymódon a (4.3.2) alak az

$$ay'' + by' + cy = j(x) \tag{4.3.4}$$

formára redukálódik. Először a homogén esettel foglalkozunk.

Meglepő módon az $ay'' + by' + cy = 0$ differenciálegyenlet megoldásához nem kell integrálni, hanem egyszerűen meg kell oldani az

$$a\lambda^2 + b\lambda + c = 0$$

másodfokú, ún. karakterisztikus egyenletet, melynek együtthatói a differenciálegyenletből származnak. Jól ismert, hogy a valós gyökök számát illetően, a diszkriminánstól függően egy másodfokú egyenlet háromféle lehet. Mindegyik esetben más-más módon adhatók meg a kívánt partikuláris megoldások.

Ha a karakterisztikus egyenlet diszkriminánsa pozitív, akkor két különböző valós gyök van, legyenek ezek λ_1 és λ_2. Ekkor a homogén egyenletnek $y_1 = e^{\lambda_1 x}$ és $y_2 = e^{\lambda_2 x}$ két lineárisan független megoldása, hiszen $W(x) = (\lambda_1 - \lambda_2)e^{(\lambda_1+\lambda_2)x} \neq 0$. Így az általános megoldás:

$$y_H = c_1e^{\lambda_1 x} + c_2e^{\lambda_2 x}.$$

Ha a karakterisztikus egyenlet diszkriminánsa nulla, akkor a két valós gyök egybeesik, jelöljük ezt az értéket λ-val. A lineárisan független megoldások $y_1 = e^{\lambda x}$ és $y_2 = xe^{\lambda x}$ lesznek, következésképp az általános megoldásra

$$y_H = c_1e^{\lambda x} + c_2xe^{\lambda x}$$

adódik ($W(x) = e^{2\lambda x} \neq 0$).

Negatív diszkrimináns esetén a karakterisztikus egyenlet gyökeit egy komplex konjugált pár adja meg, legyen ez $\lambda \pm \mu i$. Ebben az esetben $y_1 = e^{\lambda x} \cos \mu x$, $y_2 = e^{\lambda x} \sin \mu x$, $W(x) = \mu e^{2\lambda x} \neq 0$, tehát

$$y_H = e^{\lambda x} \left(c_1 \cos \mu x + c_2 \sin \mu x \right) .$$

4.3.1. Feladat. Mutassuk meg, hogy ha $\lambda \in \mathbb{R}$ gyöke a karakterisztikus egyenletnek, akkor $e^{\lambda x}$ megoldása a másodrendű differenciálegyenletnek.

4.3.2. Feladat. Lássuk be, hogy a fenti három esetben valóban a Wronski-féle determinánst adtuk meg, és hogy közülük egyik sem azonosan 0.

Ezután rátérünk az inhomogén eset vizsgálatára. Az általános megoldás meghatározásához először az inhomogén egyenletből származtatott homogén egyenlet megoldását kell – az előbb ismertetett módon – meghatározni. A második lépésben meg kell keresni az inhomogén egyenlet egy partikuláris megoldását, itt újból a Lagrange-féle állandók variálása eljárást követjük. Ennek elvét most már nem részletezzük, az elsőrendű lineáris differenciálegyenleteknél elmondottak analógiájára cselekszünk.

A homogén egyenlet $y_H = c_1 y_1 + c_2 y_2$ általános megoldásából kiindulva az inhomogén egyenlet egy partikuláris megoldását

$$y_p = c_1(x) y_1 + c_2(x) y_2$$

alakban keressük. A feladat az egyelőre ismeretlen $c_1(x)$ és $c_2(x)$ függvények meghatározása. A partikuláris megoldás első deriváltja

$$y_p' = \underbrace{c_1(x) y_1' + c_2(x) y_2'} + \underbrace{c_1'(x) y_1 + c_2'(x) y_2} .$$

A deriváltat jól látható módon két részre bontottuk. Mivel elegendő egyetlenegy partikuláris megoldás meghatározása, és két ismeretlen függvény van, így két feltételt lehet előírni y_p számára. Az egyik nyilvánvalóan az, hogy elégítse ki az inhomogén differenciálegyenletet. A másik feltétel az lesz, hogy y_p' második része 0 legyen. Tehát

$$c_1'(x) y_1 + c_2'(x) y_2 = 0 , \tag{4.3.5}$$

így $y_p' = c_1(x) y_1' + c_2(x) y_2'$. A második derivált

$$y'' = c_1'(x) y_1' + c_2'(x) y_2' + c_1(x) y_1'' + c_2(x) y_2''.$$

Az inhomogén egyenletbe való behelyettesítéssel, rövid számolás után az

$$a c_1'(x) y_1' + a c_2'(x) y_2' = j(x) , \tag{4.3.6}$$

egyenlőséghez jutunk. (4.3.5) és (4.3.6) együttesen egy kétismeretlenes, két egyenletből álló lineáris egyenletrendszert jelent az ismeretlen $c_1'(x)$ és $c_2'(x)$ függvényekre. Ennek akkor van egyértelmű megoldása, ha a rendszer determinánsa – amely éppen a Wronski-féle determináns a-szorosa – nem zérus (amely $a \neq 0$ és a korábbiak szerint teljesül):

$$\begin{vmatrix} y_1 & y_2 \\ a y_1' & a y_2' \end{vmatrix} = a \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = a W(x) \neq 0 .$$

Ezáltal lehetőség nyílik arra, hogy a (4.3.5) és a (4.3.6) egyenletekből álló rendszert a Cramer-szabállyal megoldjuk.

$$\left\{ \begin{array}{l} c_1'(x) = \dfrac{\begin{vmatrix} 0 & y_2 \\ j(x) & ay_2' \end{vmatrix}}{aW} = -\dfrac{y_2 j(x)}{aW} \ , \ c_2'(x) = \dfrac{\begin{vmatrix} y_1 & 0 \\ ay_1' & j(x) \end{vmatrix}}{aW} = \dfrac{y_1 j(x)}{aW} \\ \\ \end{array} \right\}.$$

Innen

$$c_1(x) = \int -\frac{y_2 j(x)}{aW}\, dx \quad , \quad c_2(x) = \int \frac{y_1 j(x)}{aW}\, dx,$$

majd $c_1(x)$ és $c_2(x)$ ismeretében megkapjuk a kívánt $y_p = c_1(x)y_1 + c_2(x)y_2$ partikuláris megoldást. Megjegyezzük, hogy az integrálok kiszámolásánál nem kell a szokásos integrálási konstansokkal foglalkozni, hiszen egy-egy $c_1(x)$-re ill. $c_2(x)$-re van szükség. Végül az inhomogén egyenlet általános megoldása $y_I = y_H + y_p$.

4.3.4. Példa. Tekintsük a $3y'' - 2y' + y = \sqrt[3]{e^x}$ differenciálegyenletet. Határozzuk meg azon partikuláris megoldását, melyre $y(0) = 0$ és $y'(0) = 1$ teljesül.

Először a $3y'' - 2y' + y = 0$ homogén egyenletnek kell megadni az általános megoldását. Ehhez a $3\lambda^2 - 2\lambda + 1 = 0$ karakterisztikus egyenlet gyökeire van szükség. A megoldóképletből $\lambda_{1,2} = 1/3 \pm i\sqrt{2}/3$ következik, majd a gyököket felhasználva

$$y_H = c_1 e^{\frac{1}{3}x} \cos \frac{\sqrt{2}}{3}x + c_2 e^{\frac{1}{3}x} \sin \frac{\sqrt{2}}{3}x \ , \quad (c_1, c_2 \in \mathbb{R}).$$

Az inhomogén egyenlet egy partikuláris megoldását

$$y_H = c_1(x) e^{\frac{1}{3}x} \cos \frac{\sqrt{2}}{3}x + c_2(x) e^{\frac{1}{3}x} \sin \frac{\sqrt{2}}{3}x$$

alakban keressük. $j(x) = \sqrt[3]{e^x} = e^{x/3}$, a Wronski-féle determináns

$$W = \frac{\sqrt{2}}{3} e^{\frac{2}{3}x}.$$

A Cramer-szabály szerint

$$c_1'(x) = -\frac{\frac{1}{3} e^{\frac{1}{3}x} \sin \frac{\sqrt{2}}{3}x \cdot j(x)}{\frac{\sqrt{2}}{3} e^{\frac{2}{3}x}} = -\frac{1}{\sqrt{2}} \sin \frac{\sqrt{2}}{3}x \ ,$$

$$c_2'(x) = \frac{\frac{1}{3} e^{\frac{1}{3}x} \cos \frac{\sqrt{2}}{3}x \cdot j(x)}{\frac{\sqrt{2}}{3} e^{\frac{2}{3}x}} = \frac{1}{\sqrt{2}} \cos \frac{\sqrt{2}}{3}x \ .$$

Az integrálásoknál nem lép fel komplikáció,

$$c_1(x) = \frac{3}{2} \cos \frac{\sqrt{2}}{3}x \quad , \quad c_2(x) = \frac{3}{2} \sin \frac{\sqrt{2}}{3}x,$$

amiből

$$y_p = \frac{3}{2} \cos \frac{\sqrt{2}}{3}x \cdot e^{\frac{1}{3}x} \cos \frac{\sqrt{2}}{3}x + \frac{3}{2} \sin \frac{\sqrt{2}}{3}x \cdot e^{\frac{1}{3}x} \sin \frac{\sqrt{2}}{3}x = \frac{3}{2} e^{\frac{1}{3}x}.$$

Végül

$$y_I = c_1 e^{\frac{1}{3}x} \cos \frac{\sqrt{2}}{3}x + c_2 e^{\frac{1}{3}x} \sin \frac{\sqrt{2}}{3}x + \frac{3}{2} e^{\frac{1}{3}x} \ , \quad (c_1, c_2 \in \mathbb{R}).$$

Miután meghatároztuk az inhomogén egyenlet általános megoldását, a feladatban szereplő partikuláris megoldást célozzuk meg. Az $y(0) = 0$ kezdeti feltételből y_I-be való behelyettesítéssel rögtön $c_1 = -3/2$ adódik. Mivel $y'(0) = c_1/3 + \sqrt{2}c_2/3 + 1/2$, így c_1 ismeretében $c_2 = 3/\sqrt{2}$, tehát a keresett partikuláris megoldás

$$y(x) = -\frac{3}{2}e^{\frac{1}{3}x}\cos\frac{\sqrt{2}}{3}x + \frac{3}{\sqrt{2}}e^{\frac{1}{3}x}\sin\frac{\sqrt{2}}{3}x + \frac{3}{2}e^{\frac{1}{3}x}.$$

□

Az állandó együtthatós lineáris inhomogén differenciálegyenletek egy partikuláris megoldásának meghatározására gyakran előnyös alkalmazni az ún. próbafüggvény módszerét, amivel (a Lagrange módszerhez viszonyítva) kevesebb számolással és integrálás nélkül kaphatjuk meg a kívánt partikuláris megoldást.

Tekintsük (4.3.4) azon speciális eseteit, mikor $j(x)$ polinom- vagy exponenciális vagy trigonometrikus függvény, esetleg ezekből képzett összeg vagy szorzat. A próbafüggvény módszer lényege, hogy (4.3.4) $y_p(x)$ partikuláris megoldását a $j(x)$-szel megegyező típusú függvények körében keresi. Az adott típuson belül $y_p(x)$ együtthatóit az inhomogén egyenletbe való behelyettesítés után határozzuk meg. Az alábbi példák $y_p(x)$ elképzelt partikulási megoldás tipizálását mutatják.

A görög betűkkel ismert konstansokat, míg latin nagy betűkkel a meghatározandó együtthatókat jelöljük. Ha $j(x) = \alpha x^2 + \beta x + \gamma$ alakú, akkor $y_p(x) = Ax^2 + Bx + C$ lesz a próbafüggvény. $j(x) = \sin(\alpha x + \beta)$ vagy $\cos(\alpha x + \beta)$ esetén $y_p(x) = A\sin(\alpha x + \beta) + B\cos(\alpha x + \beta)$, míg $j(x) = (\alpha x + \beta)e^{\gamma x}$ esetén $(Ax + B)e^{Cx}$ alakban keresssük a partikuláris megoldást.

A próbafüggvény felállításánál probléma merül fel, ha a homogén egyenlet általános megoldásának egyik tagja (együttható nélkül) megegyezik $j(x)$-szel vagy annak egyik tagjával. Ekkor ún. *rezonancia* lép fel, és ilyenkor a korábbi elv alapján felírt próbafüggvényt, vagy annak egyik tagját még meg kell szorozni x-szel.

4.3.5. Példa. A próbafüggvény módszert is felhasználva megoldjuk az

$$y'' - y = e^x + 3e^{2x} \tag{4.3.7}$$

differenciálgyenletet. A származtatott homogén egyenlet általános megoldása $y_H(x) = c_1e^x + c_2e^{-x}$, ugyanis a $\lambda^2 - 1 = 0$ karakterisztikus egyenletből $\lambda_{1,2} = \pm 1$ következik.

Az inhomogén egyenlet egy partikuláris megoldására a próbafüggvény (4.3.7) jobb oldala alapján $y_p(x) = Ae^x + Be^{2x}$ lenne, de az $y_p(x)$-ben és $y_H(x)$-ben levő közös e^x-es tag miatt rezonancia van. Emiatt módosítani kell a próbafüggvényt $y_p(x) = Axe^x + Be^{2x}$-re. Az eredeti inhomogén differenciálegyenletbe $y_p(x)$-et, valamint $y_p''(x) = Axe^x + 2Ae^x + 4Be^{2x}$-et visszahelyettesítve

$$e^x + 3e^{2x} = y_p''(x) - y_p(x) = 2Ae^x + 3Be^{2x}$$

adódik. Ebből megkapjuk az $A = 1/2$, $B = 1$ együtthatókat.

Tehát $y_p(x) = xe^x/2 + e^{2x}$ továbbá $y_I(x) = c_1e^x + c_2e^{-x} + xe^x/2 + e^{2x}$. □

4.3.6. Példa. Most az

$$y'' - 4y' + 20y = 100x \tag{4.3.8}$$

differenciálegyenletet fogjuk megoldani, részben a próbafüggvény módszer alkalmazásával. A $\lambda^2 - 4\lambda + 20 = 0$ karakterisztikus egyenlet $\lambda_{1,2} = 2 \pm 4i$ megoldásaiból felírhatjuk a származtatott homogén egyenlet általános megoldását: $y_H(x) = e^{2x}(c_1\cos 4x + c_2\sin 4x)$.

Most nincs közös tag (4.3.8) jobb oldalában és $y_H(x)$-ben (azaz nincs rezonancia), emiatt a próbafüggvény $y_p(x) = Ax + B$. Az inhomogén egyenletbe visszahelyettesítve

$$100x = y'' - 4y' + 20y = 20Ax + (20B - 4A)$$

következik. A $20A = 100$ és $20B - 4A = 0$ egyenletekből $A = 5$, $B = 1$ adódik. Tehát $y_p(x) = 5x + 1$ és $y_I(x) = e^{2x}(c_1 \cos 4x + c_2 \sin 4x) + 5x + 1$. □

4.3.3. Feladat. Mi lesz a próbafüggvény, ha rezonanciát nem tételezünk fel?
a) $j(x) = x^3 + 1$;
b) $j(x) = 3x \cos 2x$;
c) $j(x) = 2x^2 + e^{-3x}$.

4.3.4. Feladat. Mi lesz a próbafüggvény, ha feltételezzük a rezonanciát?
a) $j(x) = -e^{-x}$;
b) $j(x) = \sin 5x$

4.3.5. Feladat. Oldjuk meg az $y'' + y' - 6y = 5e^{2x}$ differenciálegyenletet.

Ellenőrző kérdések

1. Melyik három alapeset fordulhat elő a másodrendű lineáris állandó együtthatós homogén differenciálegyenlet megoldásánál?

2. Mit értünk karakterisztikus egyenlet alatt?

3. Másodrendű lineáris differenciálegyenleteknél milyen a homogén ill. inhomogén egyenlet általános megoldásának szerkezete?

4. Mit értünk Wronski-féle determináns alatt?

5. Milyen esetekben alkalmazható a próbafüggvény módszere?

6. Mit értünk rezonancián és hogyan módosul ekkor a próbafüggvény?

A feladatok megoldásai

4.3.1.
Az $y(x) = e^{\lambda x}$ függvényt a differenciálegyenletbe helyettesítve ($y'(x) = \lambda e^{\lambda x}$, $y''(x) = \lambda^2 e^{\lambda x}$) az

$$e^{\lambda x}(a\lambda^2 + b\lambda + c) = 0$$

egyenlőséghez jutunk. Mivel $\lambda \in \mathbb{R}$ gyöke a karakterisztikus egyenletnek, ezért a bal oldal második tényezője zérus.

Megjegyezzük, hogy komplex λ esetén is igaz a feladat állítása, de a korábbi tanulmányok alapján (valós (és nem komplex) függvénytan) csak erre vállalkozhattunk.

4.3.2.
Két különböző valós gyök esetén

$$W = \begin{vmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{vmatrix} = \lambda_2 e^{\lambda_1 x} e^{\lambda_2 x} - \lambda_1 e^{\lambda_1 x} e^{\lambda_2 x} = (\lambda_2 - \lambda_1) e^{(\lambda_1 + \lambda_2)x},$$

és a különböző gyökök miatt $\lambda_1 \neq \lambda_2$, tehát $W \neq 0$.

Egybeeső (vagy másképp fogalmazva egyetlen) valós gyök(ök) esetén

$$W = \begin{vmatrix} e^{\lambda x} & xe^{\lambda x} \\ \lambda e^{\lambda x} & e^{\lambda x} + \lambda x e^{\lambda x} \end{vmatrix} = e^{2\lambda x} + \lambda x e^{2\lambda x} - \lambda x e^{2\lambda x} = e^{2\lambda x}.$$

Az exponenciális függvény pozitivitása miatt nyilvánvalóan most is $W \neq 0$ teljesül.

Komplex konjugált gyökpár mellett

$$W = \begin{vmatrix} e^{\lambda x}\cos\mu x & e^{\lambda x}\sin\mu x \\ \lambda e^{\lambda x}\cos\mu x - \mu e^{\lambda x}\sin\mu x & \lambda e^{\lambda x}\sin\mu x + \mu e^{\lambda x}\cos\mu x \end{vmatrix} =$$

$= \lambda e^{2\lambda x}\sin\mu x\cos\mu x + \mu e^{2\lambda x}\cos^2\mu x - \lambda e^{2\lambda x}\sin\mu x\cos\mu x + \mu e^{2\lambda x}\sin^2\mu x = \mu e^{2\lambda x}$. Itt $\mu \neq 0$ biztosítja azt, hogy a Wronski-féle determináns nem zérus.

4.3.3.

a) $y_p(x) = Ax^3 + Bx^2 + Cx + D$;
b) $y_p(x) = (Ax + B)(C\cos 2x + D\sin 2x)$;
c) $y_p(x) = (Ax^2 + Bx + C) + De^{-3x}$.

4.3.4.

a) $y_p(x) = Axe^{-x}$;
b) $y_p(x) = Ax\sin 5x + B\cos 5x$.

4.3.5.

A megadott inhomogén egyenlet általános megoldása

$$y_I(x) = c_1e^{2x} + c_2e^{-3x} + e^{2x}\left(x - \frac{1}{5}\right).$$

5. fejezet

Mátrixok

5.1. Nevezetes mátrixok, mátrixműveletek

Ebben a fejezetben a mátrixok fogalmával és a velük elvégezhető műveletekkel ismerkedünk meg. A továbbiak szempontjából nagyon fontosak lesznek a mátrixok összeadására és valós számmal való szorzására vonatkozó megfigyeléseink, mert azok azt mutatják, hogy a mátrixok ún. vektorteret alkotnak.

5.1.1. Definíció. Legyenek az a_{ij} $(i = 1, \dots, k;\ j = 1, \dots, n)$ számok tetszőleges valós számok. Ezen számok – a továbbiakban elemek – k sorba és n oszlopba való

$$\begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kj} & \dots & a_{kn} \end{bmatrix} \tag{5.1.1}$$

elrendezését $k \times n$-es vagy $k \times n$ típusú **mátrixnak** nevezzük (ejtsd: k-szor n-es mátrix). $\diamondsuit$

Érdekességképpen megjegyezzük, hogy a mátrix latin eredetű szó és anyaméhet jelent. Az összes $k \times n$ típusú mátrix halmazát $\mathbb{R}^{k \times n}$ fogja jelölni. Az első index (most k) a mátrix sorainak számát, míg a második (most n) az oszlopainak számát adja meg. Mátrixok jelölésére a vastag latin nagybetűket használjuk. Mivel az (5.1.1) mátrix elemei az a_{ij} számok, ezért célszerű a mátrixot $\mathbf{A}$-val jelölni. Amennyiben ki akarjuk hangsúlyozni, hogy hány sora ill. oszlopa van a mátrixnak, akkor használhatjuk az $\mathbf{A}_{k \times n}$ jelölést is. Ezeken kívül még hasznos az $\mathbf{A} = [a_{ij}]_{k \times n}$ jelölés, amely nemcsak a sorok és oszlopok számát, de azt is megmutatja, hogy mivel jelöltük az elemeket. Az a_{ij} elemet az $\mathbf{A} = [a_{ij}]_{k \times n}$ mátrix i-edik sorának j-edik elemének vagy j-edik oszlopának i-edik elemének nevezzük. Ha az $\mathbf{A}$ mátrix elemeire nem akarunk új jelölést bevezetni, akkor az $\mathbf{A}$ mátrix i-edik sorának j-edik elemét $(\mathbf{A})_{ij}$-vel is jelölhetjük.

5.1.2. Definíció. Legyenek $\mathbf{A}$ és $\mathbf{B}$ $k \times n$-es mátrixok. Ekkor $\mathbf{A} = \mathbf{B}$ (ill. $\mathbf{A} < \mathbf{B}$, $\mathbf{A} \leq \mathbf{B}$, $\mathbf{A} > \mathbf{B}$, $\mathbf{A} \geq \mathbf{B}$), ha $a_{ij} = b_{ij}$ ($a_{ij} < b_{ij}$, $a_{ij} \leq b_{ij}$, $a_{ij} > b_{ij}$, $a_{ij} \geq b_{ij}$) minden $i = 1, \dots, k$ és $j = 1, \dots, n$ esetén. $\diamondsuit$

5.1.1. Feladat. Tekintsük a

$$\begin{bmatrix} 2 & -1 & 0.5 \\ -4 & 0 & 0.3 \end{bmatrix}$$

mátrixot! Hány sora ill. oszlopa van? Mi a 2. sorának 3. eleme? Mekkora a 2. sorban szereplő elemek összege?

5.1.3. Definíció. Egy mátrixot **kvadratikus (négyzetes) mátrixnak** hívunk, ha sorainak száma megegyezik oszlopainak számával. ♢

5.1.4. Definíció. Az $\mathbb{R}^{n\times 1}$-beli mátrixokat (n elemű) **(oszlop)vektoroknak** nevezzük. Az $\mathbb{R}^{n\times 1}$ halmazt egyszerűbben $\mathbb{R}^n$-nel fogjuk jelölni. ♢

Azok a mátrixok, melyeknek csak egy oszlopuk van kitüntetett szerepet fognak játszani a mátrixok között, emiatt kaptak külön elnevezést. A vektorokat föléhúzott vastag kisbetűkkel jelöljük, pl. $\overline{\mathbf{a}}, \overline{\mathbf{b}}$, stb.

A mátrixok fontos szerepet töltenek be a matematikai és gazdasági számításokban. Tömör alakban írhatók fel velük a lineáris egyenletrendszerek ill. más optimalizálási feladatok. Városok közti távolságokat, bizonyos munkafolyamatokhoz szükséges időtartamokat ill. költségeket is célszerű mátrixokkal szemléltetni. Most műveleteket definiálunk köztük. Először az összeadással és a valós számmal való szorzással foglalkozunk.

5.1.5. Definíció. Legyen $\mathbf{A} = [a_{ij}]_{k\times n}$ és $\mathbf{B} = [b_{ij}]_{k\times n}$ két azonos sor- és oszlopszámú mátrix. Az $\mathbf{A}$ és $\mathbf{B}$ **mátrixok összegén** azt az $\mathbf{A} + \mathbf{B} = [c_{ij}]_{k\times n}$ mátrixot értjük, amely elemeire $c_{ij} = a_{ij} + b_{ij}$ $(i = 1, \ldots, k;\ j = 1, \ldots, n)$. ♢

5.1.6. Definíció. Legyen $\mathbf{A} = [a_{ij}]_{k\times n}$ egy tetszőleges mátrix és α egy tetszőleges valós szám. Ekkor az $\mathbf{A}$ **mátrix és az** α **valós szám szorzatán** azt az $\alpha\mathbf{A} = [c_{ij}]_{k\times n}$ mátrixot értjük, melynek elemeire $c_{ij} = \alpha \cdot a_{ij}$ $(i = 1, \ldots, k;\ j = 1, \ldots, n)$. ♢

5.1.2. Feladat. Tekintsük az

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 0.5 \\ -4 & 0 & 0.3 \\ 2 & 2 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & 1 & 0.5 \\ -2 & 0 & 0.3 \\ 1 & 0 & 7 \end{bmatrix}$$

mátrixokat! Határozzuk meg az $\mathbf{A} + \mathbf{B}$ és az $5\mathbf{A}$ mátrixokat!

5.1.3. Feladat. Tekintsük az

$$\overline{\mathbf{a}} = \begin{bmatrix} -1 \\ -7 \\ 2 \end{bmatrix}, \quad \overline{\mathbf{b}} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

vektorokat! Határozzuk meg az $\overline{\mathbf{a}} + \overline{\mathbf{b}}$ és $-2\overline{\mathbf{b}}$ vektorokat!

Vizsgáljuk meg a mátrixösszeadás és a valós számmal való szorzás tulajdonságait! A valós számok tulajdonságai alapján automatikusan adódnak a következő tulajdonságok.

1. A mátrixösszeadás kommutatív, azaz tetszőleges $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{k\times n}$ mátrixok esetén $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$.

2. A mátrixösszeadás asszociatív, azaz tetszőleges $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{k\times n}$ mátrixok esetén $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$.

3. Tekintsük azt az $\mathbf{O}$-val jelölt $\mathbb{R}^{k\times n}$-beli mátrixot, melynek minden eleme nulla. Ezt a mátrixot nullmátrixnak fogjuk hívni. Nyilvánvalóan $\mathbf{A}+\mathbf{O} = \mathbf{A}$ minden $\mathbf{A} \in \mathbb{R}^{k\times n}$ esetén.

4. Minden $\mathbf{A} = [a_{ij}]_{k\times n}$ mátrixnak létezik egy ún. additív inverze (inverze az összeadásra nézve), mellyel összeadva a nullmátrixot kapjuk. Ez a mátrix a $(-1)\mathbf{A} = [-a_{ij}]_{k\times n}$ mátrix.

5. Tetszőleges $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{k\times n}$ és $\alpha \in \mathbb{R}$ esetén $\alpha(\mathbf{A}+\mathbf{B}) = \alpha\mathbf{A}+\alpha\mathbf{B}$.

6. Tetszőleges $\mathbf{A} \in \mathbb{R}^{k\times n}$ és $\alpha, \beta \in \mathbb{R}$ esetén $(\alpha+\beta)\mathbf{A} = \alpha\mathbf{A}+\beta\mathbf{A}$.

7. Tetszőleges $\mathbf{A} \in \mathbb{R}^{k\times n}$ és $\alpha, \beta \in \mathbb{R}$ esetén $(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A})$.

8. Tetszőleges $\mathbf{A} \in \mathbb{R}^{k\times n}$ esetén $1\mathbf{A} = \mathbf{A}$.

Természetesen nem véletlen, hogy pontosan a fenti tulajdonságokat soroltuk fel. Az összeadás és a valós számmal való szorzás más matematikai objektumok (pl. függvények, polinomok, helyvektorok, stb.) között is értelmezhető, és ugyanolyan tulajdonságokkal rendelkeznek mint a mátrixok összeadása és számmal való szorzása. Ezen műveleti tulajdonságokra fog alapulni a 6.1. fejezetben tárgyalt vektortér fogalma.

Most a mátrixok közti szorzás műveletét definiáljuk. Szorzást csak olyan mátrixok között értelmezünk, melyeknél az első mátrix oszlopainak száma megegyezik a második sorainak számával.

5.1.7. Definíció. Legyen $\mathbf{A} = [a_{ij}]_{k\times m}$ és $\mathbf{B} = [b_{ij}]_{m\times n}$ két mátrix. Az **$\mathbf{A}$ és $\mathbf{B}$ mátrixok szorzatán** azt az $\mathbf{AB} = [c_{ij}]_{k\times n}$ mátrixot értjük, melyre

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{im}b_{mj} = \sum_{l=1}^{m} a_{il}b_{lj}$$

minden $i = 1, \ldots, k$ és $j = 1, \ldots, n$ esetén. $\diamondsuit$

Ezt a szorzási módot sor-oszlop szorzásnak hívjuk, mert az első mátrix sorainak elemeit rendre a második megfelelő oszlopainak elemeivel szorozzuk össze.

5.1.4. Feladat. Számítsuk ki az $\mathbf{AB}$ és $\mathbf{BA}$ szorzatokat az

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 2 \\ 3 & 1 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 3 & 1 \\ -3 & 1 \end{bmatrix}$$

mátrixokra.

A mátrixszorzás definíciója látszólag bonyolultnak tűnhet, de a továbbiakban látni fogjuk mennyire jól alkalmazható. Kérdezhetnénk, hogy miért nem úgy szorzunk össze két azonos méretű mátrixot, hogy elemeiket rendre összeszorozzuk egymással (hasonlóan az összeadáshoz). Erre most azt tudjuk válaszolni, hogy akkor a mátrixok teljesen hasonló struktúrát alkotnának a valós számokkal, és így értelmezésüknek nem sok haszna lenne. Most egy példával világítunk rá a mátrixszorzás definíciójának "életrevalóságára".

Tekintsük a

$$\left.\begin{array}{rcrcl} 3x & + & 2y & = & 1 \\ x & - & y & = & 2 \end{array}\right\}$$

egyenletrendszert. Az ismeretlenek együtthatóiból alkotott mátrixot jelöljük $\mathbf{A}$-val, a jobboldalon lévő elemekből álló oszlopvektort jelöljük $\overline{\mathbf{b}}$-vel, valamint alkossunk egy $\overline{\mathbf{x}}$ oszlopvektort a két ismeretlennel, azaz

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & -1 \end{bmatrix}, \quad \overline{\mathbf{b}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \overline{\mathbf{x}} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Így a lineáris egyenletrendszer a most definiált mátrixszorzással $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$ alakban írható, ami párhuzamba állítható egy egyismeretlenes lineáris egyenlet alakjával: $ax = b$, ahol $a, b \in \mathbb{R}$ adott számok és az x valós számot kell meghatározni.

Vizsgáljuk meg a mátrixszorzás tulajdonságait. Ehhez egy olyan halmazt kell definiálni, melynek elemei között értelmezett a szorzás és az nem vezet ki a halmazból. Könnyű észrevenni, hogy a kvadratikus mátrixok bármilyen sorrendben összeszorozhatók és a szorzat is ugyanolyan típusú mátrix lesz.

Tekintsük tehát az $\mathbb{R}^{n\times n}$ halmazt, azaz az $n \times n$-es kvadratikus mátrixok halmazát és a rajta értelmezett szorzás műveletet. Egy művelet esetén először annak kommutativitását és asszociativitását szoktuk ellenőrizni. Így teszünk most is.

A szorzás műveletéről azt fogjuk megmutatni, hogy nem kommutatív. Ehhez elegendő egy ellenpéldát találni. Tekintsük az

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$$

mátrixokat. Ezeket kétféle sorrendben összeszorozva kapjuk, hogy

$$\mathbf{AB} = \begin{bmatrix} 2 & 8 \\ -1 & 1 \end{bmatrix}, \quad \mathbf{BA} = \begin{bmatrix} 2 & -2 \\ 4 & 1 \end{bmatrix},$$

ami mutatja, hogy a szorzás valóban nem kommutatív. Ha valamely $\mathbf{A}$ és $\mathbf{B}$ mátrixokra teljesül az $\mathbf{AB} = \mathbf{BA}$ egyenlőség, akkor azt mondjuk, hogy a mátrixok kommutálnak .

Most megmutatjuk, hogy a mátrixszorzás asszociatív. Ehhez azt kell igazolnunk, hogy tetszőleges $\mathbf{A} = [a_{ij}]_{n\times n}, \mathbf{B} = [b_{ij}]_{n\times n}$ és $\mathbf{C} = [c_{ij}]_{n\times n}$ mátrixokra teljesül, hogy $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$. A kívánt összefüggés leolvasható az

$$((\mathbf{AB})\mathbf{C})_{ij} = \sum_{k=1}^{n} (\mathbf{AB})_{ik} \cdot c_{kj} = \sum_{k=1}^{n} (\sum_{l=1}^{n} a_{il} \cdot b_{lk}) \cdot c_{kj} = \sum_{k=1}^{n} (\sum_{l=1}^{n} a_{il} \cdot b_{lk} \cdot c_{kj})$$

és

$$(\mathbf{A}(\mathbf{BC}))_{ij} = \sum_{l=1}^{n} a_{il} \cdot (\mathbf{BC})_{lj} = \sum_{l=1}^{n} a_{il} \cdot (\sum_{k=1}^{n} b_{lk} \cdot c_{kj}) = \sum_{l=1}^{n} (\sum_{k=1}^{n} a_{il} \cdot b_{lk} \cdot c_{kj})$$

egyenlőségből, hiszen az l és k szerinti összegzés felcserélhető egymással. Megjegyezzük, hogy a mátrixszorzás asszociativitása nem csak kvadratikus mátrixokra igaz, hanem tetszőleges olyan mátrixokra is, melyekre elvégezhető a szorzás.

Mielőtt tovább vizsgálnánk a szorzás tulajdonságait, célszerű bevezetni két új fogalmat.

5.1.8. Definíció. Egy **mátrix főátlóján** azon elemek összességét értjük, melyek sorill. oszlopindexe megegyezik. ♢

5.1.9. Definíció. Egy mátrixot **diagonális mátrixnak** hívunk, ha a főátlóján kívüli elemei nullával egyenlők. ♢

Egy $\mathbf{A} = [a_{ij}]_{k\times n}$ mátrix esetén a főátlót tehát az a_{ii} elemek alkotják, ahol $i = 1, \ldots, \min\{k, n\}$. Azt, hogy az $\mathbf{A} = [a_{ij}]_{k\times n}$ mátrix nem diagonálmátrix, úgy is megfogalmazhatjuk, hogy van olyan i és j index ($i \neq j$), melyekre $a_{ij} \neq 0$.

5.1.5. Feladat. Mutassunk példát diagonális és nem diagonális mátrixokra.

A következőkben azt vizsgáljuk, hogy van-e olyan mátrix, amellyel az $n\times n$-es négyzetes mátrixokat szorozva azok nem változnak meg. Könnyű rájönni, hogy az egyetlen ilyen mátrix az a diagonális mátrix lesz, melynek a főátlójában egyesek állnak (a többi eleme nulla). Ezt a mátrixot egységmátrixnak hívjuk, és $\mathbf{E}$-vel jelöljük. Ha ki akarjuk hangsúlyozni az egységmátrix méretét is, akkor használhatjuk az $\mathbf{E}_n$ jelölést. Vegyük észre, hogy akár balról akár jobbról szorzunk egy mátrixot az egységmátrixszal az változatlan marad.

A negyedik vizsgált tulajdonsága a szorzásnak az, hogy létezik-e a négyzetes mátrixoknak multiplikatív (szorzásra vonatkozó) inverzük.

5.1.10. Definíció. Ha az $\mathbf{A} \in \mathbb{R}^{n\times n}$ mátrixhoz létezik egy olyan $\mathbf{A}^{-1}$-gyel jelölt mátrix, melyre $\mathbf{A}\mathbf{A}^{-1} = \mathbf{E}$, akkor az $\mathbf{A}$ mátrixot **invertálható (vagy reguláris) mátrixnak** hívjuk. Az $\mathbf{A}^{-1}$ mátrixot az **$\mathbf{A}$ mátrix inverzének** nevezzük. ♢

Igazolható, hogy ha egy mátrixnak van inverze, akkor az $\mathbf{A}^{-1}\mathbf{A} = \mathbf{E}$ egyenlőség is teljesül. (Ez amiatt nem nyilvánvaló, mert a szorzás nem kommutatív.) Rögtön láthatjuk, hogy nincs minden mátrixnak inverze. Ha $\mathbf{A}$ olyan mátrix, melynek mondjuk az i-edik sora csupa nulla elemekből áll, akkor azt bármilyen mátrixszal is szorozzuk, a szorzat i-edik sora is nulla elemekből fog állni a sor-oszlop szorzás miatt, tehát az nem lehet az egységmátrix. Azokat a mátrixokat, melyeknek nincs inverzük szinguláris mátrixoknak hívjuk. Egy mátrix inverzének meghatározása nem egyszerű feladat. A 7.5. fejezetben megmutatjuk, hogy hogyan lehet mátrixok inverzét bázistranszformáció segítségével meghatározni.

Az összeadás és szorzás műveleteket a disztributivitási (széttagolhatósági) szabály köti össze. Ellenőrizzük le tehát, hogy vajon a mátrixösszeadás a mátrixszorzásra nézve disztributív-e. Tekintsük az $\mathbf{A} = [a_{ij}]_{n\times n}, \mathbf{B} = [b_{ij}]_{n\times n}$ és $\mathbf{C} = [c_{ij}]_{n\times n}$ mátrixokat, és vizsgáljuk az

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

egyenlőségeket. A disztributivitást mindkét irányban le kell ellenőrizni, mivel a mátrixszorzás nem kommutatív. Emiatt szerepel itt két disztributivitási szabály. Az első egyenlőség teljesülését mutatják az

$$(\mathbf{A}(\mathbf{B}+\mathbf{C}))_{ij} = \sum_{k=1}^{n} a_{ik}\cdot(\mathbf{B}+\mathbf{C})_{kj} = \sum_{k=1}^{n} a_{ik}\cdot(b_{kj}+c_{kj}) = \sum_{k=1}^{n}(a_{ik}\cdot b_{kj} + a_{ik}\cdot c_{kj}),$$

$$(\mathbf{A}\mathbf{B}+\mathbf{A}\mathbf{C})_{ij} = (\mathbf{A}\mathbf{B})_{ij} + (\mathbf{A}\mathbf{C})_{ij} = \sum_{k=1}^{n} a_{ik}\cdot b_{kj} + \sum_{k=1}^{n} a_{ik}\cdot c_{kj} = \sum_{k=1}^{n}(a_{ik}\cdot b_{kj} + a_{ik}\cdot c_{kj})$$

egyenlőségek. A másodikat hasonló módon lehet igazolni. Megemlítjük, hogy a disztributivitási szabály nem csak kvadratikus mátrixokra teljesül, hanem tetszőleges olyan mátrixokra is, melyekkel a műveletek elvégezhetőek.

Most tétel formájában megfogalmazzuk a mátrixszorzás tulajdonságaival kapcsolatos megfigyeléseinket.

5.1.11. Tétel. *Az $\mathbb{R}^{n\times n}$ négyzetes mátrixok halmazán értelmezett mátrixszorzás nem kommutatív; asszociatív; létezik $\mathbf{E}$ egységmátrix, melyre $\mathbf{EA} = \mathbf{AE} = \mathbf{A}$ teljesül minden $\mathbf{A}$ mátrixra; nincs minden $\mathbf{A}$ mátrixnak olyan $\mathbf{A}^{-1}$ inverze, melyre $\mathbf{A}^{-1}\mathbf{A} = \mathbf{AA}^{-1} = \mathbf{E}$, valamint az összeadás a szorzásra nézve disztributív.*

Az egyszerűség kedvéért mátrixokra is értelmezzük a pozitív egész kitevőjű hatványokat. Az l-tényezős $\mathbf{A}\cdot\ldots\cdot\mathbf{A}$ szorzatot ezentúl $\mathbf{A}^l$ fogja jelölni. Egy $\mathbf{A}$ mátrix nulladik hatványát az egységmátrixnak definiáljuk, azaz $\mathbf{A}^0 = \mathbf{E}$.

5.1.6. Feladat. Milyen mátrixokra teljesül az $(\mathbf{A}+\mathbf{B})(\mathbf{A}-\mathbf{B}) = \mathbf{A}^2 - \mathbf{B}^2$ egyenlőség?

5.1.12. Definíció. Egy $\mathbb{R}^{k\times n}$-beli **$\mathbf{A}$ mátrix transzponáltján** azt az $\mathbf{A}^\top \in \mathbb{R}^{n\times k}$ mátrixot értjük, melynek i-edik sorának j-edik eleme megegyezik az $\mathbf{A}$ mátrix j-edik sorának i-edik elemével. Azaz $(\mathbf{A}^\top)_{ij} = (\mathbf{A})_{ji}$ minden $j = 1,\ldots,k$ és $i = 1,\ldots,n$ index esetén. $\diamondsuit$

Transzponálás során a mátrix soraiból a transzponált mátrix oszlopai, oszlopaiból pedig a sorai lesznek. Az elemek úgy rendeződnek át, mintha a mátrixot tükröztük volna a főátlójára. Vegyük észre, hogy a transzponálás során a mátrix főátlójának elemei a helyükön maradnak, hiszen a transzponált definíciója szerint $(\mathbf{A}^\top)_{ii} = (\mathbf{A})_{ii}$. Most felsorolunk még néhány tulajdonságot, melyeket az Olvasó könnyen igazolhat a transzponálás definíciójának közvetlen alkalmazásával. Tetszőleges $\mathbf{A}$ mátrix esetén

$$(\mathbf{A}^\top)^\top = \mathbf{A},$$

tetszőleges $\mathbf{A},\mathbf{B} \in \mathbb{R}^{k\times n}$ mátrixok és α valós szám esetén

$$(\mathbf{A}+\mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top,$$

$$(\alpha\mathbf{A})^\top = \alpha\mathbf{A}^\top$$

és tetszőleges $\mathbf{A} \in \mathbb{R}^{k\times m}$ és $\mathbf{B} \in \mathbb{R}^{m\times n}$ mátrixok esetén

$$(\mathbf{AB})^\top = \mathbf{B}^\top\mathbf{A}^\top.$$

5.1.7. Feladat. Határozzuk meg az

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

mátrix főátlóját, majd transzponáljuk a mátrixot. Mi lesz az $\mathbf{A}^\top$ mátrix főátlója?

5.1.13. Definíció. Egy $\mathbf{A}$ kvadratikus **mátrix szimmetrikus**, ha transzponáltja saját maga, azaz ha teljesül az $\mathbf{A}^\top = \mathbf{A}$ egyenlőség, ill. **antiszimmetrikus** vagy **ferdén szimmetrikus**, ha $\mathbf{A}^\top = -\mathbf{A}$. $\diamondsuit$

Mivel transzponálás során a sorok és oszlopok helyet cserélnek egymással, egy mátrix csak akkor lehet szimmetrikus vagy ferdén szimmetrikus, ha sorainak és oszlopainak száma megegyezik. Emiatt már csak ilyen mátrixokra definiáltuk a fogalmakat.

5.1.8. Feladat. Igazoljuk, hogy egy ferdén szimmetrikus mátrix főátlójában csak nullák szerepelhetnek.

5.1.9. Feladat. Milyen számokat írjunk az egyes mátrixok szabadon hagyott helyeire ahhoz, hogy szimmetrikus ill. ferdén szimmetrikus mátrixokat nyerjünk?

$$\mathbf{A} = \begin{bmatrix} 2 & . \\ -1 & . \\ 1 & -2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & 2 & . \\ . & 5 & . \\ 1 & -1 & . \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & . & -1 \\ 3 & . & 1 \\ . & . & . \end{bmatrix}$$

5.1.14. Tétel. *Minden $\mathbf{A} \in \mathbb{R}^{n\times n}$ kvadratikus mátrix felírható egy szimmetrikus és egy ferdén szimmetrikus mátrix összegeként.*

Bizonyítás. Legyen $\mathbf{S}$ az $(1/2)(\mathbf{A} + \mathbf{A}^\top)$ mátrix. Ez a mátrix szimmetrikus, ami az

$$\mathbf{S}^\top = [(1/2)(\mathbf{A} + \mathbf{A}^\top)]^\top = (1/2)(\mathbf{A} + \mathbf{A}^\top)^\top = (1/2)(\mathbf{A}^\top + \mathbf{A}) = (1/2)(\mathbf{A} + \mathbf{A}^\top) = \mathbf{S}$$

egyenlőségből következik. Legyen $\mathbf{F}$ az $(1/2)(\mathbf{A} - \mathbf{A}^\top)$ mátrix. Ez a mátrix ferdén szimmetrikus, mert

$$\mathbf{F}^\top = [(1/2)(\mathbf{A} - \mathbf{A}^\top)]^\top = (1/2)(\mathbf{A} - \mathbf{A}^\top)^\top = (1/2)(\mathbf{A}^\top - \mathbf{A}) = -(1/2)(\mathbf{A} - \mathbf{A}^\top) = -\mathbf{F}.$$

Végül megmutatjuk, hogy $\mathbf{A} = \mathbf{S} + \mathbf{F}$, ami mutatni fogja, hogy $\mathbf{A}$ előáll egy szimmetrikus ($\mathbf{S}$) és egy ferdén szimmetrikus ($\mathbf{F}$) mátrix összegeként.

$$\mathbf{S}+\mathbf{F} = (1/2)(\mathbf{A}+\mathbf{A}^\top)+(1/2)(\mathbf{A}-\mathbf{A}^\top) = (1/2)\mathbf{A}+(1/2)(\mathbf{A}^\top)+(1/2)\mathbf{A}-(1/2)(\mathbf{A}^\top) = \mathbf{A}. \blacksquare$$

5.1.10. Feladat. Írjuk fel az

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

mátrixot egy szimmetrikus és egy ferdén szimmetrikus mátrix összegeként.

A transzponálás műveletének segítségével helytakarékosan írhatók fel az $\mathbb{R}^n$ tér oszlopvektorai. Pl. az $\overline{\mathbf{a}} = [1, 2, 3, 4]^\top$ vektor egy $\mathbb{R}^4$-beli oszlopvektor, de a transzponálás segítségével – helyet megtakarítva – sorvektorként írható fel. Ezt a módszert e jegyzet írása közben is többször alkalmaztuk. Megjegyezzük, hogy bár nem szoktunk egy mátrixban az elemek közé vesszőt tenni, egy sorvektor esetén – a félreértések elkerülése miatt – mégis vesszővel választjuk el egymástól az elemeket.

Külön kiemelünk most néhány speciális mátrixszorzatot, amelyek a későbbiek során gyakran előfordulnak majd.

Mátrix és oszlopvektor szorzata. Ha $\mathbf{A} \in \mathbb{R}^{k\times n}$ egy tetszőleges mátrix és $\overline{\mathbf{x}} = [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$ egy tetszőleges vektor (n-szer 1-es mátrix), akkor ezek összeszorozhatók és szorzatuk egy k elemű vektor lesz, azaz $\mathbf{A}\overline{\mathbf{x}} \in \mathbb{R}^k$.

5.1.11. Feladat. Határozzuk meg az $\mathbf{A}\overline{\mathbf{x}}$ szorzatot, ha

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 3 \\ 3 & 1 & -1 \\ 4 & 0 & 3 \end{bmatrix}, \quad \overline{\mathbf{x}} = \begin{bmatrix} -1 \\ 3 \\ 3 \end{bmatrix} !$$

Oszlopvektor és sorvektor szorzata. Tekintsünk most két oszlopvektort, pl. az $\overline{\mathbf{x}} = [x_1, \ldots, x_n]^\top$ és $\overline{\mathbf{y}} = [y_1, \ldots, y_n]^\top$ vektorokat! Az $\overline{\mathbf{x}}\overline{\mathbf{y}}^\top$ szorzat, amely nyilván elvégezhető, eredménye egy $n \times n$-es mátrix lesz.

$$\overline{\mathbf{x}}\overline{\mathbf{y}}^\top = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1, y_2 \ldots, y_n \end{bmatrix} = \begin{bmatrix} x_1 \cdot y_1 & \ldots & x_1 \cdot y_n \\ x_2 \cdot y_1 & \ldots & x_2 \cdot y_n \\ \vdots & \ddots & \vdots \\ x_n \cdot y_1 & \ldots & x_n \cdot y_n \end{bmatrix}.$$

Sorvektor és oszlopvektor szorzata. A fenti két oszlopvektor az $\overline{\mathbf{x}}^\top\overline{\mathbf{y}}$ alakban is összeszorozható. Az eredmény egy 1×1-es mátrix, melynek egyetlen eleme $x_1\cdot y_1+\ldots+x_n\cdot y_n$, azaz

$$\overline{\mathbf{x}}^\top\overline{\mathbf{y}} = \begin{bmatrix} x_1, x_2, \ldots, x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 \cdot y_1 + x_2 \cdot y_2 + \ldots + x_n \cdot y_n \end{bmatrix}.$$

5.1.15. Definíció. Legyen $\overline{\mathbf{x}} = [x_1, \ldots, x_n]^\top$ és $\overline{\mathbf{y}} = [y_1, \ldots, y_n]^\top$ két $\mathbb{R}^n$-beli vektor. Ekkor az $x_1\cdot y_1+x_2\cdot y_2+\ldots+x_n\cdot y_n$ mennyiséget az $\overline{\mathbf{x}}$ és $\overline{\mathbf{y}}$ vektorok **skaláris szorzatának** nevezzük és az $\langle\overline{\mathbf{x}}, \overline{\mathbf{y}}\rangle$ szimbólummal jelöljük. $\diamondsuit$

A skaláris szorzás tehát két vektorhoz egy számot (skalárt) rendel hozzá. Ez a szám az $\overline{\mathbf{x}}^\top\overline{\mathbf{y}}$ vektor egyetlen eleme. A skaláris szorzattal az $\overline{\mathbf{x}}^\top\overline{\mathbf{y}}$ vektor az $\overline{\mathbf{x}}^\top\overline{\mathbf{y}} = [\langle\overline{\mathbf{x}}, \overline{\mathbf{y}}\rangle]$ alakban is írható. Megjegyezzük, hogy gyakran – praktikus okokból – azonosítják az $\overline{\mathbf{x}}^\top\overline{\mathbf{y}}$ szorzatot magával az $x_1 \cdot y_1 + \ldots + x_n \cdot y_n$ valós számmal is.

Az 5.1.15. definíció alapján könnyen igazolhatók a skaláris szorzás következő tulajdonságai.

1. Minden $\overline{\mathbf{x}}, \overline{\mathbf{y}} \in \mathbb{R}^n$ esetén $\langle\overline{\mathbf{x}}, \overline{\mathbf{y}}\rangle = \langle\overline{\mathbf{y}}, \overline{\mathbf{x}}\rangle$.
2. Minden $\overline{\mathbf{x}}, \overline{\mathbf{y}} \in \mathbb{R}^n$ és $\lambda \in \mathbb{R}$ esetén $\langle\lambda\overline{\mathbf{x}}, \overline{\mathbf{y}}\rangle = \lambda\langle\overline{\mathbf{x}}, \overline{\mathbf{y}}\rangle$.
3. Minden $\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{\mathbf{z}} \in \mathbb{R}^n$ esetén $\langle\overline{\mathbf{x}} + \overline{\mathbf{y}}, \overline{\mathbf{z}}\rangle = \langle\overline{\mathbf{x}}, \overline{\mathbf{z}}\rangle + \langle\overline{\mathbf{y}}, \overline{\mathbf{z}}\rangle$.
4. Minden $\overline{\mathbf{x}} \in \mathbb{R}^n$ esetén $\langle\overline{\mathbf{x}}, \overline{\mathbf{x}}\rangle \geq 0$, és $\langle\overline{\mathbf{x}}, \overline{\mathbf{x}}\rangle$ pontosan akkor nulla, ha $\overline{\mathbf{x}} \neq \overline{\mathbf{0}}$.

Az $\mathbb{R}^1, \mathbb{R}^2$ és $\mathbb{R}^3$ halmazok elemei és a számegyenes, a koordinátasík ill. a koordinátatér helyvektorai között sokszor célszerű megfeleltetést létesíteni a következők szerint. Az $[x] \in \mathbb{R}^1$ vektort feleltessük meg a számegyenes origójából az x-et reprezentáló pontba mutató helyvektorral, az $[x, y]^\top \in \mathbb{R}^2$ vektort a síkbeli derékszögű koordinátarendszer origóját

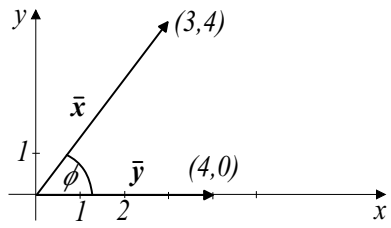
az (x, y) koordinátájú ponttal összekötő vektorral, és hasonlóan eljárva az $[x, y, z]^\top \in \mathbb{R}^3$ vektort feleltessük meg a térbeli derékszögű koordinátarendszer origóját az (x, y, z) ponttal összekötő vektorral.

A skaláris szorzás és a fenti megfeleltetés segítségével meghatározhatjuk helyvektorok hosszát és szögét. Tekintsük pl. az 5.1.1. ábrát, ahol $\overline{\mathbf{x}} = [3, 4]^\top$ és $\overline{\mathbf{y}} = [4, 0]^\top$. Számítsuk ki az $\langle \overline{\mathbf{x}}, \overline{\mathbf{x}} \rangle$ skaláris szorzat gyökét.

$$\sqrt{\langle \overline{\mathbf{x}}, \overline{\mathbf{x}} \rangle} = \sqrt{3 \cdot 3 + 4 \cdot 4} = 5.$$

Látható, hogy a fenti kifejezés a Pitagorasz-tétel miatt éppen az $\overline{\mathbf{x}}$ vektor hosszát adja meg. Ezen megfigyelésre alapozva vezetjük be a következő definíciót.

5.1.16. Definíció. Egy $\overline{\mathbf{x}} \in \mathbb{R}^n$ **vektor hosszán** vagy **normáján** az $\|\overline{\mathbf{x}}\| = \sqrt{\langle \overline{\mathbf{x}}, \overline{\mathbf{x}} \rangle}$ értéket értjük. $\diamondsuit$



5.1.1. ábra. *Síkvektorok szemléltetése a hossz és szög értelmezéséhez.*

A hossz definíciója tehát $\mathbb{R}^1, \mathbb{R}^2$ és $\mathbb{R}^3$ esetén valóban a helyvektorok hosszát adja meg. Vegyük észre, hogy egy vektor hossza csak abban az esetben nulla, ha a vektor a nullvektor. A hossz minden vektor esetén kiszámítható a skaláris szorzás 4. tulajdonsága miatt.

Számítsuk ki most az 5.1.1. ábrán látható $\overline{\mathbf{x}}$ és $\overline{\mathbf{y}}$ vektorok ϕ szögét. Nyilvánvalóan

$$\cos \phi = \frac{3}{5}$$

és így $\phi = 53.13°$. Tekintsük most a

$$\cos \phi = \frac{3}{5} = \frac{3}{\|\overline{\mathbf{x}}\|} \cdot \frac{4}{4} = \frac{3 \cdot 4}{\|\overline{\mathbf{x}}\| \cdot \|\overline{\mathbf{y}}\|} = \frac{\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle}{\|\overline{\mathbf{x}}\| \cdot \|\overline{\mathbf{x}}\|}$$

átalakítást, amely a skaláris szorzás segítségével adja meg két vektor szögének koszinuszát. Vezessük be a következő definíciót.

5.1.17. Definíció. Az $\overline{\mathbf{x}} \in \mathbb{R}^n$ és $\overline{\mathbf{y}} \in \mathbb{R}^n$ nullvektortól különböző **vektorok szögén** az

$$\phi = \arccos \frac{\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle}{\|\overline{\mathbf{x}}\| \cdot \|\overline{\mathbf{y}}\|}$$

értéket értjük. $\diamondsuit$

Ha az $\overline{\mathbf{x}}$ és $\overline{\mathbf{y}}$ vektorok valamelyike nullvektor, akkor a szögüket tetszőlegesnek szokták tekinteni. Joggal tehetjük fel a kérdést, hogy vajon a fenti definícióban mindig értelmezett-e az *arccos* függvény, azaz hogy az $\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle / (\|\overline{\mathbf{x}}\| \cdot \|\overline{\mathbf{y}}\|)$ érték mindig a $[-1, 1]$ intervallumba esik-e.

5.1.18. Tétel. *(Cauchy-Bunyakovszkij-Schwarz (CBS) egyenlőtlenség) Tetszőleges* $\overline{\mathbf{x}}, \overline{\mathbf{y}} \in \mathbb{R}^n$ *vektorra teljesül az*

$$|\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle| \leq \|\overline{\mathbf{x}}\| \cdot \|\overline{\mathbf{y}}\| \tag{5.1.2}$$

egyenlőtlenség.

Bizonyítás. Ha valamelyik vektor nullvektor, akkor nyilvánvalóan igaz az állítás. Tegyük most fel, hogy egyik vektor sem nullvektor és alkalmazzuk a skaláris szorzás tulajdonságait.

$$0 \leq \left\langle \overline{\mathbf{x}} - \frac{\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle}{\|\overline{\mathbf{y}}\|^2}\overline{\mathbf{y}}, \overline{\mathbf{x}} - \frac{\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle}{\|\overline{\mathbf{y}}\|^2}\overline{\mathbf{y}} \right\rangle = \langle \overline{\mathbf{x}}, \overline{\mathbf{x}} \rangle - 2\frac{\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle}{\|\overline{\mathbf{y}}\|^2}\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle + \frac{\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle^2}{\|\overline{\mathbf{y}}\|^4}\langle \overline{\mathbf{y}}, \overline{\mathbf{y}} \rangle =$$

$$= \langle \overline{\mathbf{x}}, \overline{\mathbf{x}} \rangle - \frac{\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle^2}{\|\overline{\mathbf{y}}\|^2} = \langle \overline{\mathbf{x}}, \overline{\mathbf{x}} \rangle - \frac{\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle^2}{\langle \overline{\mathbf{y}}, \overline{\mathbf{y}} \rangle}.$$

Az egyenlőtlenséget az

$$\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle^2 \leq \langle \overline{\mathbf{x}}, \overline{\mathbf{x}} \rangle \cdot \langle \overline{\mathbf{y}}, \overline{\mathbf{y}} \rangle$$

alakba átírva, majd gyököt vonva, a tétel állítását kapjuk. ■

A CBS-egyenlőtlenség miatt tehát

$$\frac{|\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle|}{\|\overline{\mathbf{x}}\| \cdot \|\overline{\mathbf{y}}\|} \leq 1,$$

ami mutatja, hogy két vektor szöge mindig értelmezett. A CBS-egyenlőtlenség a vektorok koordinátáival az

$$|x_1 \cdot y_1 + \ldots + x_n \cdot y_n| \leq \sqrt{x_1^2 + \ldots + x_n^2} \cdot \sqrt{y_1^2 + \ldots + y_n^2}$$

alakban is felírható.

5.1.19. Példa. Számítsuk ki az $\overline{\mathbf{x}} = [1, 2, 3]^\top$ és $\overline{\mathbf{y}} = [-2, 3, 1]^\top$ vektorok szögét!

$$\|\overline{\mathbf{x}}\| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14},$$

$$\|\overline{\mathbf{y}}\| = \sqrt{(-2)^2 + 3^2 + 1^2} = \sqrt{14}$$

és

$$\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle = -2 + 6 + 3 = 7.$$

Tehát

$$\phi = \arccos \frac{7}{\sqrt{14} \cdot \sqrt{14}} = \arccos \frac{1}{2} = 60^\circ.$$

□

5.1.12. Feladat. Számítsuk ki az $\overline{\mathbf{x}}^\top \overline{\mathbf{y}}$ és $\overline{\mathbf{x}}\overline{\mathbf{y}}^\top$ szorzatokat ill. az $\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle$ skaláris szorzatot, ha $\overline{\mathbf{x}} = [-1, 2, 3]^\top$ és $\overline{\mathbf{y}} = [2, 2, 3]^\top$.

Egy konkrét mátrix megadásánál ill. mátrixokkal való számolások során fontos szerepet játszanak az ún. blokkmátrixok.

5.1.20. Definíció. **Blokkmátrixnak** hívjuk azokat a mátrixokat, melyeknek elemei szintén mátrixok, és bármely sorának minden eleme azonos sorszámú ill. bármely oszlopának minden eleme azonos oszlopszámú. ♢

Az

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 3 \end{bmatrix}$$

mátrix megadásakor két lehetőségünk van: vagy leírjuk az egész mátrixot, vagy pedig megmondjuk, hogy hol vannak benne egyesek, hol hármasok és mely elemek nullák. A mátrix megadása azonban egyszerűsíthető, ha észrevesszük, hogy a mátrix első három oszlopa éppen egy 3×3-as egységmátrix, az utolsó három oszlopa pedig annak háromszorosa. Az $\mathbf{A}$ mátrix tehát formálisan az $\mathbf{A} = [\mathbf{E}_3\ 3\mathbf{E}_3]$ alakban írható. Így $\mathbf{A}$ tekinthető egy 1×2-es blokkmátrixnak, melynek elemei az $\mathbf{E}_3$ és $3\mathbf{E}_3$ mátrixok. A blokkok elkülönítésére használhatjuk (és sok esetben ajánlatos is használni) még az $\mathbf{A} = [\mathbf{E}_3, 3\mathbf{E}_3]$ és $\mathbf{A} = [\mathbf{E}_3|3\mathbf{E}_3]$ jelöléseket is.

5.1.21. Példa. Ha egy $\mathbf{A} \in \mathbb{R}^{k\times n}$ mátrix j-edik oszlopának elemeivel definiálunk egy $\overline{\mathbf{a}}_j \in \mathbb{R}^k$ oszlopvektort ($j = 1, \ldots, n$), akkor $\mathbf{A}$ egy $1 \times n$-es blokkmátrixként az $[\overline{\mathbf{a}}_1, \ldots, \overline{\mathbf{a}}_n]$ alakban írható fel. □

5.1.22. Példa. Ha egy $\mathbf{A} \in \mathbb{R}^{k\times n}$ mátrixhoz hozzávesszük $(n+1)$-edik oszlopként a $\overline{\mathbf{b}} \in \mathbb{R}^k$ vektort, akkor egy $[\mathbf{A}|\overline{\mathbf{b}}]$ blokkmátrixot kapunk. Ezt a blokkmátrixot szokás az $\mathbf{A}$ mátrix $\overline{\mathbf{b}}$ vektorral való kibővítésének is hívni. □

Blokkmátrixok között is lehet értelmezni a mátrixműveleteket. Ha két blokkmátrix ugyanolyan típusú – azaz ugyanannyi soruk és oszlopuk van – és a két mátrix egymásnak megfelelő elemei is rendre ugyanolyan típusú mátrixok, akkor úgy adhatjuk össze őket, hogy az egymásnak megfelelő elemeket összeadjuk. Egy blokkmátrixot úgy szorzunk meg egy valós számmal, hogy minden elemét megszorozzuk az adott valós számmal. Két blokkmátrix akkor összeszorozható, ha az első mátrix első sorában álló mátrixok oszlopainak száma rendre megegyezik a második mátrix első oszlopában álló mátrixok sorainak számával. A szorzás ezek után a mátrixoknál tanult sor-oszlop szorzás szerint történik.

5.1.23. Példa. Legyenek adottak az $\mathbf{A}_{11}, \mathbf{B}_{11} \in \mathbb{R}^{k_1\times n_1}$; $\mathbf{A}_{12}, \mathbf{B}_{12} \in \mathbb{R}^{k_1\times n_2}$; $\mathbf{A}_{21}, \mathbf{B}_{21} \in \mathbb{R}^{k_2\times n_1}$ és $\mathbf{A}_{22}, \mathbf{B}_{22} \in \mathbb{R}^{k_2\times n_2}$ mátrixok. Ekkor az

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

mátrixok összegét az

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

képlettel számíthatjuk ki. Az $\mathbf{A}$ mátrix α-szorosa ($\alpha \in \mathbb{R}$)

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

Az $\mathbf{A}$ és $\mathbf{B}$ mátrix akkor összeszorozható, ha $k_1 = n_1$ és $k_2 = n_2$, azaz ha a blokkmátrix főátlójában álló mátrixok kvadratikusak. □

5.1.13. Feladat. Legyenek adottak az $\mathbf{A}_{11} \in \mathbb{R}^{k_1 \times n_1}, \mathbf{A}_{12} \in \mathbb{R}^{k_1 \times n_2}, \mathbf{A}_{21} \in \mathbb{R}^{k_2 \times n_1}$ és $\mathbf{A}_{22} \in \mathbb{R}^{k_2 \times n_2}$ mátrixok, valamint az $\overline{\mathbf{x}}_1 \in \mathbb{R}^{m_1}$ és $\overline{\mathbf{x}}_2 \in \mathbb{R}^{m_2}$ vektorok. Milyen feltételek mellett lehet blokkmátrixokként összeszorozni az

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \overline{\mathbf{x}} = \begin{bmatrix} \overline{\mathbf{x}}_1 \\ \overline{\mathbf{x}}_2 \end{bmatrix}$$

blokkmátrixokat?

Ellenőrző kérdések

1. Definiáljuk a mátrixokat!

2. Hogyan definiáljuk az összeadást a mátrixok között? Soroljuk fel az összeadás tulajdonságait.

3. Hogyan definiáljuk mátrixok valós számmal való szorzatát? Soroljuk fel a valós számmal szorzás tulajdonságait.

4. Hogyan definiáljuk mátrixok szorzatát? Milyen tulajdonságait tanultuk a mátrixszorzásnak?

5. Mit hívunk egy mátrix főátlójának? Mikor diagonális egy mátrix?

6. Mit értünk egy mátrix transzponáltján? Milyen szabályokat tanultunk a transzponált képzésével kapcsolatban?

7. Milyen mátrixokat nevezünk szimmetrikus ill. ferdén szimmetrikus mátrixoknak?

8. Hogyan lehet felírni egy kvadratikus mátrixot egy szimmetrikus és egy ferdén szimmetrikus mátrix összegeként?

9. Hogyan értelmezzük két vektor skaláris szorzatát? Soroljuk fel a skaláris szorzás tulajdonságait!

10. Hogyan értelmezzük a vektorok hosszát és szögét?

11. Mik a blokkmátrixok és hogyan végzünk velük műveleteket?

A feladatok megoldásai

5.1.1.
A mátrixnak 2 sora van és három oszlopa. A 2. sor 3. eleme 0.3. A 2. sor összege -3.7.

5.1.2.
A műveletek elvégzése után a következő mátrixokat nyerjük.

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 5 & 0 & 1 \\ -6 & 0 & 0.6 \\ 3 & 2 & 8 \end{bmatrix}, \quad 5\mathbf{A} = \begin{bmatrix} 10 & -5 & 2.5 \\ -20 & 0 & 1.5 \\ 10 & 10 & 5 \end{bmatrix}$$

5.1.3.
A műveletek elvégzése után a következő vektorokat nyerjük.

$$\overline{\mathbf{a}} + \overline{\mathbf{b}} = \begin{bmatrix} 0 \\ -9 \\ 3 \end{bmatrix}, \quad -2\overline{\mathbf{b}} = \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix}$$

5.1.4.
Az $\mathbf{A}$ és $\mathbf{B}$ mátrixok ebben a sorrendben összeszorozhatók, hiszen az $\mathbf{A}$ oszlopszáma (3) megegyezik $\mathbf{B}$ sorszámával (szintén 3). A szorzat egy 2×2-es mátrix lesz. Első sorának első elemét úgy kapjuk meg, hogy $\mathbf{A}$ első sorának ill. $\mathbf{B}$ első oszlopának elemeit páronként összeszorozzuk, és az eredményeket összeadjuk. Ezek szerint pl. $(\mathbf{AB})_{11} = 2 \cdot 1 - 1 \cdot 3 + 2 \cdot (-3) = -7$. A többi elemet is hasonlóan számolva kapjuk, hogy

$$\mathbf{AB} = \begin{bmatrix} -7 & -1 \\ 9 & -3 \end{bmatrix}.$$

A $\mathbf{BA}$ szorzatra a

$$\mathbf{BA} = \begin{bmatrix} -1 & -2 & 3 \\ 9 & -2 & 5 \\ -3 & 4 & -7 \end{bmatrix}$$

mátrixot kapjuk.

5.1.5.
Diagonális mátrixok pl. a

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

mátrixok és nem diagonálisak pl.

$$\begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & -5 \end{bmatrix}.$$

5.1.6.
Számítsuk ki az $(\mathbf{A} + \mathbf{B})(\mathbf{A} - \mathbf{B})$ szorzatot a disztributivitási szabály felhasználásával.

$$(\mathbf{A} + \mathbf{B})(\mathbf{A} - \mathbf{B}) = \mathbf{A}(\mathbf{A} - \mathbf{B}) + \mathbf{B}(\mathbf{A} - \mathbf{B}) = \mathbf{A}^2 - \mathbf{AB} + \mathbf{BA} - \mathbf{B}^2.$$

Innét látható, hogy az adott egyenlőség akkor és csak akkor teljesül, ha az $\mathbf{A}$ és $\mathbf{B}$ mátrixok kommutálnak.

5.1.7.
A főátló az azonos sor és oszlopindexű elemek összessége. Ezt most a 2 és 0 elemek alkotják. A transzponált az

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

mátrix lesz. Ennek főátlója megegyezik $\mathbf{A}$ főátlójával.

5.1.8.
Egy ferdén szimmetrikus mátrix transzponálásakor a főátlóban szereplő elemek nem változnak, ugyanakkor a ferdén szimmetrikusság definíciója miatt (-1)-gyel kell szorzódniuk. Ez csak úgy lehet, ha minden főátlóbeli elem nulla.

5.1.9.
Az $\mathbf{A}$ mátrix nem lehet sem szimmetrikus, sem ferdén szimmetrikus mert nem kvadratikus mátrix. $\mathbf{B}$ nem lehet ferdén szimmetrikus, mert ahhoz a főátlójában nulláknak kellene

szerepeni. Szimmetrikussá úgy egészíthető ki, hogy $b_{13} = 1, b_{21} = 2, b_{23} = -1$ és b_{33} tetszőleges valós szám lehet. A $\mathbf{C}$ mátrix esetén az elemválasztás: $c_{12} = 3, c_{31} = -1, c_{32} = 1$ és c_{22} ill. c_{33} tetszőleges valós számok lehetnek, ill. $c_{22} = c_{33} = 0$ és $c_{12} = -3, c_{31} = 1, c_{32} = -1$.

5.1.10.
Az 5.1.14. tétel alapján

$$\mathbf{S} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}.$$

Ezek összege tényleg az $\mathbf{A}$ mátrixot adja.

5.1.11.

$$\mathbf{A}\overline{\mathbf{x}} = \begin{bmatrix} 4 \\ -3 \\ 5 \end{bmatrix}.$$

5.1.12.
A mátrixszorzás definícióját vektorokra alkalmazva kapjuk, hogy $\overline{\mathbf{x}}^\top\overline{\mathbf{y}} = [11]$ és

$$\overline{\mathbf{x}}\overline{\mathbf{y}}^\top = \begin{bmatrix} -2 & -2 & -3 \\ 4 & 4 & 6 \\ 6 & 6 & 9 \end{bmatrix}.$$

Az $\overline{\mathbf{x}}$ és $\overline{\mathbf{y}}$ vektorok skaláris szorzata nem más, mint az $\overline{\mathbf{x}}^\top\overline{\mathbf{y}}$ egyelemű oszlopvektor egyetlen eleme, azaz 11.

5.1.13.
Két blokkmátrix akkor szorozható össze, ha az első mátrix első sorában álló mátrixok oszlopainak száma (most n_1 és n_2) rendre megegyezik a második mátrix első oszlopában álló mátrixok sorainak számával (most m_1 és m_2). Az összeszorozhatóság feltételei tehát $m_1 = n_1$ és $m_2 = n_2$. A szorzat az

$$\mathbf{A}\overline{\mathbf{x}} = \begin{bmatrix} \mathbf{A}_{11}\overline{\mathbf{x}}_1 + \mathbf{A}_{12}\overline{\mathbf{x}}_2 \\ \mathbf{A}_{21}\overline{\mathbf{x}}_1 + \mathbf{A}_{22}\overline{\mathbf{x}}_2 \end{bmatrix}$$

2×1-es blokkmátrix lesz.

5.2. Mátrixok determinánsa

Ebben a fejezetben minden négyzetes mátrixhoz hozzárendelünk egy-egy valós számot, amit a mátrix determinánsának fogunk hívni. Egy mátrix determinánsának segítségével könnyen meg tudunk oldani majd kisebb egyenletrendszereket, el tudjuk dönteni, hogy egy mátrixnak van-e inverze ill. egy mátrix sajátértékeinek kiszámítására is használhatjuk.

Bevezetőként tekintsük az

$$\left.\begin{array}{rcrcl} a_{11}\cdot x & + & a_{12}\cdot y & = & b_1 \\ a_{21}\cdot x & + & a_{22}\cdot y & = & b_2 \end{array}\right\}$$

egyenletrendszert. Célunk egy olyan képlet meghatározása, amely az $a_{11}, a_{12}, a_{21}, a_{22}, b_1$ és b_2 együtthatók segítségével megmondja, hogy mely x és y számok lesznek az egyenletrendszer megoldásai. Szorozzuk meg az első egyenletet a_{22}-vel, a másodikat a_{12}-vel, majd vonjuk ki az első egyenletből a másodikat. Ekkor az

$$(a_{11}\cdot a_{22} - a_{12}\cdot a_{21})x = b_1\cdot a_{22} - a_{12}\cdot b_2$$

egyenlőséghez jutunk, melyből $a_{11}\cdot a_{22} - a_{12}\cdot a_{21} \neq 0$ (ekkor egyértelmű megoldása van az egyenletrendszernek) esetén x értékére, és hasonlóan eljárva y értékére a következőket kapjuk

$$x = \frac{b_1\cdot a_{22} - a_{12}\cdot b_2}{a_{11}\cdot a_{22} - a_{12}\cdot a_{21}}, \qquad y = \frac{a_{11}\cdot b_2 - b_1\cdot a_{21}}{a_{11}\cdot a_{22} - a_{12}\cdot a_{21}}.$$

Vegyük észre, hogy mindkét képlet számlálójában és nevezőjében is hasonló kifejezések állnak. Ha elkészítjük az ismeretlenek együtthatóiból álló

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{5.2.1}$$

mátrixot, akkor láthatjuk, hogy a képletek nevezőit úgy lehet kiszámolni, hogy a mátrix főátlójában lévő elemek szorzatából ki kell vonni a másik két elem szorzatát. Az így számított értéket a 2×2-es mátrix determinánsának fogjuk hívni. A számlálók is hasonlóan számíthatóak a

$$\begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix} \quad \text{ill.} \quad \begin{bmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{bmatrix} \tag{5.2.2}$$

mátrixokkal.

5.2.1. Feladat. Oldjuk meg a

$$\left.\begin{array}{rcrcl} 3\cdot x & + & 2\cdot y & = & 1 \\ x & - & y & = & 2 \end{array}\right\}$$

egyenletrendszert a 2×2-es mátrixok determinánsának segítségével.

Természetesen a három ismeretlenes és három egyenletből álló egyenletrendszerek vizsgálatával kitalálhatnánk, hogy hogyan lenne érdemes definiálni a 3×3-as mátrixok determinánsát. Hasonlóan járhatnánk el tetszőleges 4×4-es mátrixok esetén is. Ezen determinánsok szabályszerűségeit megfigyelve juthatunk el a determinánsok általános definíciójához, melyet most mi is ismertetünk. A mátrixok determinánsa persze nem csak

egyenletrendszerek megoldása esetén játszik fontos szerepet. Használni fogjuk majd pl. annak eldöntésére, hogy egy mátrixnak van-e inverze vagy nincs ill. mátrixok sajátértékeinek kiszámítására is. Először definiáljuk a permutáció fogalmát.

5.2.1. Definíció. Az $1, \ldots, n$ egész számoknak egy tetszőleges $i_1, \ldots, i_n$ sorrendjét az $1, \ldots, n$ számok egy **permutációjának** hívjuk. $\diamondsuit$

5.2.2. Tétel. *Az $1, \ldots, n$ egész számok összes lehetséges permutációinak száma $n!$.*

Bizonyítás. A bizonyítás teljes indukcióval történik. Az állítás nyilvánvalóan igaz egy elem esetére ($n = 1$), hiszen azt csak $1! = 1$ különböző módon lehet sorbarakni. Most tegyük fel, hogy igaz az állítás $n = k$ esetén és igazoljuk, hogy $n = k+1$-re is igaz. Vegyük az $1, \ldots, k$ számok egy sorbarendezését. Ilyenből $k!$ van az indukciós feltevés szerint. A $(k+1)$-es számot $k+1$ helyre illeszthetjük be ebbe a sorrendbe: a számsor elejére, végére ill. a köztük lévő $k-1$ helyre. Így $k+1$ elemet $k! \cdot (k+1) = (k+1)!$-féleképpen lehet sorbarakni. ■

5.2.3. Definíció. Azt mondjuk, hogy az $1, \ldots, n$ számok egy $i_1, \ldots, i_n$ permutációjában az i_l és i_k $(l < k)$ számok **inverzióban** állnak egymással, ha $i_l > i_k$. Az $i_1, \ldots, i_n$ permutáció inverzióinak számát $\mathcal{I}(i_1, \ldots, i_n)$ fogja jelölni. $\diamondsuit$

Akkor lesz tehát két szám inverzióban egymással, ha a nagyobb szám megelőzi a permutációban a kisebbet.

5.2.2. Feladat. Határozzuk meg, hogy a 4,3,5,6,1,2 és 6,5,4,3,1,2 permutációkban mekkora az inverziók száma!

5.2.4. Definíció. Az $\mathbf{A} = [a_{ij}]_{n \times n}$ négyzetes mátrixhoz rendelt

$$\det(\mathbf{A}) = \sum_{i_1, \ldots, i_n} (-1)^{\mathcal{I}(i_1, \ldots, i_n)} \cdot a_{1i_1} \cdot a_{2i_2} \cdot \ldots \cdot a_{ni_n}$$

számot, ahol az összegzést az $1, \ldots, n$ számok összes $i_1, \ldots, i_n$ permutációjára kell elvégezni, az **A mátrix determinánsának** nevezzük. $\diamondsuit$

A determináns számítási módja, vagy más szóval kifejtése, tehát a következő. A mátrix első sorából kijelölünk egy elemet. A második sorból is választunk egy elemet ügyelve arra, hogy ne essen az első sorból választott elemmel egy oszlopba. Hasonlóan járunk el a harmadik sorral is egy olyan elemet választva, amely nincs azokban az oszlopokban, ahonnét az első és második sorbeli elemet választottuk. Az eljárást folytatjuk a mátrix utolsó soráig, ahol már nem lesz választási lehetőségünk, hiszen egy oszlop kivételével az összes foglalt lesz már. A kiválasztott elemeket összeszorozzuk, majd ezt a szorzatot $+1$-gyel vagy (-1)-gyel szorozzuk aszerint, hogy az oszlopindexek sorozatában az inverziók száma páros vagy páratlan. Ezzel elkészül a determinánst definiáló összeg egy tagja. A többi tagot úgy kapjuk, hogy a fenti eljárást megismételjük az összes lehetséges különböző sorrendben választva az oszlopokat. Összesen tehát $n!$ tag fog szerepelni a kifejtésben.

5.2.5. Példa. Az

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

mátrix determinánsa $\det(\mathbf{A}) = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$. Ha az első sorból az első elemet választjuk ki, akkor a másodikból már csak a másodikat lehet. Az oszlopindexek tehát 1 és 2, azaz az inverziók száma nulla. Az $a_{11} \cdot a_{22}$ tényező tehát pozitív előjellel szerepel a determináns kifejtésében. Ha az első sorból a második elemet választjuk ki, akkor a másodikból már csak az elsőt lehet. Az oszlopindexek tehát 2 és 1, azaz az inverziók száma egy. Az $a_{12} \cdot a_{21}$ tényező tehát negatív előjellel szerepel. Mivel az oszlopindexek csak kétféleképpen permutálhatók, ezért a determináns kifejtésében több tag nincsen, azaz $\det(\mathbf{A}) = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$. □

5.2.6. Példa. Az

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

mátrix determinánsa

$$\det(\mathbf{A}) = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \quad (5.2.3)$$

Egy 3×3-as mátrix determinánsa általában $3! = 6$ tagból fog állni, hiszen ennyiféleképpen lehet az $1, 2$ és 3 oszlopindexeket permutálni. Válasszuk mondjuk az első sorból a harmadik elemet, a másodikból a másodikat és a harmadikból az elsőt. Ezek oszlopindexei rendre $3, 2$ és 1. Mivel $\mathcal{I}(3, 2, 1) = 3$, emiatt a determinánst definiáló összegben az $a_{13} \cdot a_{22} \cdot a_{31}$ szorzat (-1)-es szorzóval fog szerepelni. Hasonlóan kaphatjuk meg a másik öt tag előjelét is.

Hogy ne kelljen mindig végiggondolni az inverziószámot az egyes szorzatok esetében amikor egy 3×3-as mátrix determinánsát számoljuk, célszerű megjegyezni az ún. Sarrus-szabályt. Írjuk le az $\mathbf{A}$ mátrix első két oszlopának elemeit jobbra a mátrix mellé.

$$\begin{bmatrix} a_{11} & & a_{12} & & a_{13} \\ & \searrow & & \diagdown\!\!\searrow & \\ a_{21} & & a_{22} & & a_{23} \\ & \nearrow & & \diagdown\!\!\searrow & \\ a_{31} & & a_{32} & & a_{33} \end{bmatrix} \begin{matrix} \diagdown\!\!\searrow \\ \\ \diagdown\!\!\searrow \\ \\ \end{matrix} \begin{matrix} a_{11} & & a_{12} \\ & \nearrow & \\ a_{21} & & a_{22} \\ & \searrow & \\ a_{31} & & a_{32} \end{matrix}$$

Induljunk el a mátrix felső sorának azon elemeiből, melyekből indul ki nyíl és kövessük a nyilakat az alsó sorig (a főátlón és még két, a főátlóval párhuzamos egyenes mentén haladunk). Szorozzuk össze az érintett elemeket. Most induljunk el az alsó sorból, szintén kövessük a nyilakat és szorozzuk össze az érintett elemeket. Ha most az első három szorzatot pozitív, míg a második hármat negatív előjelekkel összeadjuk, akkor így éppen a mátrix determinánsát kapjuk. Végül külön felhívjuk a figyelmet arra, hogy a Sarrus-szabály csak 3×3-as mátrixokra alkalmazható. □

5.2.3. Feladat. Számítsuk ki a mátrixok determinánsát.

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

Egy $\mathbf{A}$ mátrix determinánsának jelölésére a $\det(\mathbf{A})$ jelölésen kívül használjuk még az $|\mathbf{A}|$ jelölést is. Ekkor a mátrix két szögletes zárójelét el szoktuk hagyni az egyszerűség

kedvéért. Tehát pl. a

$$\begin{bmatrix} 2 & 3 \\ -2 & 1 \end{bmatrix}$$

mátrix determinánsát

$$\begin{vmatrix} 2 & 3 \\ -2 & 1 \end{vmatrix}$$

fogja jelölni.

A determináns definíciójában nincs szükségünk az inverziók pontos számára, hiszen csak azt kell megállapítanunk, hogy páros vagy páratlan sok inverzió van-e a permutációban. Ezt egyszerűbben is el lehet dönteni, mint az inverziók számának pontos meghatározása.

5.2.7. Tétel. *Az $1, \ldots, n$ számok egy permutációját akkor és csak akkor tudjuk páros sok cserével visszaalakítani a természetes $1, \ldots, n$ sorrendre, ha a permutációban az inverziók száma páros volt. Akkor és csak akkor szükséges páratlan sok csere a visszaalakításhoz, ha az inverziók száma páratlan volt.*

Bizonyítás. Ha két szomszédos elemet felcserélünk a permutációban, akkor az inverziók számának paritása (páros vagy páratlan volta) megváltozik. Ha eddig inverzióban voltak, akkor a csere után már nem lesznek, ill. ha nem voltak inverzióban, akkor a csere után már inverzióban fognak állni. A többi elemmel nem változik az inverzió. Ha két tetszőleges elemet akarunk felcserélni a permutációban, akkor azt megtehetjük úgy, hogy több lépésben mindig csak szomszédos elemeket cserélünk. Ha pl. az i_l és i_k $(l < k)$ elemeket akarjuk kicserélni, akkor először az l-edik elemet kicseréljük az $l+1$-edikkel. Ezután az $l+1$-ediket az $l+2$-edikkel, stb. Ezt ismételjük, míg el nem jutunk a $k-1$-edik és k-adik elemig, majd visszafelé cserélgetünk addig, amíg az eredetileg k-adik elem az l-edik helyre nem kerül. Ezalatt összesen $2(k-l)-1$, azaz páratlan sok szomszédcserét fogunk végrehajtani. Összefoglalva tehát, minden egyes elemcsere a permutációban megváltoztatja az inverziók számának paritását. Így egy permutációt akkor és csak akkor tudunk páros sok cserével a természetes sorrendbe hozni (melynek inverziója nulla, azaz páros), ha az eredeti inverziószám páros volt. A permutációt akkor és csak akkor tudjuk páratlan sok cserével a természetes sorrendbe hozni, ha az eredeti inverziószám is páratlan volt. ■

5.2.4. Feladat. Döntsük el, hogy páros vagy páratlan a $6, 2, 3, 4, 5, 1$ és $5, 3, 4, 2, 1$ permutációk inverziószáma.

Ahogy már szó volt róla, egy $n \times n$-es mátrix determinánsa $n!$ darab n-tényezős szorzat összegeként számítható ki. Ez nem is okoz nagy problémát, ha kisméretű mátrixról van szó ($n \leq 4$). Ha mondjuk egy 10×10-es mátrix determinánsát szeretnénk a definíció segítségével meghatározni, akkor $10! = 3628800$ 10-tényezős szorzatot kellene összeadni, ami számítógépek segítsége nélkül reménytelen feladat. Egy 20×20-as mátrix determinánsának a definíció szerinti számolása pedig már egy nagyon gyors számítógépnek is több hónapos munkájába telne. Ezek alapján sejthetjük, hogy a determináns értékét nem annak definíciója alapján szoktuk számítani. Ez így is van. Általában azt a módszert követjük, hogy a mátrixot valamilyen módon átalakítjuk egy olyan mátrixszá, amelynek determinánsa egyenlő az eredetiével (vagy könnyen követhető, hogy az átalakítás milyen változással jár a determináns értékében), de sokkal könnyebben számítható. Ezekben az átalakításokban segítenek a következő tételek.

5.2.8. Tétel. *Egy* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *mátrixnak és transzponáltjának ugyanaz a determinánsa, azaz* $\det(\mathbf{A}) = \det(\mathbf{A}^\top)$.

Bizonyítás. A determináns kifejtésének mindegyik tagja egy olyan szorzat, amely a mátrix minden sorából és oszlopából csak egy elemet tartalmaz. Emiatt ez a szorzat szerepelni fog a transzponált determinánsának kifejtésében is. Így csak azt kell megmutatnunk, hogy a két szorzat ugyanolyan előjellel fog szerepelni. Ez abból következik, hogy azon cserék számának paritása, melyek előállítják a természetes sorrendet egy permutációból, megegyezik azon cserék paritásával, melyekkel a természetes sorrendből az adott permutációt lehet előállítani. ■

A most következő állításokat elegendő lesz vagy sorokra vagy oszlopokra bizonyítani, mert az 5.2.8. tétel miatt az állítások helyessége ekkor oszlopokra ill. sorokra automatikusan adódik.

5.2.9. Tétel. *Ha egy mátrix valamelyik sora vagy oszlopa csupa 0 elemekből áll, akkor a mátrix determinánsa 0.*

Bizonyítás. A tételt sorokra igazoljuk. A determináns kifejtésében szereplő szorzatok minden sorból tartalmaznak egy-egy elemet. Így minden szorzat tartalmazni fog nulla tényezőt, azaz mindegyik szorzat nulla. Ekkor a determináns értéke is nulla. ■

5.2.10. Tétel. *Ha egy mátrix összes főátló alatti (ill. feletti) eleme nulla (ún. háromszögmátrix), akkor a mátrix determinánsa a főátlójában lévő elemek szorzata.*

Bizonyítás. Tegyük fel, hogy az $\mathbf{A}$ mátrix főátló alatti összes eleme nulla. Ekkor a determináns kifejtésében a szorzatok képzése során az utolsó sorból csak az utolsó elemet érdemes választani (különben nulla a szorzat), az utolsó előttiből az utolsót nem választhatjuk, így az utolsó előttit kell választanunk, tehát a főátlóban lévő elemet. Hasonlóan gondolkodva kapjuk, hogy csak akkor lehet az elemek szorzata nem nulla, ha a főátlóbeli elemeket szorozzuk össze. Mivel az oszlopindexek ekkor a természetes sorrendben követik egymást, ezért az inverziók száma nulla. Így a determináns valóban a főátlóbeli elemek szorzata lesz. ■

5.2.11. Tétel. *Ha egy mátrix két sorát vagy oszlopát felcseréljük, akkor a determinánsának értéke* (-1)*-szeresére változik.*

Bizonyítás. A tételt oszlopokra igazoljuk. Két oszlop felcserélésekor a determináns kiszámításában ugyanazok a szorzatok fognak szerepelni, de az oszlopcsere miatt a szorzatokhoz tartozó inverziók számának paritása minden tagban megváltozik, így minden tag (-1)-gyel szorzódik. Ebből következik, hogy a determináns értéke is (-1)-gyel fog szorzódni. ■

5.2.12. Tétel. *Ha egy mátrix két sora vagy oszlopa megegyezik, akkor a determinánsának értéke nulla.*

Bizonyítás. Cseréljük fel a mátrix azon két sorát vagy oszlopát, melyek megegyeznek. Ekkor egyrészt a determináns nem változik, mivel ugyanazt a mátrixot kapjuk a csere után is, másrészt (-1)-szeresére változik az előző tétel miatt. Ez csak úgy lehet, ha a mátrix determinánsa nulla. ■

5.2.13. Tétel. *Ha egy mátrix egy sorát vagy oszlopát egy adott valós számmal megszorozzuk, akkor a determinánsának értéke is azzal a valós számmal szorzódik.*

Bizonyítás. Bizonyítsuk a tételt sorokra. Ha valamelyik sort végigszorozzuk az α valós számmal, akkor a determináns kiszámításakor az ebből a sorból választott elem az eredetileg választott α-szorosa lesz. Így a determináns kifejtésében minden tag α-val fog szorzódni, azaz a determináns maga is α-val szorzódik. ■

5.2.14. Tétel. *Ha egy mátrix valamelyik sorának (ill. oszlopának) minden eleme egy kéttagú összeg, akkor determinánsa egyenlő azon két mátrix determinánsának összegével, melyek közül az első megfelelő sora (ill. oszlopa) az összegek első tagjait, a második a második tagokat tartalmazza; a többi elem változatlan.*

Bizonyítás. A tételt sorokra bizonyítjuk. Tegyük fel, hogy az $\mathbf{A}$ mátrix első sora kéttagú összegekből áll. Legyen $a_{1i} = a_{i1}^{(1)} + a_{i1}^{(2)}$ $(i = 1, \dots, n)$. Ekkor a mátrix determinánsa a

$$\det(\mathbf{A}) = \sum_{i_1,\dots,i_n} (-1)^{\mathcal{I}(i_1,\dots,i_n)} \cdot (a_{1i_1}^{(1)} + a_{1i_1}^{(2)}) \cdot a_{2i_2} \cdot \ldots \cdot a_{ni_n} =$$

$$= \sum_{i_1,\dots,i_n} (-1)^{\mathcal{I}(i_1,\dots,i_n)} \cdot a_{1i_1}^{(1)} \cdot a_{2i_2} \cdot \ldots \cdot a_{ni_n} + \sum_{i_1,\dots,i_n} (-1)^{\mathcal{I}(i_1,\dots,i_n)} a_{1i_1}^{(2)} \cdot a_{2i_2} \cdot \ldots \cdot a_{ni_n}$$

alakban írható (az összegzést az $1, \dots, n$ számok összes $i_1, \dots, i_n$ permutációjára kell elvégezni). Az összeg első tagja egy olyan mátrix determinánsa, melynek az első sorában az $a_{1i}^{(1)}$ számok állnak. Hasonlóan a második tag egy olyan mátrix determinánsa, melynek első sorában az $a_{1i}^{(2)}$ számok találhatóak $(i = 1, \dots, n)$. Ezzel a tételt bebizonyítottuk. ■

5.2.15. Tétel. *Ha egy mátrix egyik sorának (oszlopának) számszorosát hozzáadjuk egy másik sorhoz (oszlophoz), akkor a mátrix determinánsa nem változik meg.*

Bizonyítás. A tételt sorokra bizonyítjuk. Tegyük fel pl. az általánosság megszorítása nélkül, hogy az $\mathbf{A}$ mátrix első sorának α-szorosát hozzáadjuk a második sorához. Ekkor az átalakított mátrix második sorában az $a_{2i} + \alpha a_{1i}$ $(i = 1, \dots, n)$ számok fognak állni. Ezek kéttagú összegek, így az előző tételt és az 5.2.13. tételt alkalmazva

$$\sum_{i_1,\dots,i_n} (-1)^{\mathcal{I}(i_1,\dots,i_n)} \cdot a_{1i_1} \cdot (a_{2i_2} + \alpha a_{1i_2}) \cdot \ldots \cdot a_{ni_n} =$$

$$= \sum_{i_1,\dots,i_n} (-1)^{\mathcal{I}(i_1,\dots,i_n)} \cdot a_{1i_1} \cdot a_{2i_2} \cdot \ldots \cdot a_{ni_n} + \sum_{i_1,\dots,i_n} (-1)^{\mathcal{I}(i_1,\dots,i_n)} \cdot a_{1i_1} \cdot (\alpha a_{1i_2}) \cdot \ldots \cdot a_{ni_n} =$$

$$= \sum_{i_1,\dots,i_n} (-1)^{\mathcal{I}(i_1,\dots,i_n)} \cdot a_{1i_1} \cdot a_{2i_2} \cdot \ldots \cdot a_{ni_n} + \alpha \sum_{i_1,\dots,i_n} (-1)^{\mathcal{I}(i_1,\dots,i_n)} \cdot a_{1i_1} \cdot a_{1i_2} \cdot \ldots \cdot a_{ni_n}.$$

Az első tag éppen az $\mathbf{A}$ mátrix determinánsa, a második pedig egy olyan mátrix determinánsának α-szorosa, melynek második sora megegyezik az első sorával, azaz nulla (lsd. 5.2.13. tétel). ■

5.2.16. Tétel. *Legyen* $\mathbf{A}$ *és* $\mathbf{B}$ *két azonos típusú kvadratikus mátrix. Ekkor*

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B}).$$

Tekintsük a 3×3-as mátrixok determinánsára nyert

$$\det(\mathbf{A}) = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

kifejezést (lásd (5.2.3)). Ez a determináns kiemelések után a

$$\det(\mathbf{A}) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

alakban írható. Vegyük észre, hogy a tagok első tényezői éppen a mátrix sorainak első elemei, a zárójelben szereplő kifejezések pedig valamilyen 2×2-es mátrixok determinánsai. Pl. az első zárójelben annak a 2×2-es mátrixnak a determinánsa szerepel, melyet az $\mathbf{A}$ mátrix első sorának és oszlopának elhagyása után kapunk.

5.2.17. Definíció. Az $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ kvadratikus mátrixnak hagyjuk el az i-edik sorát és j-edik oszlopát. Az így nyert $(n-1)\times(n-1)$-es mátrix determinánsát az a_{ij} elemhez tartozó **aldeterminánsnak** hívjuk. Szorozzuk be az aldetmináns értékét a $(-1)^{i+j}$ kifejezéssel. Az így kapott értéket az a_{ij} elemhez tartozó **előjeles aldeterminánsnak** hívjuk és $A_{(ij)}$-vel jelöljük $(i, j = 1, \ldots, n)$. $\diamondsuit$

A $(-1)^{i+j}$ kifejezéssel való szorzás igazából csak azt adja meg, hogy a kiszámított aldeterminánsnak meg kell-e változtatni az előjelét vagy sem. Az előjelek megváltoztatásának módját célszerű az ún. sakktábla szabály alapján megjegyezni. Ha egy sakktábla alapján osztjuk ki az elemekhez tartozó előjeleket (bal felő sarokban + jel van)

$$\begin{bmatrix} + & - & + & - & \ldots \\ - & + & - & + & \ldots \\ + & - & + & - & \ldots \\ - & + & - & + & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

akkor az aldeterminánsokat a mátrixban feltüntetett előjelekkel szorozva kapjuk meg az előjeles aldeterminánsokat.

5.2.18. Tétel. *(Kifejtési tétel.) Legyen $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ egy tetszőleges kvadratikus mátrix, és legyen $1 \le k \le n$ egy rögzített index. Ekkor*

$$\det(\mathbf{A}) = a_{k1}A_{(k1)} + a_{k2}A_{(k2)} + \ldots + a_{kn}A_{(kn)}$$

ill.

$$\det(\mathbf{A}) = a_{1k}A_{(1k)} + a_{2k}A_{(2k)} + \ldots + a_{nk}A_{(nk)}.$$

A tételt nem bizonyítjuk. A kifejtési tétel szerint tehát egy mátrix determinánsát úgy is megkaphatjuk, ha valamelyik sorának ill. oszlopának elemeit rendre beszorozzuk az elemekhez tartozó előjeles aldeterminánsok értékeivel és az így nyert szorzatokat összeadjuk.

5.2.5. Feladat. Határozzuk meg az

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

mátrix determinánsát a kifejtési tétel segítségével.

Két példát mutatunk arra, hogy a bizonyított tételek segítségével hogyan lehet mátrixok determinánsát kiszámítani.

5.2.19. Példa. Most meg fogjuk határozni az

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & -1 & 2 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 5 & 6 \\ -9 & 5 & 3 & 2 & 1 \end{vmatrix}$$

determináns értékét. Adjuk hozzá a mátrix első sorát a harmadikhoz. Az így kapott új mátrix

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & -1 & 2 \\ 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 \\ -9 & 5 & 3 & 2 & 1 \end{vmatrix}$$

determinánsa megegyezik a régiével (5.2.15. tétel). Ennek a determinánsnak az értékéről viszont egyből látszik, hogy nulla, mivel a harmadik és negyedik sora ugyanaz (5.2.12. tétel). □

5.2.20. Példa. Meghatározzuk az

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & -1 & 2 \\ 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 4 & 5 & 6 \\ -9 & 5 & 3 & 2 & 1 \end{vmatrix}$$

determinánst. Célként azt tűzzük ki magunk elé, hogy úgy alakítsuk át a mátrixot, hogy a főátlója alatt már csak nullák álljanak, mert ekkor a determináns értéke éppen a főátlóbeli elemek szorzata lesz (5.2.10. tétel). Vonjuk ki először a mátrix első sorának kétszeresét a második sorból. Az így nyert mátrix determinánsa megegyezik az eredetiével (5.2.15. tétel). Hasonlóan eljárva, a következő átalakítások sem változtatják meg a determinánst. Az első sor kétszeresét vonjuk ki a harmadikból, az első sort vonjuk ki a negyedikből, végül az első sor kilencszeresét adjuk hozzá az ötödikhez. Ezekután elmondhatjuk, hogy az

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -1 & -9 & -8 \\ 0 & -1 & -2 & -3 & -4 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 23 & 30 & 38 & 46 \end{vmatrix}$$

determináns értéke megegyezik az eredeti determináns értékével. Célunkat részben elértük, hiszen az első oszlopban már nullák állnak a főátló alatt. Vonjuk ki a második sort a harmadikból, adjuk hozzá a negyedikhez ill. 23-szorosát adjuk az ötödikhez.

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -1 & -9 & -8 \\ 0 & 0 & -1 & 6 & 4 \\ 0 & 0 & 0 & -8 & -7 \\ 0 & 0 & 7 & -169 & -138 \end{vmatrix}.$$

Adjuk hozzá a harmadik sor hétszeresét az ötödikhez

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -1 & -9 & -8 \\ 0 & 0 & -1 & 6 & 4 \\ 0 & 0 & 0 & -8 & -7 \\ 0 & 0 & 0 & -127 & -110 \end{vmatrix}.$$

Végül a negyedik sor $(-127/8)$-szorosát adjuk hozzá az ötödik sorhoz

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -1 & -9 & -8 \\ 0 & 0 & -1 & 6 & 4 \\ 0 & 0 & 0 & -8 & -4 \\ 0 & 0 & 0 & 0 & 9/8 \end{vmatrix}.$$

Ezen determináns értéke tehát – amely megegyezik az eredeti determináns értékével – $1 \cdot (-1) \cdot (-1) \cdot (-8) \cdot (9/8) = -9$. □

Végül ismét felhívjuk a figyelmet arra, hogy ha az utolsó példában szereplő determinánst a determináns definíciója alapján számítottuk volna ki, akkor $5! = 120$ darab ötténye zős szorzatot kellet volna összeadnunk. Ez jóval tovább tartott volna, mint a most bemutatott módszer.

Ellenőrző kérdések

1. Mit hívunk permutációnak és hogyan definiáltuk az inverziók számát egy permutációban?
2. Hogyan értelmezzük egy mátrix determinánsát?
3. Hogyan számítjuk ki egy 2×2-es és egy 3×3-as mátrix determinánsát?
4. Milyen változtatásokat lehet csinálni egy mátrixon úgy, hogy determinánsának változása nyomon követhető legyen?
5. Mit mond ki a kifejtési tétel?

A feladatok megoldásai

5.2.1.
Az együtthatókból képzett (5.2.1) mátrix determinánsa $3 \cdot (-1) - 2 \cdot 1 = -5$. Hasonlóan az (5.2.2) mátrixok determinánsának értékei rendre -5 és 5. Az egyenlet megoldásai tehát $x = (-5)/(-5) = 1$ és $y = 5/(-5) = -1$.

5.2.2.
Az első permutációban a 4-es és a 3-as inverzióban van, mert a 4-es megelőzi a 3-ast. Hasonlóan inverzióban van a 4-es az 1-gyel és a 2-vel, a 3-as az 1-gyel és a 2-vel, az 5-ös az 1-gyel és a 2-vel és a 6-os is az 1-gyel és a 2-vel. Összefoglalva: $\mathcal{I}(4,3,5,6,1,2) = 9$. Ugyanilyen okoskodással $\mathcal{I}(6,5,4,3,1,2) = 14$.

5.2.3.
$\det(\mathbf{A}) = 8,\ \det(\mathbf{B}) = -20.$

5.2.4.
Az első permutációban az 1-est és az 6-ost felcserélve a számok természetes sorrendbe kerültek. Így az 5.2.7. tételt felhasználva az inverziók száma páratlan (valóban,

$\mathcal{I}(6,2,3,4,5,1) = 9$). A másik permutációnál cseréljük fel először az 1-est és az 5-öst, majd a 2-est és a 4-est, végül a 2-es és a 3-ast. Ezzel helyreáll a természetes sorrend. Három csere kellett, tehát a permutációk száma páratlan (valóban, $\mathcal{I}(5,3,4,2,1) = 9$).

5.2.5.
Az adott mátrix esetén célszerű a kifejtési tételt az első oszlopra alkalmazni, hiszen ennek egyik eleme nulla. Az $A_{(11)}$ előjeles aldeterminánsa -1, míg $A_{(31)} = 11$. Így $\det(\mathbf{A}) = 2 \cdot (-1) + 1 \cdot 11 = 9$.

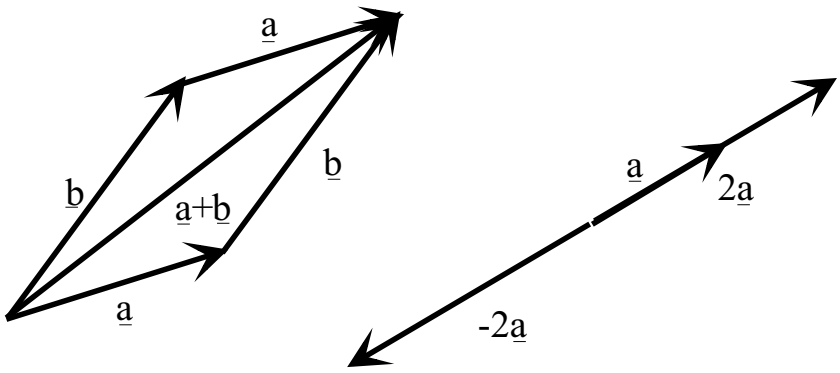
6. fejezet

Vektorterek

6.1. Vektortér fogalma, nevezetes vektorterek

Középiskolai tanulmányainkban már találkoztunk a vektorok fogalmával. A vektorokat megtanultuk összeadni és valós számokkal szorozni. Ebben a fejezetben általánosítjuk a vektorok fogalmát.

A síkbeli helyvektorok fogalma középiskolából ismert. Az ottani tanulmányaink alapján már össze tudunk adni két vektort ill. meg tudjuk szorozni őket bármilyen valós számmal. Tekintsük a 6.1.1. ábrát, melynek bal oldali részén az $\underline{a}$ és $\underline{b}$ vektorokat adtuk



6.1.1. ábra. *Síkvektorok összeadása és számmal való szorzása.*

össze. Az összeadás úgy történt, hogy a $\underline{b}$ vektor kezdőpontját az $\underline{a}$ vektor végpontjához illesztettük, a szabadon maradt kezdő- és végpontokat összekötöttük és a keletkezett szakaszt a végpont felé irányítottuk. Az így kapott vektor lesz az $\underline{a}$ és $\underline{b}$ vektorok összege. Vegyük észre, hogy ugyanezt az összegvektort kaptuk volna akkor is, ha $\underline{b}$-hez adjuk hozzá $\underline{a}$-t. Tehát a vektorösszeadás kommutatív. Egyszerűen megmutatható, hogy az összeadás asszociatív is. Van olyan vektor, amit bármelyik síkvektorhoz hozzáadhatunk anélkül hogy a síkvektor megváltozna. Ez a vektor a nullvektor. Az is igaz, hogy tetszőleges vektorhoz találhatunk olyan vektort, amivel összeadva a nullvektort kapjuk eredményül. Ez a vektor az eredeti vektorral megegyező hosszúságú, de ellenkező irányú vektor lesz.

Tekintsük most a 6.1.1. ábra jobb oldali részét. Ezen azt szemléltettük, hogy hogyan kell egy vektort valós számmal szorozni. Például a 2-vel való szorzás egyszerűen a kétszeresre nyújtást jelenti, a (-2)-vel való szorzás pedig a hossz kétszeresére nyújtásán kívül egy irányváltoztatást is jelent. Látható, hogy ha 1-gyel szorozzuk a vektort akkor az

önmaga marad. Könnyen igazolható továbbá, hogy ha egy vektort egymás után szorzunk két számmal, akkor ez egyenértékű azzal, mintha a két szám szorzatával szoroztuk volna. Két vektor összegét szorozhatjuk úgy valós számmal, hogy mindkét tagját megszorozzuk és az eredményvektorokat összeadjuk, ill. ha egy vektort két szám összegével szorzunk, akkor ugyanazt az eredményt kapjuk, mint amikor a két számmal külön-külön megszorozzuk a vektort, majd az így kapott vektorokat összeadjuk.

A síkbeli helyvektorokon kívül sok esetben találkozunk a matematikában olyan struktúrákkal, melyek a síkvektorokhoz hasonlóan "viselkednek". Ilyenek pl. a tér helyvektorai, a függvények, a polinomok vagy az 5.1. fejezetben definiált mátrixok is. Értelmezett köztük az összeadás, a valós számmal való szorzás és ezen műveletekre hasonló tulajdonságok érvényesek, mint a helyvektorok esetén. Emiatt érdemes a síkvektorokhoz hasonló tulajdonságokkal rendelkező struktúrákra külön elnevezést bevezetni. Ez az elnevezés - a síkvektorokra utalva - a vektortér.

6.1.1. Definíció. Legyen adott egy V halmaz, melynek elemei között értelmezve van egy összeadás és egy valós számmal való szorzás művelet. Azt mondjuk, hogy a V halmaz **vektorteret** (vagy **lineáris teret**) alkot a fenti két műveletre nézve, ha teljesülnek a következő tulajdonságok.

Az összeadásra:
1. $\forall \mathbf{a}, \mathbf{b} \in V$ esetén $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$,
2. $\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in V$ esetén $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$,
3. $\exists \mathbf{o} \in V$, melyre $\mathbf{a} + \mathbf{o} = \mathbf{a}$, $\forall \mathbf{a} \in V$ esetén ($\mathbf{o}$: nullvektor),
4. $\forall \mathbf{a} \in V$ elemhez $\exists \hat{\mathbf{a}} \in V$, melyre $\mathbf{a} + \hat{\mathbf{a}} = \mathbf{o}$ ($\hat{\mathbf{a}}$: $\mathbf{a}$ additív inverze);

A valós számmal való szorzásra:
5. $\forall \mathbf{a}, \mathbf{b} \in V$ és $\forall \alpha \in \mathbb{R}$ esetén $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$,
6. $\forall \mathbf{a} \in V$ és $\forall \alpha, \beta \in \mathbb{R}$ esetén $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$,
7. $\forall \mathbf{a} \in V$ és $\forall \alpha, \beta \in \mathbb{R}$ esetén $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$,
8. $\forall \mathbf{a} \in V$ esetén $1\mathbf{a} = \mathbf{a}$. $\diamondsuit$

Az, hogy értelmezve van V-n egy összeadás művelet és egy valós számmal szorzás művelet azt jelenti, hogy bármely két $\mathbf{a}, \mathbf{b} \in V$ vektor esetén az $\mathbf{a} + \mathbf{b}$ elemnek is V-ben kell lennie, továbbá, hogy bármely valós α szám és bármely $\mathbf{a} \in V$ vektor esetén $\alpha\mathbf{a}$ is V-ben van. Ezt a két tulajdonságot úgy is ki szoktuk fejezni, hogy V-nek zártnak kell lennie az összeadásra és számmal szorzásra nézve, vagy úgy, hogy az összeadás és a számmal szorzás nem vezet ki V-ből.

Egy vektortér megadásánál nem elegendő csak a V halmazt megadni, hanem azt is mindig meg kell mondani, hogy hogyan értelmezzük V elemei között a két műveletet. Ettől csak akkor szoktunk eltérni, ha már korábbról ismert a két művelet. Ekkor azt mondjuk, hogy V-n a szokásos műveletek vannak értelmezve.

A vektorterek elemeire továbbra is megtartjuk a vektor elnevezést, de ne feledjük, hogy ezek nem feltétlenül lesznek a középiskolában megszokott síkvektorok. $\mathbb{R}$ elemeire szokás a skalár kifejezést is használni.

A 3. tulajdonságban szereplő $\mathbf{o}$ elemet a vektortér nullelemének vagy nullvektorának nevezzük. A 4. tulajdonságban szereplő $\hat{\mathbf{a}}$ elem az $\mathbf{a}$ additív inverze (inverze az összeadásra nézve).

Szeretnénk felhívni a figyelmet a jelölésekre is. Ha általánosan beszélünk a vektorterekről, akkor a valós számokat mindig görög betűkkel jelöljük, a vektortér elemeit pedig megvastagított latin betűkkel. Ez segíteni fog a szereplő képletek gyorsabb átlátásában

és megértésében. Konkrét vektorterek esetén a hagyományokból eredően más jelöléseket fogunk használni, de ezekre mindig felhívjuk majd a figyelmet.

6.1.1. Feladat. Mutassuk meg, hogy egy vektortérben tetszőleges $\mathbf{a}$ elemre teljesül, hogy $\mathbf{o} + \mathbf{a} = \mathbf{a}$.

6.1.2. Feladat. Mutassuk meg, hogy egy vektortérben csak egy nullelem lehet.

6.1.3. Feladat. Mutassuk meg, hogy egy vektortérben egy vektornak csak egy additív inverze lehet.

6.1.4. Feladat. Mutassuk meg, hogy tetszőleges $\mathbf{a}$ vektor esetén $0\mathbf{a} = \mathbf{o}$.

6.1.5. Feladat. Mutassuk meg, hogy tetszőleges $\mathbf{a}$ vektor esetén $\hat{\mathbf{a}} = (-1)\mathbf{a}$.

6.1.6. Feladat. Mutassuk meg, hogy tetszőleges α valós szám esetén $\alpha\mathbf{o} = \mathbf{o}$.

Megkönnyíthetők a jelölések a 6.1.5. feladat eredményét felhasználva. Az $\mathbf{a}$ vektor inverzét jelölhetjük $-\mathbf{a}$-val. Hasonlóan pl. $\mathbf{a} - \mathbf{b}$ fogja azt a műveletet jelölni, hogy $\mathbf{a}$-hoz hozzáadjuk $\mathbf{b}$ additív inverzét. Vegyük észre, hogy ezzel tulajdonképpen kivonást definiáltunk a vektortér két eleme között.

Most példákat sorolunk fel vektorterekre. Ezek legtöbbje a későbbiek során nagyon fontos szerepet fog játszani. A vektortér-tulajdonságok teljesülésének átgondolását az Olvasóra bízzuk.

6.1.2. Példa. A legkézenfekvőbb példa vektortérre a síkbeli vagy a térbeli helyvektorok halmaza a bevezetőben ismertetett összeadásra és számmal való szorzásra nézve. □

6.1.3. Példa. Mivel a valós számok halmaza testet alkot, ezért a valós számok halmaza egyben vektortér is a szokásos összeadásra és szorzásra nézve. □

6.1.4. Példa. Az egy rögzített halmazon értelmezett függvények is vektorteret alkotnak a szokásos függvények közti összeadásra és számmal való szorzásra nézve. Ennek speciális eseteként vektorteret alkotnak pl. az $f : [0, 1] \to \mathbb{R}$ típusú függvények vagy az $f : [0, 1] \to \mathbb{R}$ típusú folytonos függvények is. □

6.1.5. Példa. Vektorteret alkotnak a legfeljebb n-edfokú polinomok is a szokásos összeadásukra és számmal való szorzásukra nézve. Az összeadásnál a polinom azonos fokszámú tagjait kell összeadni, számmal való szorzásnál pedig a polinom minden együtthatóját be kell szoroznunk az adott számmal. Az összes polinom halmaza is vektortér a most megadott műveletekre. □

6.1.6. Példa. Az 5.1. fejezet alapján az $k \times n$ típusú mátrixok is vektorteret alkotnak az ott definiált összeadásra és valós számmal való szorzásra. Ezek speciális eseteként az oszlopvektorok, az $\mathbb{R}^n$ elemei is vektorteret alkotnak. □

6.1.7. Példa. A végtelen valós számsorozatok vektorteret alkotnak a szokásos összeadásra és számmal való szorzásra. □

6.1.7. Feladat. Miért nem alkotnak vektorteret a pontosan n-edfokú polinomok?

6.1.8. Feladat. Legyen V az $f : [0,1] \to \mathbb{R}$ típusú folytonos függvények halmaza a szokásos összeadásra és számmal szorzásra nézve. Mi lesz a nulleleme ennek a vektortérnek? Mi lesz az additív inverze az $x \mapsto x(1-x)$ függvénynek?

6.1.9. Feladat. Tekintsük $\mathbb{R}^3$-ból azokat a vektorokat, melyeknek egyik eleme sem negatív. Igazoljuk, hogy ez a halmaz nem alkot vektorteret a szokásos összeadásra és számmal való szorzásra nézve.

Ellenőrző kérdések

1. Melyik két alapműveletre épül a vektortér definíciója? Ismertessük a vektortér definícióját.

2. Mi az oka a vektortér elnevezésnek?

3. Mondjunk példát vektorterekre. Mondjuk meg, hogy az egyes esetekben hogyan értelmezzük a műveleteket.

4. Milyen struktúrákat nevezünk számtesteknek?

A feladatok megoldásai

6.1.1.
Használjuk a vektortér tulajdonságait. $\mathbf{o} + \mathbf{a} = \mathbf{a} + \mathbf{o}$ az összeadás kommutativitása szerint. $\mathbf{a} + \mathbf{o} = \mathbf{a}$ a nullelem tulajdonsága szerint. Tehát $\mathbf{o} + \mathbf{a} = \mathbf{a}$.

6.1.2.
Tegyük fel indirekt, hogy két különböző nullelem van: $\mathbf{o}_1$ és $\mathbf{o}_2$ ($\mathbf{o}_1 \neq \mathbf{o}_2$). Ekkor a vektorterek 3. tulajdonságát használva az $\mathbf{o}_1 + \mathbf{o}_2 = \mathbf{o}_1$ (mivel $\mathbf{o}_2$ nullelem) és az $\mathbf{o}_2 + \mathbf{o}_1 = \mathbf{o}_2$ (mivel $\mathbf{o}_1$ nullelem) egyenlőségek is teljesülnek. Tehát $\mathbf{o}_2 = \mathbf{o}_2 + \mathbf{o}_1 = \mathbf{o}_1 + \mathbf{o}_2 = \mathbf{o}_1$, azaz $\mathbf{o}_1 = \mathbf{o}_2$, ami ellentmond az indirekt feltevésnek.

6.1.3.
Tegyük fel indirekt, hogy az $\mathbf{a}$ vektornak két különböző additív inverze is van: $\hat{\mathbf{a}}_1$ és $\hat{\mathbf{a}}_2$ ($\hat{\mathbf{a}}_1 \neq \hat{\mathbf{a}}_2$). Ekkor a vektorterek tulajdonságait használva kapjuk, hogy

$$\hat{\mathbf{a}}_1 = \hat{\mathbf{a}}_1 + \mathbf{o} = \hat{\mathbf{a}}_1 + (\mathbf{a} + \hat{\mathbf{a}}_2) = (\hat{\mathbf{a}}_1 + \mathbf{a}) + \hat{\mathbf{a}}_2 = (\mathbf{a} + \hat{\mathbf{a}}_1) + \hat{\mathbf{a}}_2 = \mathbf{o} + \hat{\mathbf{a}}_2 = \hat{\mathbf{a}}_2.$$

Ez ellentmond az indirekt feltevésnek.

6.1.4.
A vektortér tulajdonságait felhasználva kapjuk, hogy $\mathbf{a} + 0\mathbf{a} = 1\mathbf{a} + 0\mathbf{a} = (1+0)\mathbf{a} = 1\mathbf{a} = \mathbf{a}$. Adjuk hozzá az egyenlőség mindkét oldalához az $\mathbf{a}$ elem $\hat{\mathbf{a}}$ inverzét. Az összeadás kommutativitásából adódik, hogy $0\mathbf{a} = \hat{\mathbf{a}} + \mathbf{a} + 0\mathbf{a} = \hat{\mathbf{a}} + \mathbf{a} = \mathbf{o}$. Ezt akartuk megmutatni.

6.1.5.
Az kell megmutatni, hogy $\mathbf{a} + ((-1)\mathbf{a}) = \mathbf{o}$. A vektortér tulajdonságait felhasználva kapjuk, hogy $\mathbf{a} + (-1\mathbf{a}) = 1\mathbf{a} + (-1\mathbf{a}) = (1 + (-1))\mathbf{a} = 0\mathbf{a} = \mathbf{o}$.

6.1.6.
Válasszuk ki a vektortér egy tetszőleges $\mathbf{a}$ elemét. A vektortér tulajdonságait és a 6.1.5. feladat eredményét felhasználva kapjuk, hogy

$$\alpha\mathbf{o} = \alpha(\mathbf{a} + \hat{\mathbf{a}}) = \alpha\mathbf{a} + \alpha\hat{\mathbf{a}} = \alpha\mathbf{a} + \alpha((-1)\mathbf{a}) = \alpha\mathbf{a} + (-1)(\alpha\mathbf{a}) = \mathbf{o}.$$

6.1.7.
Azért nem lesz vektortér mert pl. $n = 2$ esetén az x^2 és $-x^2+x$ polinomok hozzátartoznak a halmazhoz, de az összegük $((x^2)+(-x^2+x) = x)$ nem. Ez már elsőfokú és nem másodfokú polinom.

6.1.8.
A nullelem a $[0, 1]$ intervallumon értelmezett konstans nulla függvény. Az $x \mapsto x(1 - x)$ függvény additív inverze pedig az $x \mapsto x(x - 1)$ függvény. Ugyanis tetszőleges $x \in [0, 1]$ mellett $x(1 - x) + x(x - 1) = 0$, azaz a két függvény összege éppen a vektortér nulleleme. Megjegyzendő még, hogy a nullelemnek és az additív inverznek választott két függvény, mivel folytonosak, valóban eleme a V vektortérnek.

6.1.9.
Ahhoz, hogy a megadott halmaz vektortér legyen zártnak kell lennie az összeadásra és a számmal szorzásra. Az összeadásra zárt is, de a számmal szorzásra nem. Pl. a csupa egyest tartalmazó vektor (-1)-szerese nincs benne a halmazban.

6.2. Altér, lineáris kombináció, generátorrendszer

Egy vektortér azon részhalmazát, amely önmagában is vektortér, az eredeti vektortér alterének hívjuk. Ha egy vektortér néhány elemét egy-egy valós számmal megszorozzuk és az eredményeket összeadjuk, akkor ismét egy vektortérbeli elemet kapunk. Ezt az eljárást a vektorok lineáris kombinációjának nevezzük. Most a lineáris kombinációkkal, az alterekkel és a köztük lévő kapcsolattal ismerkedünk meg részletesebben.

A legfeljebb harmadfokú és a legfeljebb másodfokú polinomok is vektorteret alkotnak, ugyanakkor mint halmazok az elsőnek részhalmaza a második. Hasonló igaz a térbeli és síkbeli helyvektorok halmazára is.

6.2.1. Definíció. Egy V vektortér azon részhalmazát, amely maga is vektortér a V-beli műveletekre nézve V egy **alterének** hívjuk. $\diamondsuit$

Minden vektortérnek van két ún. triviális altere. Az egyik saját maga, a másik pedig az a halmaz, amely csak a nullvektort tartalmazza (lásd 6.2.3. feladat). A következő tétel szükséges és elégséges feltételt fogalmaz meg arra, hogy egy vektortér egy részhalmaza mikor lesz altér.

6.2.2. Tétel. *Egy V vektortér W nemüres részhalmaza akkor és csak akkor lesz altere V-nek, ha W zárt a V-beli műveletekre nézve.*

Bizonyítás. Ha W altér, akkor nyilvánvalóan zártnak kell lennie a V-beli műveletekre. A másik irány bizonyításához tegyük fel, hogy W zárt a V-beli műveletekre nézve. Ekkor az 1.,2.,5.,6.,7., és 8. vektortér tulajdonság automatikusan teljesül. Azt kell már csak igazolnunk, hogy W-ben van nullelem, és minden elemnek van W-beli additív inverze. Legyen $\mathbf{a}$ W-nek egy tetszőleges eleme (ilyen van, hiszen W nem az üres halmaz). Ekkor $\mathbf{a}$ -1-szeresének is W-ben kell lennie, hiszen W zárt a számmal szorzásra. Tehát $-\mathbf{a} \in W$. Ebből látható, hogy minden elemnek lesz additív inverze, valamint mivel az összeadásra is zárt W, a nullelemnek is W-ben kell lennie ugyanis $\mathbf{a} + (-\mathbf{a}) = \mathbf{a} - \mathbf{a} = \mathbf{o}$. ■

6.2.1. Feladat. Adjuk meg néhány alterét a korábban megismert vektortereknek.

6.2.2. Feladat. Miért nem lesz $\mathbb{R}^2$-nek altere az a halmaz, amely az

$$\begin{bmatrix} 1 \\ \gamma \end{bmatrix}$$

alakú vektorokat tartalmazza (γ tetszőleges valós szám)?

6.2.3. Feladat. Legyen V egy tetszőleges vektortér. Igazoljuk, hogy az $\{\mathbf{o}\}$ halmaz vektortér a V-beli műveletekre nézve.

A továbbiakban azzal foglalkozunk, hogy hogyan lehet egy vektortér altereit előállítani.

6.2.3. Definíció. Legyenek az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok a V vektortér elemei és legyenek $\alpha_1, \ldots, \alpha_n$ tetszőleges valós számok. Ekkor az $\alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n \in V$ elemet az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok $\alpha_1, \ldots, \alpha_n$ számokkal alkotott **lineáris kombinációjának** nevezzük. $\diamondsuit$

Megjegyezzük, hogy mivel a számmal való szorzás és az összeadás is elvégezhető egy vektortér elemeivel, ezért az $\alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n$ lineáris kombináció valóban eleme lesz a V vektortérnek. A lineáris kombinációk kulcsszerepet játszanak a vektorterek elméletében. Az összeadás is lineáris kombináció ($\mathbf{a} + \mathbf{b}$ az $\mathbf{a}$ és $\mathbf{b}$ vektoroknak az 1, 1 számokkal alkotott lineáris kombinációja) és a skaláris szorzás is ($\alpha\mathbf{a}$ az $\mathbf{a}$ vektornak az α számmal alkotott lineáris kombinációja). Két lineáris kombináció egyenlő, ha ugyanazok a vektorok ugyanazon valós számokkal szorzódnak bennük.

6.2.4. Feladat. Határozzuk meg a legfeljebb másodfokú polinomok vektorterében az x^2, x, $x-2$ és 2 polinomoknak a 2, 0, 1 és -2 számokkal alkotott lineáris kombinációját.

6.2.5. Feladat. Határozzuk meg $\mathbb{R}^3$-ban az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2$ és $\overline{\mathbf{a}}_3$ vektorok 1,-2 és 3 számokkal alkotott lineáris kombinációját.

$$\overline{\mathbf{a}}_1 = \begin{bmatrix} 5 \\ 2 \\ 0 \end{bmatrix}, \quad \overline{\mathbf{a}}_2 = \begin{bmatrix} -1 \\ -2 \\ 3 \end{bmatrix}, \quad \overline{\mathbf{a}}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

6.2.6. Feladat. Írjuk fel a

$$\left.\begin{array}{rcrcr} 3x & + & 2y & = & 1 \\ x & - & y & = & 2 \end{array}\right\}$$

egyenletrendszert $\mathbb{R}^2$-beli vektorok lineáris kombinációjaként.

6.2.7. Feladat. Egyenlőek-e a következő lineáris kombinációk?

$$\begin{array}{ll} a) & 3\mathbf{a} + 5\mathbf{o} + 2\mathbf{b}, \quad 3\mathbf{a} + 6\mathbf{o} + 2\mathbf{b} \\ b) & 2\mathbf{a} + 3\mathbf{b} + 3\mathbf{a}, \quad 3\mathbf{a} + 2\mathbf{b} + 2\mathbf{a} \end{array}$$

Az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorok összes lehetséges lineáris kombinációjának halmazát a továbbiakban $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ fogja jelölni (L a lineáris kombináció kifejezésre utal).

6.2.4. Tétel. *Rögzített $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorok esetén az $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ halmaz altere V-nek.*

Bizonyítás. Az nyilvánvaló, hogy $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ része lesz V-nek, hiszen az zárt az összeadásra és a számmal szorzásra nézve. A 6.2.2. tétel miatt már csak azt kell igazolni, hogy $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ szintén zárt a fenti két műveletre. Legyen $\mathbf{a}$ és $\mathbf{b}$ két $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$-beli vektor. Ekkor léteznek olyan $\alpha_1, \ldots, \alpha_n$ valamint $\beta_1, \ldots, \beta_n$ konstansok, melyekkel az $\mathbf{a}$ és $\mathbf{b}$ vektorok felírhatók

$$\mathbf{a} = \alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n, \quad \mathbf{b} = \beta_1\mathbf{a}_1 + \ldots + \beta_n\mathbf{a}_n$$

alakban. Az $\mathbf{a}$ és $\mathbf{b}$ vektorok összege a vektortérbeli műveleti szabályokat alkalmazva

$$\mathbf{a} + \mathbf{b} = (\alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n) + (\beta_1\mathbf{a}_1 + \ldots + \beta_n\mathbf{a}_n) = (\alpha_1 + \beta_1)\mathbf{a}_1 + \ldots + (\alpha_n + \beta_n)\mathbf{a}_n,$$

az $\mathbf{a}$ vektornak egy tetszőleges λ valós számmal való szorzata pedig

$$\lambda\mathbf{a} = \lambda(\alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n) = (\lambda\alpha_1)\mathbf{a}_1 + \ldots + (\lambda\alpha_n)\mathbf{a}_n.$$

Ez azt jelenti, hogy $\mathbf{a} + \mathbf{b}$ és $\lambda\mathbf{a}$ is eleme lesz $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$-nek, hiszen előállnak az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok lineáris kombinációjaként. Ezt akartuk megmutatni. ■

Az $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ halmaz tehát altere V-nek, ezért azt szokás az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok által generált altérnek is hívni. Emiatt azt az eljárást, amikor képezzük az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok összes lineáris kombinációját altérgenerálásnak vagy vektortérgenerálásnak is hívjuk.

6.2.5. Definíció. Ha az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorokra $L(\mathbf{a}_1, \ldots, \mathbf{a}_n) = V$, akkor az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorokat a V vektortér egy **generátorrendszerének** nevezzük. ♢

Egy generátorrendszerben definíció szerint tehát csak véges sok vektor szerepelhet. Ha ezt mégis ki szeretnénk hangsúlyozni, akkor a generátorrendszer elnevezés helyett a véges generátorrendszer kifejezést fogjuk használni. Ha egy vektortérnek van generátorrendszere, akkor a generátorrendszer vektorainak lineáris kombinációjaként a vektortér minden eleme előállítható.

6.2.8. Feladat. Mutassuk meg, hogy az $1 + x$, x^2 és $2 + x$ polinomok generátorrendszert alkotnak a legfeljebb másodfokú polinomok vektorterében, majd határozzuk meg, hogy az $x^2 + x + 1$ polinom hogyan írható fel a fenti három polinom lineáris kombinációjaként.

6.2.9. Feladat. Generátorrendszert alkotnak-e az

$$\overline{\mathbf{a}} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \overline{\mathbf{b}} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}, \quad \overline{\mathbf{c}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

vektorok $\mathbb{R}^3$-ban?

6.2.10. Feladat. Igazoljuk, hogy ha egy generátorrendszernek az $\mathbf{a}$ és $-\mathbf{a}$ vektor is tagja, akkor a vektortér minden eleme végtelen sokféleképpen áll elő a generátorrendszer lineáris kombinációjaként.

6.2.11. Feladat. Igazoljuk, hogy az összes polinom vektorterének nincsen véges generátorrendszere.

Ellenőrző kérdések

1. Mit nevezünk altérnek?
2. Mi a szükséges és elégséges feltétele annak, hogy egy vektortér részhalmaza altere legyen a vektortérnek?
3. Mondjunk példákat a nevezetes vektorterek esetén alterekre.
4. Képezzük néhány vektor összes lineáris kombinációjának halmazát! Hogyan jelöljük és milyen tulajdonságokkal rendelkezik ez a halmaz?
5. Mit hívunk generátorrendszernek?

A feladatok megoldásai

6.2.1.
Csak a nem triviális altereket tekintve a síkbeli helyvektorok terében altér pl. azon vektorok halmaza melyek végpontjai egy az origón átmenő egyenesre illeszkednek. A térbeli helyvektorok esetén ugyanígy lehet alteret kapni. Altér még továbbá azon vektorok halmaza is, melyek végpontjai egy az origón átmenő síkra illeszkednek. A legfeljebb ötödfokú polinomok vektorterében altér a legfeljebb elsőfokú vagy pl. a legfeljebb harmadfokú polinomok halmaza. Az $f : [0, 1] \to \mathbb{R}$ típusú függvények vektorterében alteret alkotnak az $f : [0, 1] \to \mathbb{R}$ típusú folytonos függvények és ebben egy újabb altér lesz a differenciálható függvények halmaza. Az $m \times n$-es mátrixok terében alteret alkotnak pl. azok a mátrixok, melyeknek minden eleme ugyanaz a szám, vagy amelyeknek minden oszlopa egyforma.

6.2.2.
Ahhoz, hogy a megadott halmaz altér legyen két dolognak kellene teljesülnie. Egyrészt $\mathbb{R}^2$ részhalmazának kellene lennie (ez teljesül is). Másrészt zártnak kellene lennie az $\mathbb{R}^2$-beli műveletekre. Ez viszont nem teljesül.

6.2.3.
A 6.2.2. tétel miatt csak azt kell megmutatni, hogy $\{\mathbf{o}\}$ zárt a műveletekre nézve. Ez abból következik, hogy $\mathbf{o} + \mathbf{o} = \mathbf{o} \in \{\mathbf{o}\}$ és $\alpha\mathbf{o} = \mathbf{o} \in \{\mathbf{o}\}$ (α tetszőleges valós szám).

6.2.4.
A lineáris kombinációként adódó polinom a $2(x^2) + 0x + 1(x - 2) + (-2)(2) = 2x^2 + x - 6$.

6.2.5.
Az adott lineáris kombináció a

$$\begin{bmatrix} 10 \\ 9 \\ -3 \end{bmatrix}$$

vektor lesz.

6.2.6.
A keresett lineáris kombináció

$$x \begin{bmatrix} 3 \\ 2 \end{bmatrix} + y \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Vegyük észre, hogy egy egyenletrendszer megoldása során tulajdonképpen az ismeretlenekkel mint együtthatókkal képezett lineáris kombinációját állítjuk elő adott vektoroknak úgy, hogy az a jobb oldalon álló vektort adja.

6.2.7.
Az első két lineáris kombináció nem egyenlő, mert az $\mathbf{o}$ vektor nem ugyanazzal a számmal szorzódik a két lineáris kombinációban. (Vegyük észre, hogy a két lineáris kombináció annak ellenére különböző, hogy eredményül mindkettő ugyanazt a $3\mathbf{a} + 2\mathbf{b}$ vektort adja.) A második kettő egyenlő, hiszen ugyanazok a vektorok ugyanazon számokkal szorzódnak, csak más sorrendben.

6.2.8.
Azt kell igazolni, hogy az összes legfeljebb másodfokú polinom előáll a megadott polinomok lineáris kombinációjaként. Legyen pl. $ax^2 + bx + c$ egy adott polinom ($a, b, c \in \mathbb{R}$ tetszőleges). Tegyük fel, hogy ez előáll az $\alpha(1+x)+\beta(x^2)+\gamma(2+x)$ alakban. A polinomot

egyszerűbb alakra hozva kapjuk a $\beta x^2 + (\alpha + \gamma)x + (\alpha + 2\gamma)$ polinomot. Ahhoz, hogy ez egyenlő lehessen az $ax^2 + bx + c$ polinommal teljesülnie kell a következő egyenlőségeknek: $\beta = a$, $\alpha + \gamma = b$ és $\alpha + 2\gamma = c$. Az utolsó kettőből kapjuk, hogy $\alpha = 2b - c$ és $\gamma = c - b$. Könnyű végiggondolni, hogy α, β és γ előbbi választása esetén a korábbi lineáris kombináció valóban előállítja az $ax^2 + bx + c$ polinomot. Így az adott vektorrendszer generátorrendszer. (Ráadásul minden polinom csak egyféleképpen áll elő.) Az $x^2 + x + 1$ polinom felírása tehát $x^2 + x + 1 = 1(1 + x) + 1(x^2) + 0(2 + x)$.

6.2.9.
Az adott vektorok α, β és γ számokkal alkotott lineáris kombinációja a

$$\begin{bmatrix} 2\alpha + 4\beta \\ 1\alpha + 2\beta \\ \gamma \end{bmatrix}$$

vektor. Mivel a vektor második eleme első elemének kétszerese, emiatt csak az ilyen típusú vektorok állíthatók elő a megadott vektorokkal. Az adott vektorrendszer tehát nem generátorrendszer.

6.2.10.
Legyen $\mathbf{b}$ a vektortér egy tetszőleges eleme. Jelöljük α_1-gyel ill. α_2-vel, hogy $\mathbf{b}$ előállításában milyen együtthatókkal szerepelnek az $\mathbf{a}$ ill. $-\mathbf{a}$ vektorok. Ekkor tetszőleges $\alpha \in \mathbb{R}$ paraméter mellett, a lineáris kombináció $\alpha_1\mathbf{a} + \alpha_2(-\mathbf{a})$ vektorát kicserélhetjük a vele egyenlő $(\alpha_1 + \alpha)\mathbf{a} + (\alpha_2 + \alpha)(-\mathbf{a})$ vektorra. Ebből már következik, hogy végtelen sok előállítás lehetséges.

6.2.11.
Tegyük fel, hogy van véges generátorrendszer. Vegyük a generátorrendszer legmagasabb fokszámú polinomját. Legyen ennek fokszáma n. Mivel az összeadás és számmal szorzás során a polinomok fokszáma nem növekedhet, ezért az adott polinomokkal nem állítható elő n-nél nagyobb fokszámú polinom. Így a vektorok nem alkothatnak generátorrendszert.

6.3. Lineáris függetlenség, összefüggőség

Az előző fejezetben megismerkedtünk a vektorok által generált alterekkel. Ebben a fejezetben bevezetjük a független ill. összefüggő vektorok fogalmát, melyek segítségével azt tárgyaljuk, hogy a generált altér egyes elemei hányféleképpen állíthatók elő a generáló vektorokkal.

6.3.1. Definíció. Az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorokat **(lineárisan) független vektoroknak** nevezzük, ha lineáris kombinációjukkal csak úgy lehet előállítani az $\mathbf{o}$ vektort, ha minden vektor szorzója zérus. Ellenkező esetben a vektorokat **(lineárisan) összefüggőknek** hívjuk. $\diamondsuit$

Az egyszerűbb fogalmazás kedvéért azt, hogy az $\mathbf{o}$ vektor csak úgy áll elő vektorok lineáris kombinációjaként, hogy mindegyik vektort nullával szorozzuk, a továbbiakban úgy fogjuk kifejezni, hogy az $\mathbf{o}$ vektor csak triviálisan áll elő a vektorok lineáris kombinációjaként.

6.3.1. Feladat. Döntsük el, hogy függetlenek-e az x^2, $x+1$, x és 2 elemek a legfeljebb másodfokú polinomok vektorterében.

6.3.2. Feladat. Döntsük el, hogy függetlenek-e az

$$\mathbf{a} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix}$$

vektorok $\mathbb{R}^3$-ban.

6.3.3. Feladat. Igazoljuk, hogy ha $\mathbf{a}$ és $\mathbf{b}$ független vektorok egy vektortérben, akkor az $\mathbf{a}+\mathbf{b}$ és $\mathbf{a}-\mathbf{b}$ vektorok is függetlenek.

Most szükséges és elégséges feltételét adjuk meg annak, hogy egy vektorrendszer összefüggő ill. független legyen.

6.3.2. Tétel. *Legyenek adottak az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorok. Ezek akkor és csak akkor lineárisan összefüggők, ha valamelyik vektor kifejezhető a többi vektor lineáris kombinációjaként.*

Bizonyítás. Tegyük fel először, hogy az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorok lineárisan összefüggőek. Az összefüggőség definíciója szerint tehát a nullvektor felírható

$$\mathbf{o} = \alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n$$

alakban, ahol nem mindegyik együttható nulla. Legyen pl. $\alpha_i \neq 0$. Ekkor $\mathbf{a}_i$ a következő módon írható fel

$$\mathbf{a}_i = -\frac{\alpha_1}{\alpha_i}\mathbf{a}_1 - \ldots - \frac{\alpha_{i-1}}{\alpha_i}\mathbf{a}_{i-1} - \frac{\alpha_{i+1}}{\alpha_i}\mathbf{a}_{i+1} - \ldots - \frac{\alpha_n}{\alpha_i}\mathbf{a}_n.$$

Ezzel a tételt az egyik irányban bebizonyítottuk.

Most tegyük fel, hogy a vektorok közül valamelyik felírható a többi lineáris kombinációjaként. Mindenféle megszorítás nélkül feltehetjük, hogy ez az $\mathbf{a}_1$ vektor. Tehát léteznek olyan $\alpha_2, \ldots, \alpha_n$ valós számok, melyekre

$$\mathbf{a}_1 = \alpha_2\mathbf{a}_2 + \ldots + \alpha_n\mathbf{a}_n.$$

Mindkét oldalból kivonva az $\mathbf{a}_1$ vektort kapjuk, hogy

$$\mathbf{o} = -\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \ldots + \alpha_n\mathbf{a}_n.$$

Mivel az $\mathbf{o}$ vektor nem csak triviálisan állítható elő az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorokkal ezért a vektorok lineárisan összefüggők. ■

A fenti tételből következik, hogy az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorok akkor és csak akkor lineárisan függetlenek, ha egyik sem fejezhető ki a többiek lineáris kombinációjaként.

6.3.4. Feladat. Igazoljuk, hogy ha az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorok között szerepel az $\mathbf{o}$ vektor, akkor ezek a vektorok lineárisan összefüggők.

6.3.5. Feladat. Igazoljuk, hogy ha az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorok között szerepel egy vektor és ennek a vektornak egy számszorosa is, akkor a vektorok lineárisan összefüggők.

6.3.6. Feladat. Igazoljuk, hogy ha az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ független vektorok közül elveszünk néhányat, akkor a megmaradó vektorok is függetlenek lesznek.

6.3.7. Feladat. Igazoljuk, hogy ha az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ összefüggő vektorokat kiegészítjük még néhány vektorral, akkor az így kapott vektorok is összefüggőek lesznek.

Az előző tétel azzal foglalkozott, hogy független ill. összefüggő vektorok hogyan fejezhetők ki egymással. Most azt fogjuk megvizsgálni, hogy ezekkel a vektorokkal hogyan fejezhetők ki az általuk generált tér elemei.

6.3.3. Tétel. *Legyenek adottak az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorok. Ezek akkor és csak akkor lineárisan függetlenek ha $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ minden eleme csak egyféleképpen állítható elő lineáris kombinációjukként.*

Bizonyítás. Tegyük fel először indirekt, hogy a $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok függetlenek és van olyan $\mathbf{a}$ vektor $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$-ben, amely kétféleképpen is előáll lineáris kombinációként. Legyen ez a két lineáris kombináció

$$\begin{aligned} \mathbf{a} &= \alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n, \\ \mathbf{a} &= \beta_1\mathbf{a}_1 + \ldots + \beta_n\mathbf{a}_n, \end{aligned}$$

ahol valamelyik α_i eltér β_i-től. Vonjuk ki egymásból a két egyenlőséget.

$$\mathbf{o} = (\alpha_1 - \beta_1)\mathbf{a}_1 + \ldots + (\alpha_n - \beta_n)\mathbf{a}_n.$$

Mivel az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok függetlenek, ezért az $\mathbf{o}$ csak triviálisan állhat elő. Tehát az $\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n$ egyenlőségeknek teljesülniük kell, ami ellentmond az indirekt feltételnek. Ezzel a tételt az egyik irányban bebizonyítottuk.

A másik irányhoz azt kell csak észrevenni, hogy ha minden vektor csak egyféleképpen áll elő lineáris kombinációként, akkor az $\mathbf{o}$ vektor is. Ez természetesen nem lehet más, mint a triviális előállítás, azaz $\mathbf{o}$ csak triviálisan áll elő lineáris kombinációként. Tehát a vektorok függetlenek. ■

6.3.4. Tétel. *Legyenek adottak az* $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ *vektorok. Ezek akkor és csak akkor lineárisan összefüggőek, ha* $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ *minden eleme végtelen sokféleképpen állítható elő lineáris kombinációjukként.*

Bizonyítás. Legyenek az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok összefüggőek. Ekkor kifejezhető velük az $\mathbf{o}$ vektor $\mathbf{o} = \beta_1\mathbf{a}_1 + \ldots + \beta_n\mathbf{a}_n$ alakban úgy, hogy nem minden β_i $(i = 1, \ldots, n)$ együttható nulla. Legyen $\mathbf{a}$ $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$-nek egy tetszőleges vektora, amely kifejezhető $\mathbf{a} = \alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n$ alakban. Azt kell megmutatnunk, hogy $\mathbf{a}$-nak végtelen sok más előállítása is lehetséges. Tetszőleges β valós számra igaz, hogy

$$\mathbf{a} = \mathbf{a} + \beta\mathbf{o} = \alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n + \beta(\beta_1\mathbf{a}_1 + \ldots + \beta_n\mathbf{a}_n) =$$

$$= (\alpha_1 + \beta\beta_1)\mathbf{a}_1 + \ldots + (\alpha_n + \beta\beta_n)\mathbf{a}_n.$$

Mivel nem mindegyik β_i együttható nulla, ezért ez az előállítás minden β-ra más lesz. Így $\mathbf{a}$-nak tényleg végtelen sok előállítása lehetséges. ■

A fenti tételekből tehát látszik, hogy vagy az az eset áll fenn, hogy az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorokkal minden $L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$-beli vektor végtelen sokféleképpen írható fel, vagy az, hogy mindegyik egyértelműen írható fel. Nem fordulhat tehát elő az az eset, hogy bizonyos vektorok csak egyféleképpen, míg mások végtelen sokféleképpen írhatók fel. Az sem fordulhat elő hogy egy vektor csak véges sokféleképpen álljon elő lineáris kombinációként. Ez az észrevétel fontos lesz majd a lineáris egyenletrendszerek tárgyalásánál.

Ellenőrző kérdések

1. Mikor nevezünk vektorokat lineárisan független vektoroknak?

2. Mikor nevezünk vektorokat lineárisan összefüggő vektoroknak?

3. Mi a szükséges és elégséges feltétele annak, hogy vektorok lineárisan függetlenek ill. összefüggők legyenek?

4. Hányféleképpen fejezhetők ki lineáris kombinációként a független ill. az összefüggő vektorok által generált altér elemei?

A feladatok megoldásai

6.3.1.
A polinomok nem függetlenek, mert a nullelem – a nulla polinom – nem csak triviálisan állítható elő. Pl. $0 = 0(x^2) + 2(x + 1) - 2(x) - 1(2)$.

6.3.2.
Az adott vektorokkal csak úgy lehet a nullvektort előállítani, hogy az $\overline{\mathbf{a}}$ vektort nullával szorozzuk, uis. különben a lineáris kombinációként adódó vektor első eleme nem lenne nulla. Hasonlóan a $\overline{\mathbf{b}}$ vektort is nullával kell ezután szorozni, és végül $\overline{\mathbf{c}}$-t is. Tehát a nullvektor csak triviálisan állt elő, azaz a vektorok lineárisan függetlenek.

6.3.3.
Állítsuk elő az $\mathbf{o}$ vektort az $\mathbf{a} + \mathbf{b}$ és $\mathbf{a} - \mathbf{b}$ vektorok lineáris kombinációjaként

$$\mathbf{o} = \alpha(\mathbf{a} + \mathbf{b}) + \beta(\mathbf{a} - \mathbf{b})$$

alakban. Az egyenlőséget átrendezve kapjuk, hogy $\mathbf{o} = (\alpha + \beta)\mathbf{a} + (\alpha - \beta)\mathbf{b}$. Mivel $\mathbf{a}$ és $\mathbf{b}$ függetlenek, ezért $\alpha + \beta$ és $\alpha - \beta$ is nulla kell legyen. Ez csak úgy lehet, ha α is és β is nulla, tehát az $\mathbf{o}$ csak triviálisan áll elő az $\mathbf{a} + \mathbf{b}$ és $\mathbf{a} - \mathbf{b}$ vektorokkal.

6.3.4.
A 6.3.2. tétel szerint elegendő megmutatni, hogy az $\mathbf{o}$ vektor felírható a többi vektor lineáris kombinációjaként. Ez pedig teljesül, ha mindegyik vektort nullával szorozzuk.

6.3.5.
Legyen az említett vektor $\mathbf{a}$, számszorosa pedig $\alpha\mathbf{a}$ ($\alpha \in \mathbb{R}$). Elég megmutatni, hogy a vektorok közül valamelyik kifejezhető a többivel. A kifejezhető vektor az $\alpha\mathbf{a}$ lesz. Ehhez $\mathbf{a}$-t α-val kell szorozni, a többi vektort pedig nullával.

6.3.6.
A vektorok akkor és csak akkor függetlenek, ha egyik sem fejezhető ki a többi lineáris kombinációjaként. Ha elveszünk néhányat a vektorok közül, akkor ez továbbra is érvényes lesz, tehát a vektorok függetlenek maradnak.

6.3.7.
A vektorok akkor és csak akkor összefüggők, ha az egyik kifejezhető a többi lineáris kombinációjaként. Ha hozzáveszünk még ezekhez a vektorokhoz újabb vektorokat, akkor ez továbbra is érvényes lesz, tehát a vektorok összefüggőek maradnak.

6.4. Bázis, dimenzió

Ha egy vektortérben van véges generátorrendszer, akkor a vektortér minden eleme előáll a generátorrendszer elemeinek lineáris kombinációjaként. Ebben a fejezetben megkeressük a legkevesebb elemű generátorrendszereket, az ún. bázisokat, valamint definiálni fogjuk a vektorterek dimenzióját.

Ahogy láttuk, egy generátorrendszer elemeivel a vektortér minden eleme felírható lineáris kombinációként, azaz egy többnyire végtelen sok elemből álló vektortér generálható véges sok eleme segítségével. Célunk a legkevesebb elemű generátorrendszerek megkeresése egy adott vektortérben.

A fenti probléma felvetésének nyilván csak olyan vektorterek esetében van értelme, melyeknek van (véges) generátorrendszere. Az ilyen vektortereket véges dimenziós vektortereknek hívjuk. Ellenkező esetben végtelen dimenziós vektortérről beszélünk. Az összes polinom vektortere (lásd 6.2.11. feladat) vagy pl. az $f : [a, b] \to \mathbb{R}$ típusú függvények vektortere nem generálható véges sok vektora segítségével, tehát ezek végtelen dimenziós vektorterek. Ebben a jegyzetben csak a véges dimenziós vektorterekkel foglalkozunk részletesebben.

Most azt vizsgáljuk, hogy mely generátorrendszerekből lehet ill. nem lehet úgy vektorokat elhagyni, hogy a maradék vektorrendszer is generátorrendszer maradjon.

6.4.1. Tétel. *Legyen $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ a V véges dimenziós vektortér egy generátorrendszere. Akkor és csak akkor hagyhatunk el egy vektort ebből a generátorrendszerből úgy, hogy a maradék vektorrendszer még mindig generátorrendszer maradjon, ha a generátorrendszer lineárisan összefüggő vektorokat tartalmaz.*

Bizonyítás. Tegyük fel, hogy a vektorok lineárisan összefüggőek. Ebben az esetben az egyik vektor kifejezhető a többi lineáris kombinációjaként (lásd 6.3.2. tétel). Legyen ez mondjuk az $\mathbf{a}_1$ vektor. Ekkor a V vektortér generálásához elegendőek az $\mathbf{a}_2, \ldots, \mathbf{a}_n$ vektorok, hiszen a vektortér összes elemének előállításában $\mathbf{a}_1$ mindig helyettesíthető az $\mathbf{a}_2, \ldots, \mathbf{a}_n$ vektorok lineáris kombinációjával. Tehát az $\mathbf{a}_2, \ldots, \mathbf{a}_n$ vektorok is generátorrendszert alkotnak. Ezzel a tételt az egyik irányban igazoltuk.

Most tegyük fel indirekt, hogy a vektorok lineárisan függetlenek és mondjuk az $\mathbf{a}_1$ vektort még el lehet hagyni úgy, hogy a vektorrendszer generátorrendszer maradjon. Mivel a vektorok lineárisan függetlenek, így minden vektortérbeli elem csak egyféleképpen írható fel lineáris kombinációjukkal (lásd 6.3.3. tétel). Emiatt az $\mathbf{a}_1$ vektor egyetlen felírása az lenne, hogy $\mathbf{a}_1$-et eggyel, a többi vektort pedig nullával szorozzuk. Ha $\mathbf{a}_1$-et elhagynánk tehát a generátorrendszerből, akkor az $\mathbf{a}_1$ vektor már nem lenne előállítható, ami azt jelentené, hogy az $\mathbf{a}_1$ elhagyásával keletkezett vektorrendszer már nem generátorrendszer. ■

A fenti tétel miatt tehát ha egy generátorrendszer lineárisan összefüggő vektorokat tartalmaz, akkor mindaddig csökkenthetjük az elemszámát egy-egy megfelelő vektor elhagyásával, míg a maradék vektorrendszer lineárisan független nem lesz. A lineárisan független elemeket tartalmazó generátrorendszerek tartalmazzák a legkevesebb vektort abban az értelemben, hogy nem hagyhatunk el belőlük vektort úgy, hogy a maradék vektorrendszer még mindig a teljes V vektorteret generálja.

6.4.2. Definíció. Az $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ vektorokat a V véges dimenziós vektortér egy **bázisának** hívjuk, ha generátorrendszert alkotnak és lineárisan függetlenek. ♢

Ahhoz tehát, hogy egy vektorrendszerről eldöntsük, hogy bázis, a vektorok függetlenségét kell ellenőrizni és azt, hogy a vektorok generálják-e az egész vektorteret. Mivel egy véges dimenziós vektortér minden generátorrendszeréből tudunk olyan generátorrendszert csinálni, amely lineárisan független elemeket tartalmaz ezért nyilvánvaló, hogy minden véges dimenziós vektortérben van bázis.

6.4.1. Feladat. Mondjunk néhány egyszerű bázist a sík- és térbeli helyvektorok és a legfeljebb másodfokú polinomok vektorterében, $\mathbb{R}^2$-ben és $\mathbb{R}^{2\times 2}$-ben.

6.4.2. Feladat. Bázist alkotnak-e $\mathbb{R}^2$-ben az

$$\overline{\mathbf{a}}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \overline{\mathbf{a}}_2 = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$$

vektorok?

Most arra keressük a választ, hogy mit mondhatunk egy vektortér bázisainak elemszámáról. Először bizonyítjuk az ún. kicserélési tételt.

6.4.3. Tétel. *(Kicserélési tétel.) Legyen $\mathbf{a}_1, \ldots, \mathbf{a}_n$ a V véges dimenziós vektortér egy bázisa és $\mathbf{b}_1, \ldots, \mathbf{b}_m$ lineárisan független vektorok V-ben. Ekkor a bázisnak létezik m olyan eleme melyeket a $\mathbf{b}_1, \ldots, \mathbf{b}_m$ vektorokra kicserélve a keletkezett vektorrendszer továbbra is bázis marad.*

Bizonyítás. Mivel az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok bázist alkotnak (ami generátorrendszer is), ezért $\mathbf{b}_1$ felírható valamilyen $\alpha_1^{(1)}, \ldots, \alpha_n^{(1)}$ számokkal

$$\mathbf{b}_1 = \alpha_1^{(1)}\mathbf{a}_1 + \ldots + \alpha_i^{(1)}\mathbf{a}_i + \ldots + \alpha_n^{(1)}\mathbf{a}_n \tag{6.4.1}$$

alakban (az $^{(1)}$ felső index azt fejezi ki, hogy $\mathbf{b}_1$ előállításáról van szó). Mivel $\mathbf{b}_1$ biztosan különbözik az $\mathbf{o}$ vektortól (az $\mathbf{o}$ vektor nem szerepelhet lineárisan független vektorok között), azért van olyan $\alpha_i^{(1)}$ együttható amely biztosan nem nulla. Most megmutatjuk, hogy $\mathbf{a}_i$ kicserélhető $\mathbf{b}_1$-re, azaz az

$$\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{b}_1, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n$$

vektorok is bázist alkotnak. Először megmutatjuk, hogy generátorrendszert alkotnak. A (6.4.1) egyenlőségből fejezzük ki az $\mathbf{a}_i$ vektort.

$$\mathbf{a}_i = \frac{1}{\alpha_i^{(1)}}\left(\mathbf{b}_1 - \alpha_1^{(1)}\mathbf{a}_1 - \ldots - \alpha_{i-1}^{(1)}\mathbf{a}_{i-1} - \alpha_{i+1}^{(1)}\mathbf{a}_{i+1} - \ldots - \alpha_n^{(1)}\mathbf{a}_n\right).$$

Ez az egyenlőség és az a tény, hogy minden V-beli vektor előáll az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok lineáris kombinációjaként mutatja, hogy minden vektor előáll az

$$\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{b}_1, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n$$

vektorok lineáris kombinációjaként is. Most megmutatjuk, hogy a vektorok függetlenek is. Tegyük fel, hogy a vektorokkal előállítottuk az $\mathbf{o}$ vektort

$$\mathbf{o} = \lambda_1\mathbf{a}_1 + \ldots + \lambda_{i-1}\mathbf{a}_{i-1} + \beta\mathbf{b}_1 + \lambda_{i+1}\mathbf{a}_{i+1} + \ldots + \lambda_n\mathbf{a}_n$$

alakban. Ez átírható az

$$\mathbf{o} = \lambda_1\mathbf{a}_1 + \ldots + \lambda_{i-1}\mathbf{a}_{i-1} + \beta(\alpha_1^{(1)}\mathbf{a}_1 + \ldots + \alpha_i^{(1)}\mathbf{a}_i + \ldots + \alpha_n^{(1)}\mathbf{a}_n) + \lambda_{i+1}\mathbf{a}_{i+1} + \ldots + \lambda_n\mathbf{a}_n =$$

$$= (\lambda_1+\beta\alpha_1^{(1)})\mathbf{a}_1+\ldots+(\lambda_{i-1}+\beta\alpha_{i-1}^{(1)})\mathbf{a}_{i-1}+\beta\alpha_i^{(1)}\mathbf{a}_i+(\lambda_{i+1}+\beta\alpha_{i+1}^{(1)})\mathbf{a}_{i+1}+\ldots+(\lambda_n+\beta\alpha_n^{(1)})\mathbf{a}_n$$

formába, felhasználva a (6.4.1) összefüggést. Mivel az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok függetlenek, ezért a fenti egyenlőségből következik, hogy $\beta = 0$ és hogy $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_n = 0$. A vizsgált vektorok tehát valóban függetlenek. Ha tehát az $\mathbf{a}_i$ vektort kicseréljük a $\mathbf{b}_1$ vektorra, akkor továbbra is bázist kapunk.

Most tekintsük a $\mathbf{b}_2$ vektort, amely nyilván felírható

$$\mathbf{b}_2 = \alpha_1^{(2)}\mathbf{a}_1 + \ldots + \alpha_{i-1}^{(2)}\mathbf{a}_{i-1} + \mu\mathbf{b}_1 + \alpha_{i+1}^{(2)}\mathbf{a}_{i+1} + \ldots + \alpha_n^{(2)}\mathbf{a}_n \qquad (6.4.2)$$

alakban. Az itt szereplő együtthatók valamelyike biztosan különbözik nullától, hiszen $\mathbf{b}_2 \neq \mathbf{o}$. Az viszont nem lehet, hogy csak μ legyen nullától különböző, mert akkor $\mathbf{b}_2 = \mu\mathbf{b}_1$, és így $\mathbf{b}_1$ és $\mathbf{b}_2$ nem lehetne független, hiszen akkor nem lehetne az egyik a másik számszorosa. Tehát van olyan $\alpha_j^{(2)}$, amely nem nulla. Ezután, hasonlóan $\mathbf{b}_1$ becseréléséhez, igazolható, hogy az

$$\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{b}_1, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_{j-1}, \mathbf{b}_2, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n$$

vektorok bázist alkotnak. (Itt $\mathbf{a}_j$-t cseréltük ki a $\mathbf{b}_2$ vektorra.)

Ezt az eljárást folytathatjuk addig, amíg a $\mathbf{b}_1, \ldots, \mathbf{b}_m$ vektorok mind bekerülnek a bázisba. ■

A tétel egyik fontos következménye az, hogy egy véges dimenziós vektortérben egy független vektorrendszernek nem lehet több eleme, mint egy bázisnak. A másik fontos következményt tétel formájában fogalmazzuk meg.

6.4.4. Tétel. *Egy V véges dimenziós vektortér minden bázisának elemszáma ugyanannyi.*

Bizonyítás. Legyen $\mathbf{a}_1, \ldots, \mathbf{a}_n$ és $\mathbf{b}_1, \ldots, \mathbf{b}_m$ a V véges dimenziós vektortér két bázisa. Ekkor mivel az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok függetlenek a $\mathbf{b}_1, \ldots, \mathbf{b}_m$ vektorok pedig bázist alkotnak, ezért $n \leq m$. Másrészt mivel az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorok bázist alkotnak a $\mathbf{b}_1, \ldots, \mathbf{b}_m$ vektorok pedig függetlenek ezért $m \leq n$. A két egyenlőtlenségből következik tehát, hogy $n = m$. Pontosan ezt akartuk igazolni. ■

A tétel következménye, hogy egy véges dimenziós vektortérben egy generátorrendszernek nem lehet kevesebb eleme, mint egy bázisnak. Másik következmény, hogy minden véges dimenziós vektortérhez tartozni fog egy szám, amely megmondja, hogy hány elemből állnak a vektortér bázisai.

6.4.5. Definíció. Egy véges dimenziós V vektortérben a bázisok elemszámát a **vektortér dimenziójának** nevezzük. Jelölése: $\dim(V)$. ♢

E fejezet bevezetőjében már használtuk a dimenzió szót. Vegyük észre, hogy a most bevezetett dimenziófogalom pontosítja az ottanit. Nevezetesen a véges dimenzió azt jelentette, hogy van a vektortérnek véges sok elemből álló generátorrendszere. Most már látjuk, hogy ezen vektorterek bázisainak elemszáma valóban véges, így jogos a véges dimenziós vektortér elnevezés. Ugyanakkor mivel a végtelen dimenziós vektortereknek nincs véges generátorrendszerük, ezért nincs véges sok elemből álló bázisuk sem. Emiatt valóban használható a végtelen dimenziós vektortér elnevezés.

Most kiszámítjuk néhány nevezetes vektortér dimenzióját. A 6.4.1. feladatban már meghatároztuk ezen vektorterek egy-egy bázisát, így a dimenzió megállapításához csak a báziselemeket kell megszámolni.

6.4.6. Példa. Tekintsük először a síkvektorok vektorterét. Ebben bázist alkot bármely két nem egy egyenesre eső vektor. Tehát a vektortér dimenziója kettő. A térvektorok terében bázis bármely három nem egy síkra eső vektor. Ez a tér tehát háromdimenziós. Vegyük észre, hogy a kapott dimenzió értékek megfelelnek mindennapi szemléletünknek is, hiszen a sík kettő- míg a tér háromdimenziós. □

6.4.7. Példa. A legfeljebb másodfokú polinomok vektortere háromdimenziós. Általában a legfeljebb n-edfokú polinomok tere $n+1$-dimenziós. A legegyszerűbb bázist itt az $1, x, \ldots, x^n$ polinomok adják. □

6.4.8. Példa. Az $\mathbb{R}^n$ tér dimenziója n. A legegyszerűbb bázist $\mathbb{R}^n$-ben azok az $\overline{\mathbf{e}}_i$ vektorok alkotják, melyeknek az i-edik elemük egy, a többi pedig nulla ($i = 1, \ldots, n$), azaz

$$\overline{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \overline{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \ldots \quad, \overline{\mathbf{e}}_{n-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \overline{\mathbf{e}}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

□

6.4.9. Példa. Az $\mathbb{R}^{k\times n}$ tér dimenziója $k \cdot n$. A legegyszerűbb bázist $\mathbb{R}^{k\times n}$-ben azok a mátrixok alkotják, melyeknek az i-edik sorának j-edik eleme egy, a többi pedig nulla ($i = 1, \ldots, k; j = 1, \ldots, n$). □

6.4.10. Tétel. *Egy véges dimenziós vektortér bármely független vektorrendszerét kiegészíthetjük bázissá.*

Bizonyítás. Ha adott egy lineárisan független vektorrendszer, amely még nem bázis, akkor van olyan elem a vektortérben, amely nem fejezhető ki az adott vektorok lineáris kombinációjával. Vegyük ezt a vektort hozzá a vektorrendszerhez, amely nyilvánvalóan továbbra is független marad. Az eljárást addig folytatjuk, míg a vektorrendszer már tovább nem bővíthető, azaz bázis lesz. A vektortér véges dimenziója garantálja, hogy véges sok lépésen belül valóban bázishoz jutunk. ■

Most két egyszerű állítás következik, melyek segítenek eldönteni, hogy egy vektorrendszer bázis vagy sem.

6.4.11. Tétel. *Legyen V egy n-dimenziós vektortér.*

a) Az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorrendszer akkor és csak akkor bázis V-ben, ha elemei lineárisan függetlenek.

b) Az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ vektorrendszer akkor és csak akkor bázis V-ben, ha a vektorok generátorrendszert alkotnak.

Bizonyítás. A tételek egyik iránya a bázis definíciójából következik. Az a) esetben a másik irány a kicserélési tétel közvetlen következménye. A b) esethez tegyük fel indirekt,

hogy a vektorok generátorrendszert alkotnak, de nem alkotnak bázist. Ekkor el tudunk hagyni a vektorrendszerből néhány vektort úgy, hogy bázist kapjunk (lásd 6.4.1. tétel). Ennek a bázisnak n-nél kevesebb eleme lenne, ami nem lehet hiszen a vektortér minden bázisa n elemet tartalmaz. ■

6.4.3. Feladat. Mondjunk olyan vektortereket, melyek dimenziója 6.

6.4.4. Feladat. Határozzuk meg azokat a mátrixok által alkotott vektortereket, melyek dimenziója 2003.

6.4.5. Feladat. Bázist alkotnak-e $\mathbb{R}^2$-ben az

$$\overline{\mathbf{a}}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \overline{\mathbf{a}}_2 = \begin{bmatrix} 0 \\ 5 \end{bmatrix}, \quad \overline{\mathbf{a}}_3 = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$$

vektorok?

Ellenőrző kérdések

1. Mely generátorrendszerekből lehet úgy elhagyni vektort, hogy a megmaradt vektorrendszer is generátorrendszer legyen?

2. Milyen vektorrendszert hívunk bázisnak?

3. Mit mond ki a kicserélési tétel? Milyen következményei van ennek a tételnek?

4. Mit mondhatunk egy vektortér két bázisának elemszámairól. Mit hívunk dimenziónak?

5. Mondjunk néhány példát véges ill. végtelen dimenziós vektorterekre. Mekkora ezek dimenziója?

A feladatok megoldásai

6.4.1.
A síkvektorok vektorterében bázist alkot bármely két vektor, melyek nem esnek egy egyenesre. Leggyakrabban két egymásra merőleges vektort szoktunk bázisnak választani. A térvektorok terében bázis bármely három vektor, melyek nem esnek egy síkba. Általában három egymásra merőleges vektort szoktunk bázisnak választani. A legfeljebb másodfokú polinomok terében a legegyszerűbb bázis az $1, x$ és x^2 polinomok. $\mathbb{R}^2$-ben az

$$\overline{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \overline{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

vektorokat szoktuk bázisnak választani. $\mathbb{R}^{2\times 2}$-ben bázist alkotnak az

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

mátrixok.

6.4.2.
Ahhoz, hogy megállapítsuk, hogy a két vektor bázist alkot-e két dolgot kell megvizsgálnunk: azt, hogy generátorrendszert alkotnak-e és azt, hogy függetlenek-e. Kezdjük a függetlenséggel. Ha a nullvektort akarjuk a két vektorral előállítani, akkor $\mathbf{a}_1$-et biztosan nullával kell szoroznunk, mert különben a lineáris kombinációjuk első eleme nem

lehetne nulla. Ezekután már $\mathbf{a}_2$-t is nullával kell szorozni. Most vizsgáljuk meg, hogy generátorrendszert alkotnak-e. Ha azt alkotnak, akkor minden $x, y \in \mathbb{R}$ számhoz kell találnunk $\alpha, \beta \in \mathbb{R}$ valós számokat, melyekre

$$\alpha \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 5 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Könnyen leellenőrizhetjük, hogy az $\alpha = x/2$ és $\beta = (2y - 3x)/10$ választás megfelelő. Így a vektorok bázist alkotnak.

6.4.3.
Hatdimenziós pl. a legfeljebb ötödfokú polinomok vektortere, valamint $\mathbb{R}^6$, $\mathbb{R}^{2\times3}$, $\mathbb{R}^{3\times2}$ ill. $\mathbb{R}^{1\times6}$.

6.4.4.
Az $k \times n$-es mátrixok vektorterének dimenziója $k \cdot n$. Mivel 2003 prímszám, ezért 2003 csak kétféleképpen áll elő szorzatként: $2003 = 1 \cdot 2003 = 2003 \cdot 1$. Tehát csak két típusú mátrixokból álló vektortérnek 2003 a dimenziója, a $\mathbb{R}^{2003\times1} = \mathbb{R}^{2003}$ és $\mathbb{R}^{1\times2003}$ tereknek.

6.4.5.
$\mathbb{R}^2$ dimenziója kettő, tehát három elemű bázisa nem lehet. Így a vektorok nem alkotnak bázist.

6.5. Koordináták, elemi bázistranszformáció

Egy bázisban minden vektor csak egyféleképpen írható fel lineáris kombinációként. Ennek a ténynek a segítségével definiáljuk a koordináták fogalmát. Megvizsgáljuk, hogy hogyan változnak meg egy vektor koordinátái egy másik bázisra való áttérés során. Az itt bemutatott módszer – az elemi bázistranszformáció – nagyon fontos és gyakori eszközünk lesz a továbbiakban.

Az előző fejezetben megmutattuk, hogy a legkevesebb elemet tartalmazó generátorrendszerek a bázisok. A bázisok elemei lineárisan független rendszert alkotnak. Ebből rögtön következik a 6.3.3. tétel alkalmazásával, hogy egy bázis vektorainak segítségével minden vektortérbeli vektor pontosan egyféleképpen írható csak fel lineáris kombinációként. (Itt emlékeztetünk arra, hogy két lineáris kombinációt akkor tekintettünk egyenlőnek, ha bennük ugyanazon vektorok ugyanazon valós számokkal szorzódnak.)

6.5.1. Definíció. Legyen $\mathbf{a}_1, \ldots, \mathbf{a}_n$ egy n-dimenziós V vektortér bázisa. Ekkor egy tetszőleges $\mathbf{a} \in V$ vektor $\mathbf{a}_1, \ldots, \mathbf{a}_n$ bázisra vonatkozó **koordinátáinak** hívjuk az

$$\mathbf{a} = \alpha_1 \mathbf{a}_1 + \ldots + \alpha_n \mathbf{a}_n$$

előállításban szereplő $\alpha_1, \ldots, \alpha_n$ konstansokat. $\diamondsuit$

Általában olyan sorrendben adjuk meg a koordinátákat, amilyen sorrendben a bázisvektorokhoz tartoznak. Az α_i számot az $\mathbf{a}_i$ vektorhoz tartozó (vagy az $\mathbf{a}_i$ vektorra vonatkozó) koordinátának is szoktuk nevezni.

6.5.1. Feladat. Határozzuk meg az $x^2 + 2x - 3$ polinom koordinátáit a legfeljebb másodfokú polinomok vektorterének $x^2, x, 1$ és $1+x, 2+x, x^2/2$ bázisában.

6.5.2. Feladat. Mik lesznek a koordinátái a $\overline{\mathbf{b}}$ vektornak az $\overline{\mathbf{a}}_1$ és $\overline{\mathbf{a}}_2$ vektorok által generált vektortérben?

$$\overline{\mathbf{b}} = \begin{bmatrix} 13 \\ 2 \end{bmatrix}, \quad \overline{\mathbf{a}}_1 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \quad \overline{\mathbf{a}}_2 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}.$$

Fontos megjegyezni, hogy a koordináták mindig csak egy rögzített bázisra vonatkoztatva érvényesek, koordináták említése a bázis megadása nélkül semmitmondó. Gondoljuk csak el, hogy a "Kati 5 évvel fiatalabb" állítás is semmitmondó, ha nem mondjuk meg, hogy kihez viszonyítjuk az életkort. Kati lehet 5 évvel fiatalabb, de 10 évvel idősebb is attól függően, hogy kivel hasonlítjuk össze az életkorát. Ha Zsuzsi helyett mondjuk Évivel hasonlítjuk össze, akkor megváltozik Kati "viszonylagos" életkora. Ehhez hasonló jelenség játszódik le akkor is, amikor egyik bázisról áttérünk egy másikra: megváltoznak a vektorok koordinátái. Most azt vizsgáljuk, hogy hogyan változnak meg a koordináták, ha a bázis egyik vektorát kicseréljük egy másik vektorra. Ezt az eljárást hívjuk elemei bázistranszformációnak.

6.5.2. Tétel. *(Elemei bázistranszformáció.) Legyen $\mathbf{a}_1, \ldots, \mathbf{a}_n \in V$ a V vektortér egy bázisa. Ebben a bázisban egy tetszőleges $\mathbf{x}$ vektor előállítása legyen*

$$\mathbf{x} = \chi_1 \mathbf{a}_1 + \ldots + \chi_n \mathbf{a}_n.$$

Legyen továbbá $\mathbf{a} = \alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n$ *egy olyan vektor, amelyre* $\alpha_k \neq 0$ *valamilyen* $k = 1, \ldots, n$ *mellett. Ekkor az* $\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}, \mathbf{a}, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n$ *vektorok is bázist alkotnak, és ebben a bázisban az* $\mathbf{x}$ *vektor előállítása*

$$\mathbf{x} = (\chi_1 - \lambda\alpha_1)\mathbf{a}_1 + (\chi_2 - \lambda\alpha_2)\mathbf{a}_2 + \ldots + (\chi_{k-1} - \lambda\alpha_{k-1})\mathbf{a}_{k-1} +$$
$$+\lambda\mathbf{a} + (\chi_{k+1} - \lambda\alpha_{k+1})\mathbf{a}_{k+1} + \ldots + (\chi_n - \lambda\alpha_n)\mathbf{a}_n,$$

ahol $\lambda = \chi_k/\alpha_k$.

Bizonyítás. Mivel az α_k együtthatók között van nullától különböző, ezért $\mathbf{a}$ mint egyedüli vektor lineárisan független. Emiatt a kicserélési tétel (6.4.3. tétel) és annak bizonyításában elmondottak szerint az $\mathbf{a}$ vektor becserélhető az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ bázisba egy olyan $\mathbf{a}_k$ vektor helyébe, melynek együtthatója nem nulla ($\alpha_k \neq 0$). Az így kapott

$$\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}, \mathbf{a}, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n$$

vektorok továbbra is bázist alkotnak. Az

$$\mathbf{a} = \alpha_1\mathbf{a}_1 + \ldots + \alpha_{k-1}\mathbf{a}_{k-1} + \alpha_k\mathbf{a}_k + \alpha_{k+1}\mathbf{a}_{k+1}$$

egyenlőségből kifejezzük az $\mathbf{a}_k$ vektort

$$\mathbf{a}_k = -\frac{\alpha_1}{\alpha_k}\mathbf{a}_1 - \ldots - \frac{\alpha_{k-1}}{\alpha_k}\mathbf{a}_{k-1} + \frac{1}{\alpha_k}\mathbf{a} - \frac{\alpha_{k+1}}{\alpha_k}\mathbf{a}_{k+1} - \ldots - \frac{\alpha_n}{\alpha_k}\mathbf{a}_n, \tag{6.5.1}$$

és az $\mathbf{x} = \chi_1\mathbf{a}_1 + \ldots + \chi_k\mathbf{a}_k + \ldots + \chi_n\mathbf{a}_n$ képletbe helyettesítve kapjuk a tételben szereplő előállítást. ■

Az elemi bázistranszformáció fontossága miatt most leírjuk szavakkal is, hogy hogyan kell elvégezni. Először megállapítjuk, hogy az adott $\mathbf{a}$ vektor melyik bázisvektor helyébe cserélhető be. Olyan vektor helyébe lehet becserélni, melynek együtthatója $\mathbf{a}$ előállításában nem nulla. Legyen ez mondjuk az $\mathbf{a}_k$ vektor. Ekkor kiszámítjuk a $\lambda = \chi_k/\alpha_k$ értéket, ahol χ_k és α_k rendre az $\mathbf{x}$ és $\mathbf{a}$ vektorok előállításában az $\mathbf{a}_k$ vektor együtthatói. Ezekután az $\mathbf{x}$ vektor új bázisbeli koordinátáit úgy kaphatjuk meg, hogy $\mathbf{x}$ régi koordinátáiból ki kell vonni $\mathbf{a}$ koordinátáinak λ-szorosait, kivéve az $\mathbf{a}$ vektor koordinátáját, ami λ lesz. Tekintsünk most egy példát.

6.5.3. Példa. Tegyük fel, hogy az $\mathbf{a}_1, \mathbf{a}_2$ és $\mathbf{a}_3$ vektorok bázist alkotnak egy vektortérben. Ebben a bázisban az $\mathbf{x}$ vektor az

$$\mathbf{x} = 3\mathbf{a}_1 + \mathbf{a}_2 + 2\mathbf{a}_3$$

alakban írható fel. Cseréljük ki az $\mathbf{a}_3$ báziselemet az $\mathbf{a} = 2\mathbf{a}_1 + 3\mathbf{a}_2 + 5\mathbf{a}_3$ vektorra (ha lehet), és írjuk fel az $\mathbf{x}$ vektort ebben az új bázisban.

Az $\mathbf{a}_3$ vektor tényleg kicserélhető az $\mathbf{a}$ vektorra, mivel $\mathbf{a}$ előállításában $\mathbf{a}_3$ együtthatója nem nulla (nevezetesen 5). Kiszámítjuk λ-t, amely $\mathbf{a}_3$-nak az $\mathbf{x}$-beli ill. $\mathbf{a}$-beli együtthatójának hányadosa. Tehát $\lambda = 2/5$. $\mathbf{x}$ új koordinátáit úgy kapjuk meg, hogy a régiekből kivonjuk $\mathbf{a}$ megfelelő koordinátáinak λ-szorosát. Így $\mathbf{a}_1$ együtthatója $\mathbf{x}$ előállításában 3-ról $3 - (2/5)\cdot 2 = 11/5$-re változik. Hasonlóan $\mathbf{a}_2$ együtthatója 1-ről $1 - (2/5)\cdot 3 = -1/5$-re módosul. Az $\mathbf{a}$ vektor együtthatója (mivel ezt cseréltük be a bázisba) $\mathbf{x}$ előállításában λ lesz. Ezek alapján $\mathbf{x}$ előállítása az új bázisban

$$\mathbf{x} = \frac{11}{5}\mathbf{a}_1 - \frac{1}{5}\mathbf{a}_2 + \frac{2}{5}\mathbf{a}.$$

□

6.5.3. Feladat. Igazoljuk, hogy a 6.5.3. példában $\mathbf{x}$ régi és új előállítása valóban ugyanazt a vektort adja meg.

A számítások megkönnyítése és átláthatósága miatt az elemi bázistranszformáció lépéseit táblázatos formában szoktuk elvégezni. Most megmutatjuk az előző példán ennek menetét.

6.5.4. Példa. Készítsük el a

	$\mathbf{a}$	$\mathbf{x}$
$\mathbf{a}_1$	2	3
$\mathbf{a}_2$	3	1
$\mathbf{a}_3$	5	2

táblázatot és figyeljük meg annak szerkezetét. A táblázat bal oldali oszlopában vannak az éppen aktuális bázis elemei, a legfelső sorban pedig az a vektor, amit be akarunk cserélni a bázisba ill. az a vektor, melynek kíváncsiak vagyunk az új koordinátáira. A táblázat többi oszlopában szereplő számok a fenti sorban szereplő vektorok koordinátáit jelentik a bal oldalon látható bázisban. Pl. az $\mathbf{a}_3$ sorában és $\mathbf{a}$ oszlopában szereplő 5-ös szám azt jelenti, hogy $\mathbf{a}$-nak az $\mathbf{a}_3$-hoz tartozó koordinátája 5.

A bázistranszformációt a következő módon végezzük (vessük össze a megoldási menetet a 6.5.3. példával). Az $\mathbf{a}$ vektor becserélhető $\mathbf{a}_3$ helyébe, mert az $\mathbf{a}_3$ sorában és $\mathbf{a}$ oszlopában álló elem nem nulla (nevezetesen 5). Ezt az elemet fogjuk ún. generálóelemnek választani és úgy jelöljük, hogy a táblázatban bekeretezzük az 5-ös számot.

	$\mathbf{a}$	$\mathbf{x}$
$\mathbf{a}_1$	2	3
$\mathbf{a}_2$	3	1
$\mathbf{a}_3$	$\boxed{5}$	2

Kiszámítjuk λ értékét, és a táblázat legalsó sorába írjuk az $\mathbf{x}$ vektornak megfelelő rubrikába.

	$\mathbf{a}$	$\mathbf{x}$
$\mathbf{a}_1$	2	3
$\mathbf{a}_2$	3	1
$\mathbf{a}_3$	$\boxed{5}$	2
		2/5

Most a tanult módon megváltoztatjuk $\mathbf{x}$ koordinátáit. Az $\mathbf{x}$ vektor koordinátáiból rendre kivonjuk a generálóelem oszlopában lévő koordináták $\lambda = 2/5$-szorosait. A generálóelem sorában pedig $\mathbf{x}$ koordinátáját $\lambda = 2/5$-re változtatjuk. Az így nyert táblázat a következő lesz.

	$\mathbf{x}$
$\mathbf{a}_1$	$3 - (2/5) \cdot 2 = 11/5$
$\mathbf{a}_2$	$1 - (2/5) \cdot 3 = -1/5$
$\mathbf{a}$	2/5
	2/5

A táblázat bal oldala már az új bázist tartalmazza. Ebben az új bázisban $\mathbf{a} = 0\mathbf{a}_1 + 0\mathbf{a}_2 + 1\mathbf{a}$, ezért $\mathbf{a}$ oszlopát már el is hagytuk a táblázatból. Az $\mathbf{x}$ vektor oszlopából leolvashatóak $\mathbf{x}$ új bázisbeli koordinátái: $\mathbf{x} = (11/5)\mathbf{a}_1 - (1/5)\mathbf{a}_2 + (2/5)\mathbf{a}$. □

6.5.4. Feladat. Tekintsük a

	a	**x**
$\mathbf{a}_1$	3	2
$\mathbf{a}_2$	3	−1
$\mathbf{a}_3$	0	2
$\mathbf{a}_4$	1	0

bázistranszformációs táblázatot. Adjuk meg $\mathbf{a}_2$ és $\mathbf{a}$ előállítását a táblázat által mutatott aktuális bázisban. Cseréljük be az $\mathbf{a}$ vektort az $\mathbf{a}_3$ báziselem helyébe. Mik lesznek $\mathbf{x}$ új koordinátái? Cseréljük be az $\mathbf{a}$ vektort az $\mathbf{a}_1$ báziselem helyébe. Mik lesznek $\mathbf{x}$ új koordinátái?

6.5.5. Feladat. Írjuk fel a $2x^2+x+1$ polinom koordinátáit az $x^2, x, 1$ bázisban. Bázistranszformáció segítségével írjuk fel a polinom koordinátáit az $x^2, x, x+1$ bázisban is.

A legtöbb esetben persze nem csak egy elemét szeretnénk kicserélni egy éppen aktuális bázisnak, hanem többet. Ezt elemi bázistranszformációk többszöri alkalmazásával hajthatjuk végre. Lássunk erre egy példát.

6.5.5. Példa. Határozzuk meg, hogy hogyan állítható elő az $\overline{\mathbf{x}}$ vektor az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3$ $\mathbb{R}^3$-beli bázis segítségével,

$$\overline{\mathbf{x}} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \quad \overline{\mathbf{a}}_1 = \begin{bmatrix} -2 \\ 4 \\ 6 \end{bmatrix}, \quad \overline{\mathbf{a}}_2 = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \quad \overline{\mathbf{a}}_3 = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}.$$

Vegyük észre, hogy az $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \overline{\mathbf{e}}_3$ vektorok által alkotott bázisban egyszerűen fel tudjuk írni az $\overline{\mathbf{x}}, \overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2$ és $\overline{\mathbf{a}}_3$ vektorokat. Pl. $\overline{\mathbf{x}} = 3\overline{\mathbf{e}}_1 + \overline{\mathbf{e}}_2 + 2\overline{\mathbf{e}}_3$ (lásd a 6.4.8. példát az $\overline{\mathbf{e}}_i$ jelöléssel kapcsolatban). Legyen ez a kiinduló bázisunk és készítsük el a bázistranszformációs táblázatot.

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{x}}$
$\overline{\mathbf{e}}_1$	**-2**	3	2	3
$\overline{\mathbf{e}}_2$	4	0	3	1
$\overline{\mathbf{e}}_3$	6	2	5	2
		−3/2	−1	−3/2

A táblázatban már választottunk is generálóelemet, $\overline{\mathbf{a}}_1$ bevihető $\overline{\mathbf{e}}_1$ helyébe. A báziscsere miatt nem csak az $\overline{\mathbf{x}}$ koordinátái változnak, hanem az $\overline{\mathbf{a}}_2$ és $\overline{\mathbf{a}}_3$ vektoroké is. Ezen vektorokhoz tartozó λ értékeket is feltüntettük a táblázatban. A bázistranszformáció utáni új táblázat a következő lesz.

	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{x}}$
$\overline{\mathbf{a}}_1$	−3/2	−1	−3/2
$\overline{\mathbf{e}}_2$	6	7	7
$\overline{\mathbf{e}}_3$	**11**	11	11
		1	1

Vigyük be most az $\overline{\mathbf{a}}_2$ vektort az $\overline{\mathbf{e}}_3$ helyébe. Ezzel elérjük, hogy a λ értékek egész számok

legyenek, melyekkel könnyebb számolni.

	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{x}}$
$\overline{\mathbf{a}}_1$	1/2	0
$\overline{\mathbf{e}}_2$	[1]	1
$\overline{\mathbf{a}}_2$	1	1
		1

Most már csak az $\overline{\mathbf{a}}_3$ vektort vihetjük be $\overline{\mathbf{e}}_2$ helyébe.

	$\overline{\mathbf{x}}$
$\overline{\mathbf{a}}_1$	$-1/2$
$\overline{\mathbf{a}}_3$	1
$\overline{\mathbf{a}}_2$	0

Tehát az $\overline{\mathbf{x}}$ vektor az $\overline{\mathbf{x}} = -(1/2)\overline{\mathbf{a}}_1 + \overline{\mathbf{a}}_3$ alakban írható fel. □

Most néhány olyan észrevétel következik, melyek meggyorsítják a bázistranszformáció elvégzését. Vegyük észre először, hogy ha a generálóelem oszlopában valamelyik elem nulla, akkor annak a sora nem fog megváltozni a transzformáció során (nullának a λ-szorosát, azaz nullát vonunk ki a régi koordinátákból). Másik észrevétel, hogy ha a generálóelem sorában van nulla valahol, akkor az ehhez az elemhez tartozó oszlop változatlan marad a transzformációban, mert erre az oszlopra a λ nulla lesz. A generálóelem sora úgy változik, hogy minden elemet végig kell osztani a generálóelemmel, hiszen ebbe a sorba a λ értékeket kell írni. Végül megjegyezzük, hogy helytakarékossági okokból és a számítások meggyorsításának érdekében a táblázatokat egy sorba szoktuk írni. A 6.5.5. példa táblázatait pl. ezután

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{x}}$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{x}}$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{x}}$	$\overline{\mathbf{x}}$
$\mathbf{a}_1 \not\in \mathbf{e}_1$	[-2]	3	2	3	$-3/2$	-1	$-3/2$	1/2	0	$-1/2$
$\mathbf{a}_3 \not\in \mathbf{e}_2$	4	0	3	1	6	7	7	[1]	1	1
$\mathbf{a}_2 \not\in \mathbf{e}_3$	6	2	5	2	[11]	11	11	1	1	0
		$-3/2$	-1	$-3/2$		1	1		1	

alakban fogjuk leírni, ahol dupla elválasztó vonalak jelzik az egyes táblázatok határait. A bal oldalon a régi bázis elemeit áthúzva írtuk be az új bázis elemeit. A bázisvektorok cseréjének sorrendjét a bekeretezett generálóelemekből azonnal látni. Ha már gyakorlatot szereztünk a bázistranszformációk elvégzésében, akkor az alsó sorból a λ számokat elhagyhatjuk.

6.5.6. Feladat. Írjuk fel az x^2+x+1 polinomot az $x^2+1, x-2, x^2-1$ bázisban lineáris kombinációként.

Az Operációkutatás tantárgy keretében az elemi bázistranszformációkat lineáris programozási feladatok megoldására fogjuk használni. Egy ún. célfüggvény maximális (vagy minimális) értékét kell majd meghatároznunk. Ez úgy történik, hogy valamilyen szabály szerint vektorokat viszünk be egy bázisba és közben figyeljük a célfüggvény értékeit. Lineáris programozási feladatoknál előfordulhatnak olyan esetek, hogy egy bázisba bevitt vektort újra ki kell hoznunk a bázisból egy másik vektor becserélésének szükségessége

miatt. Ezekben az esetekben célszerű a bázisból kikerülő vektorok koordinátáit is feljegyezni és a további bázistranszformációk során ezekkel is számolni. (Eddig a bázisból kicserélt vektorokat már elhagytuk a táblázatból, hiszen biztosak lehettünk benne, hogy nem fogjuk őket visszacserélni.) A bázisból kikerülő vektor koordinátáit a (6.5.1) képlet segítségével számíthatjuk ki.

6.5.7. Feladat. Az $\mathbb{R}^3$ vektortér $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \overline{\mathbf{e}}_3$ bázisának $\overline{\mathbf{e}}_3$ tagját cseréljük ki az

$$\overline{\mathbf{a}} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$$

vektorra. Mik lesznek $\overline{\mathbf{e}}_3$ koordinátái az új bázisban?

Az elemi bázistranszformáció nagyon fontos módszer a matematikában. Használni fogjuk pl. lineáris egyenletrendszerek megoldására, mátrixok rangjának és inverzének kiszámítására ill. determinánsok meghatározására. Ahogy említettük már, lineáris programozási feladatok kapcsán is fogunk majd vele találkozni az Operációkutatás tantárgy keretében.

Ellenőrző kérdések

1. Hányféleképpen lehet egy vektortér elemeit egy bázisának lineáris kombinációival előállítani?

2. Mit hívunk egy vektor koordinátáinak? Milyen kapcsolatban vannak a koordináták a bázisokkal?

3. Mit nevezünk elemi bázistranszformációnak? Ismertessük a rá vonatkozó tételt.

4. Hogyan hajtjuk végre táblázattal az elemi bázistranszformációt. Szemléltessük az elmondottakat egy példán.

5. Mik lesznek a koordinátái egy régi bázisból kikerülő vektornak az új bázisban?

A feladatok megoldásai

6.5.1.
A polinom lineáris kombinációs előállítása a két bázisban

$$x^2 + 2x - 3 = 1(x^2) + 2(x) - 3(1), \quad x^2 + 2x - 3 = 7(1 + x) + (-5)(2 + x) + 2(x^2/2).$$

Tehát az első bázisban a koordináták $1, 2, -3$, míg a másodikban $7, -5, 2$. Tehát a koordináták függnek a bázis megválasztásától.

6.5.2.
A

$$\begin{bmatrix} 13 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 5 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

előállításból következik, hogy a koordináták 2 és 1.

6.5.3.
Induljunk ki az $\mathbf{x} = (11/5)\mathbf{a}_1 - (1/5)\mathbf{a}_2 + (2/5)\mathbf{a}$ új előállításból. Ekkor az

$$\mathbf{a} = 2\mathbf{a}_1 + 3\mathbf{a}_2 + 5\mathbf{a}_3$$

előállítást alkalmazva átrendezések után a következő egyenlőségeket kapjuk

$$\mathbf{x} = \frac{11}{5}\mathbf{a}_1 - \frac{1}{5}\mathbf{a}_2 + \frac{2}{5}\mathbf{a} = \frac{11}{5}\mathbf{a}_1 - \frac{1}{5}\mathbf{a}_2 + \frac{2}{5}(2\mathbf{a}_1 + 3\mathbf{a}_2 + 5\mathbf{a}_3) = 3\mathbf{a}_1 + \mathbf{a}_2 + 2\mathbf{a}_3,$$

ami mutatja, hogy a két előállítás ugyanazon vektort adja.

6.5.4.
Mivel $\mathbf{a}_2$ a bázisban van, ezért $\mathbf{a}_2 = 0\mathbf{a}_1 + 1\mathbf{a}_2 + 0\mathbf{a}_3 + 0\mathbf{a}_4$. Az $\mathbf{a}$ vektor koordinátái $\mathbf{a}$ oszlopából olvashatók le, azaz $\mathbf{a} = 3\mathbf{a}_1 + 3\mathbf{a}_2 + 0\mathbf{a}_3 + 1\mathbf{a}_4$. Mivel $\mathbf{a}$ oszlopában nulla áll $\mathbf{a}_3$ sorában, ezért az $\mathbf{a}$ vektor nem becserélhető az $\mathbf{a}_3$ helyébe. Az $\mathbf{a}_1$ vektor becserélhető, becserélés után a

	$\mathbf{x}$
$\mathbf{a}$	$2/3$
$\mathbf{a}_2$	-3
$\mathbf{a}_3$	2
$\mathbf{a}_4$	$-2/3$

táblázatot kapjuk. Így $\mathbf{x}$ új előállítása $\mathbf{x} = (2/3)\mathbf{a} - 3\mathbf{a}_2 + 2\mathbf{a}_3 - (2/3)\mathbf{a}_4$.

6.5.5.
A $2x^2 + x + 1$ polinom koordinátái az $x^2, x, 1$ bázisban nyilvánvalóan 2,1, és 1. Az új bázisbeli koordinátáit úgy kapjuk meg, hogy a bázis 1-es elemét kicseréljük $x + 1$-re. A kezdő bázistranszformációs táblázat tehát

	$x+1$	$2x^2+x+1$
x^2	0	2
x	1	1
1	$\boxed{1}$	1
		1

a bázistranszformáció utáni pedig

	$2x^2+x+1$
x^2	2
x	0
$x+1$	1
	1

Emiatt az új bázisbeli koordináták $2, 0$ és 1 ($2x^2 + x + 1 = 2(x^2) + 0x + 1(x + 1)$).

6.5.6.
Az $x^2, x, 1$ bázisban könnyen felírhatjuk az adott polinomot. Erről a bázisról áttérünk a feladatban megadott $p_1 = x^2 + 1$, $p_2 = x - 2$, $p_3 = x^2 - 1$ bázisra elemei bázistranszformációk segítségével.

	p_1	p_2	p_3	x^2+x+1	p_2	p_3	x^2+x+1	p_3	x^2+x+1	x^2+x+1
$p_1 \not\in x^2$	$\boxed{1}$	0	1	1	0	1	1	1	1	2
$p_2 \not\in x$	0	1	0	1	$\boxed{1}$	0	1	0	1	1
$p_3 \not\in 1$	1	-2	-1	1	-2	-2	0	$\boxed{-2}$	2	-1
		0	1	1		0	1		-1	

A keresett előállítás tehát $x^2 + x + 1 = 2(x^2 + 1) + 1(x - 2) - 1(x^2 - 1)$.

6.5.7.
Mivel $\overline{\mathbf{a}}$ harmadik eleme nem nulla, ezért valóban becserélhető a vektor a bázisba. A (6.5.1) képlet szerint $\overline{\mathbf{e}}_3$ új koordinátái $1/2, 3/2$ és $-1/2$, azaz $\overline{\mathbf{e}}_3 = (1/2)\overline{\mathbf{e}}_1 + (3/2)\overline{\mathbf{e}}_2 - (1/2)\overline{\mathbf{a}}$.

7. fejezet

Alkalmazások

7.1. Vektorrendszerek és mátrixok rangja

Ebben a fejezetben értelmezzük a vektorrendszerek és mátrixok rangját, melyet bázistranszformációval számolunk ki.

Bázistranszformáció segítségével könnyen el tudjuk dönteni, hogy bizonyos vektorok függetlenek vagy sem. Gondoljunk a kicserélési tételre (lásd 6.4.3. tétel), amely szerint lineárisan független vektorok bármely rendszere becserélhető egy bázis elemei közé. Ezek szerint, ha egy vektorrendszer minden elemét be tudjuk cserélni a bázisba, akkor a vektorok függetlenek; ellenkező esetben összefüggőek.

7.1.1. Definíció. A V véges dimenziós vektortérben az $\mathbf{a}_1, \ldots, \mathbf{a}_n$ **vektorok rangjának** hívjuk a $\dim(L(\mathbf{a}_1, \ldots, \mathbf{a}_n))$ számot, azaz a vektorok által generált vektortér dimenzióját. Jelölése: $r(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. $\diamondsuit$

Mivel a definíció szerint $r(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \dim(L(\mathbf{a}_1, \ldots, \mathbf{a}_n))$, ezért a vektorrendszerben nem lehet a rangjánál több lineárisan független vektor. Emiatt azt is szoktuk mondani, hogy egy vektorrendszer rangja a vektorrendszerből kiválasztható lineárisan független vektorok maximális számát adja meg.

Nyilvánvaló, hogy $0 \leq r(\mathbf{a}_1, \ldots, \mathbf{a}_n) \leq n$, és ha a vektorok lineárisan függetlenek, akkor $r(\mathbf{a}_1, \ldots, \mathbf{a}_n) = n$. Szintén könnyű látni, hogy $r(\mathbf{a}_1, \ldots, \mathbf{a}_n) \leq \dim(V)$.

A rang úgy számolható ki, hogy a vektorrendszer elemeit becseréljük egy bázis elemeinek helyére. Amennyit be tudunk vinni közülük a bázisba, az lesz a vektorrendszer ranga. Lássunk egy példát.

7.1.2. Példa. Az

$$\overline{\mathbf{a}}_1 = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}, \quad \overline{\mathbf{a}}_2 = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \quad \overline{\mathbf{a}}_3 = \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix}, \quad \overline{\mathbf{a}}_4 = \begin{bmatrix} 2 \\ 8 \\ -4 \end{bmatrix}$$

vektorrendszer rangját fogjuk meghatározni. Mivel $\dim(\mathbb{R}^3) = 3$, ezért a fent szereplő vektorok nem lehetnek lineárisan függetlenek, így biztosan $r(\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3, \overline{\mathbf{a}}_4) \leq 3$. Cseréljük be a vektorokat az $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \overline{\mathbf{e}}_3$ bázisvektorok helyébe elemi bázistranszformációt alkalmazva.

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$
$\overline{\mathbf{a}}_1 \not\overline{\mathbf{e}}_1$	$\boxed{1}$	2	3	2	3	3	2	0	2
$\overline{\mathbf{a}}_2 \not\overline{\mathbf{e}}_2$	4	1	5	8	$\boxed{-7}$	-7	0	1	0
$\overline{\mathbf{e}}_3$	-2	4	2	-4	8	8	0	0	0
		2	3	2		1	0		

Mivel $\overline{\mathbf{a}}_3$ és $\overline{\mathbf{a}}_4$ már nem becserélhető, így $r(\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3, \overline{\mathbf{a}}_4) = 2$. □

7.1.3. Definíció. Egy $\mathbf{A} \in \mathbb{R}^{k\times n}$ mátrix oszlopvektorai által alkotott vektorrendszer rangját az **A mátrix rangjának** hívjuk. Jelölése: $r(\mathbf{A})$. ♢

A mátrixok rangját a fenti definícióval a vektorrendszerek rangjára vezettük vissza. Az ottani észrevételek alapján, ha $\mathbf{A} \in \mathbb{R}^{k\times n}$, akkor $r(\mathbf{A}) \leq \min\{n, k\}$. Ez egyrészt abból következik, hogy a mátrix oszlopvektorai a k-dimenziós $\mathbb{R}^k$ térből kerülnek ki, melyből maximálisan csak k független vektort lehet kiválasztani; másrészt az oszlopvektorok darabszáma n, így n-nél több független vektor sem lehet köztük.

7.1.4. Példa. Az

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & 3 & -1 \\ 3 & 1 & 4 & 2 \\ 1 & 1 & 1 & 2 \end{bmatrix}$$

mátrix rangjának meghatározásához azt kell eldönteni, hogy a mátrix oszlopvektorai közül (legyenek ezek $\overline{\mathbf{a}}_1, \ldots, \overline{\mathbf{a}}_4$) hány cserélhető be az $\overline{\mathbf{e}}_1, \ldots, \overline{\mathbf{e}}_4$ bázisba. A bázistranszformáció táblázata:

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_3$
$\overline{\mathbf{a}}_1 \not\overline{\mathbf{e}}_1$	$\boxed{1}$	2	3	4	2	3	4	-1	0	-1
$\overline{\mathbf{a}}_4 \not\overline{\mathbf{e}}_2$	2	-1	3	-1	-5	-3	-9	7	$\boxed{1}$	7
$\overline{\mathbf{a}}_3 \not\overline{\mathbf{e}}_3$	3	1	4	2	-5	-5	-10	5	0	$\boxed{5}$
$\overline{\mathbf{a}}_2 \not\overline{\mathbf{e}}_4$	1	1	1	2	$\boxed{-1}$	-2	-2	2	2	-12

Mivel négy oszlopvektort tudunk becserélni, ezért $r(\mathbf{A}) = 4$. □

Ellenőrző kérdések

1. Hogyan definiáljuk egy vektorrendszer rangját?
2. Hogyan vezetjük vissza a mátrixok rangját vektorrendszerek rangjára?
3. Milyen összefüggés van egy vektorrendszer rangja és a vektorok függetlensége ill. összefüggősége között?

7.2. Mátrixok determinánsának kiszámítása

Az 5.2. fejezetben megismertük, hogy hogyan lehet kvadratikus mátrixok determinánsát kiszámítani a determináns tulajdonságai alapján. Most a determinánst bázistranszformáció segítségével számítjuk ki.

Ebben a fejezetben azt vizsgáljuk, hogy hogyan lehet egy kvadratikus mátrix determinánsát bázistranszformáció segítségével kiszámítani. Legyen adott egy $\mathbf{A} = [a_{ij}]_{n\times n}$ kvadratikus mátrix. Oszlopvektorait jelöljék rendre az $\overline{\mathbf{a}}_1, \ldots, \overline{\mathbf{a}}_n$ vektorok. Tegyük fel, hogy az $\overline{\mathbf{e}}_1, \ldots, \overline{\mathbf{e}}_n$ bázisvektorokat ki akarjuk cserélni az $\overline{\mathbf{a}}_1, \ldots, \overline{\mathbf{a}}_n$ vektorokra. Hajtsunk végre egy elemi bázistranszformációt és vizsgáljuk meg, hogy hogyan változik a mátrix a bázistranszformáció során. A

	$\overline{\mathbf{a}}_1$	$\ldots$	$\overline{\mathbf{a}}_k$	$\ldots$	$\overline{\mathbf{a}}_n$
$\overline{\mathbf{e}}_1$	a_{11}	$\ldots$	a_{1k}	$\ldots$	a_{1n}
$\overline{\mathbf{e}}_2$	a_{21}	$\ldots$	a_{2k}	$\ldots$	a_{2n}
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$\overline{\mathbf{e}}_{l-1}$	$a_{l-1,1}$	$\ldots$	$a_{l-1,k}$	$\ldots$	$a_{l-1,n}$
$\overline{\mathbf{e}}_l$	a_{l1}	$\ldots$	$\boxed{a_{lk}}$	$\ldots$	a_{ln}
$\overline{\mathbf{e}}_{l+1}$	$a_{l+1,1}$	$\ldots$	$a_{l+1,k}$	$\ldots$	$a_{l+1,n}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$\overline{\mathbf{e}}_n$	a_{n1}	$\ldots$	a_{nk}	$\ldots$	a_{nn}

táblázatból indulva cseréljük be az $\overline{\mathbf{a}}_k$ vektort az $\overline{\mathbf{e}}_l$ vektor helyébe.

	$\overline{\mathbf{a}}_1$	$\ldots$	$\overline{\mathbf{a}}_k$	$\ldots$	$\overline{\mathbf{a}}_n$
$\overline{\mathbf{e}}_1$	$a_{11} - a_{1k}a_{l1}/a_{lk}$	$\ldots$	0	$\ldots$	$a_{1n} - a_{1k}a_{ln}/a_{lk}$
$\overline{\mathbf{e}}_2$	$a_{21} - a_{2k}a_{l1}/a_{lk}$	$\ldots$	0	$\ldots$	$a_{2n} - a_{2k}a_{ln}/a_{lk}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$\overline{\mathbf{e}}_{l-1}$	$a_{l-1,1} - a_{l-1,k}a_{l1}/a_{lk}$	$\ldots$	0	$\ldots$	$a_{l-1,n} - a_{l-1,k}a_{ln}/a_{lk}$
$\overline{\mathbf{a}}_k$ ~~$\overline{\mathbf{e}}_l$~~	a_{l1}/a_{lk}	$\ldots$	1	$\ldots$	a_{ln}/a_{lk}
$\overline{\mathbf{e}}_{l+1}$	$a_{l+1,1} - a_{l+1,k}a_{l1}/a_{lk}$	$\ldots$	0	$\ldots$	$a_{l+1,n} - a_{l+1,k}a_{ln}/a_{lk}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$\overline{\mathbf{e}}_n$	$a_{n1} - a_{nk}a_{l1}/a_{lk}$	$\ldots$	0	$\ldots$	$a_{nn} - a_{nk}a_{ln}/a_{lk}$

A bázistranszformáció után nyert újabb mátrix (most feltüntettük a bázisba cserélt $\overline{\mathbf{a}}_k$ vektor koordinátáit is az új bázisban) úgy is előállítható az eredetiből, hogy annak l-edik sorát végigosztjuk a_{lk}-val, majd az így kapott sort rendre az $a_{1k}, \ldots, a_{l-1,k}, a_{l+1,k}, \ldots, a_{nk}$ valós számokkal szorozva kivonjuk rendre az $1, \ldots, l-1, l+1, \ldots, n$ sorszámú sorokból. Amikor az l-edik sort végigosztjuk a_{lk}-val, akkor a mátrix determinánsa $1/a_{lk}$-val szorzódik (lásd 5.2.13. tétel). A sorok kivonása közben a determináns értéke nem változik meg (lásd 5.2.15. tétel). Megállapíthatjuk tehát, hogy egy elemi bázistranszformáció során a bázistranszformációs táblázatban szereplő új mátrix determinánsa a régi mátrix determinánsának $1/a_{lk}$-szorosa.

Tegyük fel, hogy az $\mathbf{A}$ mátrix minden oszlopvektorát be tudtuk cserélni az $\overline{\mathbf{e}}_1, \ldots, \overline{\mathbf{e}}_n$ bázisba, és hogy a soronkénti generálóelemek az $\boxed{g_{1i_1}}, \ldots, \boxed{g_{ni_n}}$ számok voltak. A báziscserék után egy olyan mátrixot fogunk kapni a táblázatban, amelyben minden oszlopban és sorban pontosan egy darab 1-es szerepel (a generálóelemek helyén), a többi elem nulla.

Ezen mátrix determinánsa $(-1)^{\mathcal{I}(i_1,\ldots,i_n)}$, és mivel egy bázistranszformáció során a mátrix determinánsa a generálóelemmel osztódik

$$\frac{\det(\mathbf{A})}{\boxed{g_{1i_1}}\cdot\ldots\cdot\boxed{g_{ni_n}}}=(-1)^{\mathcal{I}(i_1,\ldots,i_n)}.$$

Az egyenlőséget átrendezve kapjuk, hogy

$$\det(\mathbf{A})=(-1)^{\mathcal{I}(i_1,\ldots,i_n)}\cdot\boxed{g_{1i_1}}\cdot\ldots\cdot\boxed{g_{ni_n}}. \tag{7.2.1}$$

A fenti képletekben bekereteztük a szereplő elemeket, ezzel is kihangsúlyozva, hogy azok a bázistranszformációk során szereplő generálóelemeket jelentik. Vegyük észre, hogy ha $\mathbf{A}$ minden oszlopvektorát be tudjuk vinni a bázisba, akkor $\det(\mathbf{A})\neq 0$, hiszen a generálóelemek a (7.2.1) felírásban nem lehetnek nullák. A $(-1)^{\mathcal{I}(i_1,\ldots,i_n)}$ kifejezés értéke csak $+1$ vagy (-1) lehet. Elegendő tehát csak a kitevő paritását eldönteni. A kitevő paritása megegyezik azon sorcserék (vagy oszlopcserék) számának paritásával, melyekkel a bázistranszformációk után nyert mátrix egységmátrixszá rendezhető át.

Most tekintsük azt az esetet, amikor a bázistranszformációt nem lehet végigcsinálni, azaz $\mathbf{A}$ oszlopvektorai közül nem lehet mindet becserélni az $\overline{\mathbf{e}}_1,\ldots,\overline{\mathbf{e}}_n$ bázisba. Ez akkor fordulhat csak elő, ha $\mathbf{A}$ oszlopvektorai lineárisan összefüggőek. Ebben az esetben az egyik oszlopvektor kifejezhető a többi lineáris kombinációjaként (lásd 6.3.2. tétel). Az általánosság megszorítása nélkül feltehetjük, hogy pl. $\overline{\mathbf{a}}_1$ felírható

$$\overline{\mathbf{a}}_1=\alpha_2\overline{\mathbf{a}}_2+\ldots+\alpha_n\overline{\mathbf{a}}_n \tag{7.2.2}$$

alakban. Vonjuk ki most $\mathbf{A}$ első oszlopából rendre az $\alpha_2,\ldots,\alpha_n$ számokkal beszorzott $2,\ldots,n$ sorszámú oszlopokat. A (7.2.2) összefüggés miatt, az így nyert mátrix első oszlopa csupa nulla elemekből fog állni, azaz a mátrix determinánsa nulla (lásd 5.2.9. tétel). Ugyanakkor a műveletek során a mátrix determinánsa nem változott (lásd 5.2.15. tétel), ami azt jelenti, hogy az eredeti mátrix determinánsa is nulla volt.

A fentiek alapján egy $\mathbf{A}\in\mathbb{R}^{n\times n}$ mátrix determinánsa a következő módon számítható bázistranszformációval. Egy bázistranszformációs táblázat segítségével az $\mathbf{A}$ oszlopvektorait becseréljük az $\overline{\mathbf{e}}_1,\ldots,\overline{\mathbf{e}}_n$ bázisvektorok helyébe. Ha nem minden vektor becserélhető, akkor $\det(\mathbf{A})=0$. Ha minden vektor becserélhető, akkor a (7.2.1) képlet segítségével számítható ki a determináns.

A fent leírtak szerint, egy mátrixnak akkor és csak akkor nem nulla a determinánsa, ha oszlopvektorai lineárisan függetlenek. Ezt az észrevételünket tétel formájában is megfogalmazzuk.

7.2.1. Tétel. *Egy $\mathbf{A}\in\mathbb{R}^{n\times n}$ kvadratikus mátrixra akkor és csak akkor teljesül, hogy $\det(\mathbf{A})\neq 0$, ha $r(\mathbf{A})=n$.*

Bizonyítás. Az állítás a tétel előtti gondolatmenet közvetlen következménye. ■

7.2.2. Példa. Az

$$\mathbf{A}=\begin{bmatrix}2&-1&0&2\\1&3&1&1\\0&2&-1&-1\\4&1&2&1\end{bmatrix}$$

mátrix determinánsának meghatározásához másoljuk be a mátrixot egy bázistranszformációs táblázatba, ahol a régi bázist az $\overline{\mathbf{e}}_1,\ldots,\overline{\mathbf{e}}_4$ vektorok alkotják, majd végezzünk elemi

bázistranszformációkat bejuttatva az $\overline{\mathbf{a}}_1, \ldots, \overline{\mathbf{a}}_4$ vektorokat a bázisba (a bázisba bevitt vektorok koordinátáit nem írtuk ki).

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_2$
$\overline{\mathbf{a}}_2 \not\in \overline{\mathbf{e}}_1$	2	−1	0	2	−7	−2	0	−11	2	$\boxed{-41}$
$\overline{\mathbf{a}}_1 \not\in \overline{\mathbf{e}}_2$	$\boxed{1}$	3	1	1	3	1	1	5	0	5
$\overline{\mathbf{a}}_3 \not\in \overline{\mathbf{e}}_3$	0	2	−1	−1	2	$\boxed{-1}$	−1	−2	1	−17
$\overline{\mathbf{a}}_4 \not\in \overline{\mathbf{e}}_4$	4	1	2	1	−11	−2	−3	−15	$\boxed{-1}$	15

A generálóelemek szorzata $1 \cdot (-1) \cdot (-1) \cdot (-41) = -41$. Ahhoz, hogy megkapjuk $\det(\mathbf{A})$ értékét már csak azt kell észrevenni, hogy a táblázat első oszlopában szereplő új bázisvektorok természetes sorrendjének visszaállításához páratlan sok cserére van szükség (nevezetesen egyre, az $\overline{\mathbf{a}}_2$ és $\overline{\mathbf{a}}_1$ vektorokat kell felcserélnünk). Emiatt a generálóelemek szorzatát még (-1)-gyel be kell szoroznunk és így

$$\det(\mathbf{A}) = (-1) \cdot (-41) = 41.$$

Megjegyezzük, hogy a generálóelemek szorzatát a négy bázistranszformáció után nyert

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

mátrix determinánsával kellene beszoroznunk, amely $(-1)^{\mathcal{I}(2,1,3,4)} = (-1)^1$, azaz valóban (-1). □

7.2.3. Példa. A

$$\mathbf{B} = \begin{bmatrix} 2 & -1 & 1 & 3 \\ 1 & 3 & 4 & -2 \\ 0 & 2 & 2 & -2 \\ 4 & 1 & 5 & 3 \end{bmatrix}$$

mátrix determinánsának meghatározásához másoljuk be a mátrixot a bázistranszformációs táblázatba, és végezzünk elemi bázistranszformációkat.

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$
$\overline{\mathbf{e}}_1$	2	−1	1	3	−7	−7	7	0	0
$\overline{\mathbf{a}}_1 \not\in \overline{\mathbf{e}}_2$	$\boxed{1}$	−2	4	1	3	4	−2	1	1
$\overline{\mathbf{a}}_2 \not\in \overline{\mathbf{e}}_3$	0	2	2	−2	$\boxed{2}$	2	−2	1	−1
$\overline{\mathbf{e}}_4$	4	1	5	3	−11	−11	11	0	0

Mivel nem tudtuk $\mathbf{B}$ összes oszlopát bevinni a bázisba, így $\det(\mathbf{B}) = 0$. □

Ellenőrző kérdések

1. Milyen kapcsolat van egy mátrix determinánsa és rangja között?

2. Hogyan változik egy mátrix determinánsa egy olyan bázistranszformáció során, amikor valamelyik oszlopvektorát visszük be az $\overline{\mathbf{e}}_1, \ldots, \overline{\mathbf{e}}_n$ bázisba?

3. Hogyan lehet egy mátrix determinánsát kiszámítani bázistranszformáció segítségével?

7.3. Lineáris egyenletrendszerek megoldása

Ez a fejezet a lineáris algebra egyik klasszikus feladatával, a lineáris egyenletrendszerek megoldhatóságával és megoldási módszereivel foglalkozik. Az eddig tanultak közül felhasználjuk a lineáris kombináció, vektorrendszer rangja, mátrix rangja, elemi bázistranszformáció, és determináns fogalmakat.

Legyenek adottak az a_{ij}, b_i, $(i = 1, \ldots, k,\ j = 1, \ldots, n)$ valós számok. Az

$$\left.\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ & & & & & & & \vdots & \\ a_{k1}x_1 & + & a_{k2}x_2 & + & \ldots & + & a_{kn}x_n & = & b_k \end{array}\right\} \tag{7.3.1}$$

objektumot lineáris egyenletrendszernek hívjuk, ahol $x_1, x_2, \ldots, x_n$ ismeretlenek, vagy más néven változók. Az a_{ij}, b_i, $(i = 1, \ldots, k,\ j = 1, \ldots, n)$ számok az egyenletrendszer együtthatói. A fenti sémából jól látható, hogy egy lineáris egyenletrendszer lineáris egyenletekből áll. A "lineáris" jelző azt a tényt fejezi ki, hogy az összes ismeretlen az első hatványon szerepel az egyenletekben. Geometriai szempontból az egyes egyenletek felfoghatók egy-egy ún. hipersík egyenletének, amely $n = 2$ esetén egy egyenes, $n = 3$ esetén egy sík szokásos egyenletét adja.

Ha a (7.3.1) egyenletrendszerben az $x_1, x_2, \ldots, x_n$ ismeretlenek helyére rendre a $t_1, t_2, \ldots, t_n$ valós számokat helyettesítjük be és így k darab egyenlőséget kapunk, akkor azt mondjuk, hogy

$$\{\ x_1 = t_1\ ,\ x_2 = t_2\ ,\ \ldots\ ,\ x_n = t_n\ \} \tag{7.3.2}$$

egy megoldása a fenti egyenletrendszernek ((7.3.2) tehát egyszerre megoldása (7.3.1) összes egyenletének). Ebben a fejezetben azt vizsgáljuk, hogyan lehet meghatározni egy adott lineáris egyenletrendszer esetén az összes megoldást. Mielőtt erre rátérnénk, még két különböző formába átfogalmazzuk a vizsgált problémát.

Bevezetve az

$$\overline{\mathbf{a}}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \end{bmatrix},\ \overline{\mathbf{a}}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{bmatrix},\ \ldots,\ \overline{\mathbf{a}}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{kn} \end{bmatrix},\ \overline{\mathbf{b}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

vektorokat, a (7.3.1) egyenletrendszer felírható

$$\overline{\mathbf{a}}_1 x_1 + \overline{\mathbf{a}}_2 x_2 + \ldots + \overline{\mathbf{a}}_n x_n = \overline{\mathbf{b}}$$

alakban is. Eszerint egy lineáris egyenletrendszert megoldani annyit jelent, mint megkeresni az adott $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n$ vektorok összes olyan lineáris kombinációját, amelyek a $\overline{\mathbf{b}}$ vektort állítják elő (lásd 6.2.6. feladat).

Az

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{bmatrix}$$

együtthatómátrixot felhasználva (7.3.1) az

$$\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}} \tag{7.3.3}$$

formát ölti, ahol az $\overline{\mathbf{x}}$ vektorváltozó rendre az $x_1, x_2, \ldots, x_n$ ismeretleneket gyűjti össze: $\overline{\mathbf{x}} = [x_1, x_2, \ldots, x_n]^\top$. Most úgy fogalmazhatjuk meg a problémát, hogy keressük az összes olyan $\overline{\mathbf{x}} \in \mathbb{R}^n$ vektort, amely kielégíti a (7.3.3) vektoregyenletet.

A lineáris egyenletrendszerek most felírt három alakja egymással ekvivalens, bármelyikből bármelyik könnyedén előállítható. A későbbiekben gyakran csak a legkényelmesebbnek tartott alakra fogunk hivatkozni.

Ha $\overline{\mathbf{b}}$ nullvektor, akkor az egyenletrendszert homogénnek, egyébként inhomogénnek nevezzük. Világos, hogy homogén esetben van az egyenletrendszernek megoldása, hiszen $\overline{\mathbf{x}} = \overline{\mathbf{o}}$ kielégíti a (7.3.3) egyenletet, mert $\mathbf{A}\overline{\mathbf{o}} = \overline{\mathbf{o}} = \overline{\mathbf{b}}$. Most tekintsünk három egyszerű példát, melyek az egyenletrendszerek megoldásának három fő esetét szemléltetik.

7.3.1. Példa. Az

$$\left.\begin{array}{rcrcl} x_1 & + & x_2 & = & 0 \\ x_1 & + & x_2 & = & 1 \end{array}\right\}$$

inhomogén lineáris egyenletrendszernek nincs megoldása, ugyanis az első és második egyenlet ellentmond egymásnak. □

7.3.2. Példa. Az

$$\left.\begin{array}{rcrcl} x_1 & + & x_2 & = & 0 \\ 2x_1 & + & 2x_2 & = & 0 \end{array}\right\} \tag{7.3.4}$$

homogén lineáris egyenletrendszernek nyilvánvalóan megoldása $\{\, x_1 = 0\,,\ x_2 = 0\,\}$, de ezen kívül további megoldások is fellépnek. Ennek az az oka, hogy az első egyenletet kettővel megszorozva a másodikat kapjuk, azaz a második egyenlet nem ad új információt az első után, akár el is hagyhatjuk azt a rendszerből. Emiatt a (7.3.4) rendszernek végtelen sok megoldása lesz, mert az $x_1 + x_2 = 0$ egyenletben az egyik változó (pl. x_1) szabadon választható, s ezután a másik a tetszőlegesen választottal írható fel ($x_2 = -x_1$). □

7.3.3. Példa. Az

$$\left.\begin{array}{rcrcl} x_1 & + & x_2 & = & -2 \\ 2x_1 & + & x_2 & = & 11 \end{array}\right\}$$

inhomogén lineáris egyenletrendszernek $\{\, x_1 = 13\,,\ x_2 = -15\,\}$ az egyetlen megoldása. □

A lineáris egyenletrendszerek vizsgálata során az alábbi kérdéskörökkel foglalkozunk.

1. Megoldhatóság (hogyan dönthető el egy egyenletrendszerről, hogy megoldható vagy nem).
2. Megoldások száma (hány megoldása van az egyenletrendszernek).
3. Megoldások áttekinthető leírása.
4. Megoldási módszerek (milyen algoritmusokat alkalmazva kapjuk meg az összes megoldást).

Ezután a most felsoroltakat fejtjük ki részletesen.

7.3.1. Lineáris egyenletrendszerek megoldhatósága

A következő részben azt vizsgáljuk, hogy a lineáris egyenletrendszer együtthatóiból hogyan dönthető el annak megoldhatósága ill. nem megoldhatósága. Az

$$\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}} \tag{7.3.5}$$

egyenletrendszerben $\mathbf{A}$ az együtthatómátrix, legyen továbbá

$$\mathbf{A}_b = \left[\mathbf{A} \mid \overline{\mathbf{b}}\right]$$

az ún. kibővített mátrix. Egy korábbi átfogalmazás szerint a (7.3.5) lineáris egyenletrendszer megoldása a $\overline{\mathbf{b}}$ vektor

$$\overline{\mathbf{b}} = x_1\overline{\mathbf{a}}_1 + x_2\overline{\mathbf{a}}_2 + \ldots + x_n\overline{\mathbf{a}}_n$$

alakú lineáris kombinációinak megkeresését jelenti, ahol $\overline{\mathbf{a}}_i$ $(i = 1, 2, \ldots, n)$ az $\mathbf{A}$ mátrix oszlopvektora. A most említett fogalmak segítségével megfogalmazzuk a megoldhatóság szükséges és elégséges feltételét.

7.3.4. Tétel. *Az* $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$ *lineáris egyenletrendszer akkor és csak akkor megoldható ha*

$$r(\mathbf{A}) = r(\mathbf{A}_b). \tag{7.3.6}$$

Bizonyítás. Az $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$ egyenletrendszer akkor és csak akkor megoldható, ha $\overline{\mathbf{b}} \in L(\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n)$, ami pontosan azt jelenti, hogy

$$\dim\left(L(\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n)\right) = \dim\left(L(\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n, \overline{\mathbf{b}})\right). \tag{7.3.7}$$

Mivel (7.3.7) ekvivalens az $r(\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n) = r(\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n, \overline{\mathbf{b}})$ egyenlőséggel, így $r(\mathbf{A}) = r(\mathbf{A}_b)$. Ezzel beláttuk az állítást. ■

> 7.3.1. Feladat. Mutassuk meg az előző tétel segítségével, hogy homogén egyenletrendszernek van megoldása.

7.3.5. Példa.

$$\left.\begin{array}{rcrcrcr} x_1 & + & x_2 & + & 3x_3 & = & 4 \\ 5x_1 & + & 6x_2 & + & 7x_3 & = & 8 \end{array}\right\} \tag{7.3.8}$$

megoldható, mert $r(\mathbf{A}) = r(\mathbf{A}_b) = 2$. □

7.3.6. Példa.

$$\left.\begin{array}{rcrcrcr} -x_1 & - & 2x_2 & - & 3x_3 & = & -4 \\ 5x_1 & + & 6x_2 & + & 7x_3 & = & 8 \\ 6x_1 & + & 8x_2 & + & 10x_3 & = & 0 \end{array}\right\} \tag{7.3.9}$$

nem megoldható, mert $r(\mathbf{A}) = 2 < r(\mathbf{A}_b) = 3$. □

7.3.7. Példa.

$$\left.\begin{array}{rcrcr} x & + & y & = & 1 \\ x & - & 2y & = & 11 \end{array}\right\} \tag{7.3.10}$$

szintén megoldható, mert $r(\mathbf{A}) = r(\mathbf{A}_b) = 2$. □

Bár (7.3.8) és (7.3.10) mindegyikének van megoldása, a következő fejezetben jelentős különbség mutatkozik majd a két egyenletrendszer között. Az utóbbi lineáris egyenletrendszernél az ismeretleneket a középiskolában megszokott hagyományos módon jelöltük. Ez mindig megtehető, de sokszor előnyösebb az eddig használt indexelt forma.

7.3.2. Lineáris egyenletrendszerek megoldásainak száma

Az előző fejezet jelöléseit használva, tekintsük az

$$\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$$

egyenletrendszert, ahol az $\overline{\mathbf{x}}$ vektor n ismeretlent tartalmaz. Ha nem megoldható, akkor a megoldások száma 0. Ha megoldható, akkor két eset lehetséges, amelyekről a következő tétel ad felvilágosítást.

7.3.8. Tétel. *Ha az $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$ lineáris egyenletrendszer megoldható ($r(\mathbf{A}) = r(\mathbf{A}_b)$), akkor*

- *$r(\mathbf{A}) = n$ esetén 1 megoldás van, azaz a megoldás egyértelmű;*
- *$r(\mathbf{A}) < n$ esetén végtelen sok megoldás van, amelyek $s = n - r(\mathbf{A})$ szabad paraméterrel jellemezhetők.*

Bizonyítás. Mivel az egyenletrendszer megoldható, így $\overline{\mathbf{b}} \in L(\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n)$. Jelölje az $\mathbf{A}$ mátrix rangját s.

Ha $s = n$ akkor az $\mathbf{A}$ mátrix $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n$ oszlopvektorai lineárisan függetlenek, ezért a 6.3.3. tétel szerint $\overline{\mathbf{b}}$ csak egyféleképpen fejezhető ki az $\mathbf{A}$ mátrix oszlopvektorainak lineáris kombinációjával.

Ha $s < n$ akkor $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n$ lineárisan összefüggők, és a 6.3.4. tétel alapján $\overline{\mathbf{b}}$ végtelen sokféleképpen írható fel $\mathbf{A}$ oszlopvektorainak lineáris kombinációjaként. ■

Az előző rész példáit végignézve, a (7.3.8) egyenletrendszernek végtelen sok megoldása van egy szabad paraméterrel, mert $n = 3$ ismeretlen van, miközben $r(\mathbf{A}) = 2$. (7.3.9) megoldásainak száma nulla, (7.3.10) megoldása egyértelmű $n = r(\mathbf{A}) = 2$ miatt.

7.3.3. Lineáris egyenletrendszerek megoldásainak leírása

Legegyszerűbb a helyzet, ha az

$$\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$$

egyenletrendszer nem oldható meg, ekkor azt lehet írni, hogy "*nincs megoldás*".

Amennyiben a megoldás egyértelmű, akkor az $x_1, x_2, \ldots, x_n$ ismeretlenek által rendre felvett $t_1, t_2, \ldots, t_n$ értékek felsorolásával lehet a megoldást lejegyezni:

$$\{\, x_1 = t_1 \,,\; x_2 = t_2 \,,\; \ldots \,,\; x_n = t_n \,\}\,.$$

Végtelen sok megoldás esetén $s = n - r(\mathbf{A})$ szabad változó van, legyenek ezek – az általánosság megszorítása nélkül – rendre $x_1, x_2, \ldots, x_s$. A most felsorolt s változó tehát tetszőleges értékeket vehet fel, míg az $x_{s+1}, x_{s+2}, \ldots, x_n$ ismeretlenek kötöttek, amelyek a szabad változók $\mathcal{L}$ lineáris függvényeiként adhatók meg. Ezek szerint a megoldások áttekinthető leírása a következőképpen alakul.

- *szabad változók:* $x_1, x_2, \ldots, x_s$;
- *kötött változók:*

$$\begin{aligned} x_{s+1} &= \mathcal{L}_{s+1}(x_1, x_2, \ldots, x_s), \\ x_{s+2} &= \mathcal{L}_{s+2}(x_1, x_2, \ldots, x_s), \\ &\vdots \\ x_n &= \mathcal{L}_n(x_1, x_2, \ldots, x_s). \end{aligned}$$

Megjegyezzük, hogy általában a szabad változók rendszere nem egyértelmű, abba az ismeretlenek többféleképpen is beválaszthatók. Speciális esetben előfordulhat, hogy egy kötött változó nulla lesz, mert az őt leíró lineáris függvény együtthatói nullák.

7.3.9. Példa. Szemléltessük az elhangzottakat a (7.3.8) egyenletrendszer végtelen sok megoldásának leírásával. $s = 3 - 2 = 1$ szabad paraméter lesz, a később ismertetendő megoldási módszer szerint ez lehet pl. x_3. Szintén a megoldási algoritmus végeztével derül ki, hogy az említett lineáris függvények a következők: $\mathcal{L}_1(x_3) = -2+x_3$, $\mathcal{L}_2(x_3) = 3-2x_3$. Tehát

- *szabad változó:* x_3;
- *kötött változók:*

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

□

7.3.4. Lineáris egyenletrendszerek megoldási módszerei

Jelen jegyzet felépítéséhez leginkább a lineáris egyenletrendszerek elemi bázistranszformációval történő megoldása illeszkedik, ezt fogjuk részletesen tárgyalni. Ennek előnye, hogy viszonylag egyszerű és konstruktív módszer. A konstruktivitása abban áll, hogy az eljárás végeztével megállapítható a megoldhatóság vagy éppen a nem megoldhatóság, leolvasható a megoldások száma és maguk a megoldások is. További előny, hogy az algoritmus egyszerűen programozható. Másik hasonlóan jól alkalmazható (általunk nem vizsgált) eljárás az ún. Gauss-elimináció.

Mindkét módszer hátránya, hogy az egyenletek és ismeretlenek számának növekedésével – a kerekítési hibák miatt – az eredmények megbízhatósága csökken. Vannak azonban különböző iterációs (közelítő) módszerek, melyek kiküszöbölik a most említett problémákat, ám ezek számítógépek használatára épülnek.

A fejezet végén ismertetjük a Cramer-szabályt is, amely csak abban a speciális esetben használható, ha az ismeretlenek és egyenletek száma megegyezik, továbbá ha az együtthatómátrix determinánsa nem 0. Ilyen feltételek mellett az egyenletrendszernek egyértelmű megoldása lesz, amely éppen a Cramer-szabállyal írható fel.

Lineáris egyenletrendszerek megoldása elemi bázistranszformációval

Tekintsünk egy tetszőleges lineáris egyenletrendszert az

$$x_1\overline{\mathbf{a}}_1 + x_2\overline{\mathbf{a}}_2 + \ldots + x_n\overline{\mathbf{a}}_n = \overline{\mathbf{b}}$$

alakban ($\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n, \overline{\mathbf{b}} \in \mathbb{R}^k$). Célunk az, hogy elemi bázistranszformációkat végrehajtva meghatározzuk az összes megoldást. Keressük tehát az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n$ vektorok összes olyan lineáris kombinációját, melyek a $\overline{\mathbf{b}}$ vektort állítják elő.

Korábban már láttuk milyen előnyös, ha az $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \ldots, \overline{\mathbf{e}}_k$ bázisát tekintjük az $\mathbb{R}^k$ vektortérnek. Az elemi bázistranszformációs módszer lényege az, hogy az eredeti $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \ldots, \overline{\mathbf{e}}_k$ vektorok közül a lehető legtöbbet kicserélünk az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n$ vektorokra. Hogy valamely $\overline{\mathbf{a}}_i$ vektort melyik $\overline{\mathbf{e}}_j$ helyére visszük be – ha erre egyáltalán lehetőség van – azt a kényelmi vagy egyéb szempontok határozzák meg. Ha elfogytak az aktuális bázisba bevihető $\overline{\mathbf{a}}_i$

vektorok, akkor vége az eljárásnak, ezután kezdődhet a megoldás(ok) leolvasása az elemi bázistranszformáció táblázatából. Az algoritmus végrehajtása során a $\overline{\mathbf{b}}$ vektor koordinátái állandóan változnak, hiszen minden lépésben más és más bázisban kerül felírásra $\overline{\mathbf{b}}$. Végül kideríthető, hogy $\overline{\mathbf{b}}$ valóban kifejezhető-e az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \ldots, \overline{\mathbf{a}}_n$ vektorok lineáris kombinációjaként, vagy szükség van ehhez bizonyos $\overline{\mathbf{e}}_j$ vektor(ok)ra is. Vegyük észre, hogy ha már nem tudunk újabb $\overline{\mathbf{a}}_i$ vektort becserélni egyik $\overline{\mathbf{e}}_j$ helyére sem, de maradt még a bázisban legalább egy $\overline{\mathbf{e}}_j$ vektor, ez nem jelenti automatikusan az egyenletrendszer megoldhatatlanságát. Előfordulhat ugyanis, hogy $\overline{\mathbf{b}}$-nek a bázisban bennmaradt összes $\overline{\mathbf{e}}_j$ vektorhoz tartozó koordinátája 0, tehát formálisan ugyan szükség van $\overline{\mathbf{b}}$ előállításához a bázisban bennmaradt $\overline{\mathbf{e}}_j$ vektorokra is, de a 0 együttható miatt azok elhagyhatók $\overline{\mathbf{b}}$ felírásából. Tekintsünk most néhány példát a módszer szemléltetésére.

7.3.10. Példa. Írjuk fel a

$$\left.\begin{array}{rcrcrcr} -2x_1 & + & -2x_2 & + & x_3 & = & 3 \\ x_1 & + & 2x_2 & + & 3x_3 & = & -1 \\ 2x_1 & + & 3x_2 & + & 5x_3 & = & 0 \end{array}\right\} \tag{7.3.11}$$

egyenletrendszert

$$x_1 \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} \tag{7.3.12}$$

alakban. Jelölje rendre $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3, \overline{\mathbf{b}}$ a felírásban szereplő négy darab háromdimenziós vektort. A most következő táblázat első része e négy vektor $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \overline{\mathbf{e}}_3$ kiindulási bázisvektorokra vonatkozó koordinátáit tartalmazza. Ha eltekintünk a matematikai háttértől, akkor úgy is fogalmazhatunk, hogy egyszerűen csak bemásoltuk az adott egyenletrendszer együtthatóit a táblázatba.

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}$	$\overline{\mathbf{b}}$
$\overline{\mathbf{a}}_3 \not\in \overline{\mathbf{e}}_1$	-2	-2	1	3	2	7	1	5	5	1
$\overline{\mathbf{a}}_1 \not\in \overline{\mathbf{e}}_2$	1	2	3	-1	2	3	-1	1	3	2
$\overline{\mathbf{a}}_2 \not\in \overline{\mathbf{e}}_3$	2	3	5	0	-1	-1	2	1	-2	-3

Megvizsgálva a táblázat kezdő bázisvektorok utáni első részét látható, hogy az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3$ vektorok bármelyike becserélhető az $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \overline{\mathbf{e}}_3$ bármelyike helyére, mert az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3$ vektorok egyik koordinátája sem zérus. (Elvileg $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2$ bármelyikét lecserélhetnénk a $\overline{\mathbf{b}}$ vektorral is, de ez nem vinné előbbre az egyenletrendszer megoldását, hiszen éppen $\overline{\mathbf{b}}$ koordinátáit szeretnénk meghatározni.) Először – a kényelmi szempontokat figyelembe véve az 1 generálóelemet választva – vigyük be az $\overline{\mathbf{a}}_1$ vektort a régi bázis $\overline{\mathbf{e}}_2$ vektora helyére. A táblázat második részében megtaláljuk, hogy a végrehajtás utáni $\overline{\mathbf{e}}_1, \overline{\mathbf{a}}_1, \overline{\mathbf{e}}_3$ bázisban hogyan alakulnak a többi vektor koordinátái. Ezek egyike sem nulla, így szabad kezünk van a második generálóelem kiválasztásában. Vegyük azonban észre, hogy nem lenne célszerű a második sorban sem a 2-t, sem a 3-at bekeretezni, mert ezzel éppen az előbb bevitt $\overline{\mathbf{a}}_1$ vektort távolítanánk el a bázisból. Válasszuk második generálóelemnek -1-et, tehát az új aktuális bázis $\overline{\mathbf{e}}_1, \overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2$ lesz, az erre vonatkozó új koordináták a táblázat harmadik részében vannak. Most egyedül az 5 generálóelem választása célszerű, mert így lehet csak megszabadulni az utolsó régi bázisvektortól, $\overline{\mathbf{e}}_1$-től, ezzel az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3$ vektorok mindegyike bejutott a bázisba. Az algoritmus véget ért, az eredmény leolvasása következik.

Rögtön megállapíthatjuk, hogy a 3×3-as együtthatómátrix rangja 3, így a kibővített mátrixé is, amely megegyezik az ismeretlenek számával. Az egyenletrendszerek megoldhatóságára ill. megoldásszámára vonatkozó tételek alapján (7.3.11) egyértelműen megoldható. Felírásához részben a táblázat utolsó oszlopát használjuk, amely az elemi bázistranszformációk végrehajtása utáni bázisvektorokat tartalmazza a megfelelő sorrendben. Az előtte levő oszlopban (a táblázat eddigi logikája szerinti negyedik részében) éppen a $\overline{\mathbf{b}}$ vektor végső bázisvektorokra vonatkozó koordinátái találhatók. Azaz

$$\overline{\mathbf{b}} = 2\overline{\mathbf{a}}_1 - 3\overline{\mathbf{a}}_2 + \overline{\mathbf{a}}_3,$$

és ez az előállítás egyértelmű. Mivel x_1, x_2, x_3 rendre az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3$ vektorok együtthatói, ezért a (7.3.11) lineáris egyenletrendszer egyetlen megoldása:

$$\{\, x_1 = 2\ ,\ x_2 = -3\ ,\ x_3 = 1\,\}\,.$$

Geometriai szemszögből azt állapíthatjuk meg, hogy az egyenletrendszerben szereplő 3 síknak egy közös pontja van. □

A következő három példában már nem részletezzük az elemi bázistranszformációkat, hanem a megoldások táblázatból való leolvasására koncentrálunk.

7.3.11. Példa. Adott az

$$\left.\begin{array}{rcrcrcr} 2x & - & 5y & + & 3z & = & -4 \\ 3x & - & 2y & + & z & = & 3 \\ x & - & 3y & + & 5z & = & -6 \\ 4x & - & 4y & - & z & = & 5 \end{array}\right\} \tag{7.3.13}$$

egyenletrendszer. A megoldás során a

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}$	$\overline{\mathbf{b}}$
$\overline{\mathbf{a}}_2$ $\not\overline{\mathbf{e}}_1$	2	-5	3	-4	1	-7	8	-7	8	1
$\overline{\mathbf{a}}_3$ $\not\overline{\mathbf{e}}_2$	3	-2	1	3	7	-14	21	35	-35	-1
$\overline{\mathbf{a}}_1$ $\not\overline{\mathbf{e}}_3$	1	-3	5	-6	-3	5	-6	-16	18	2
$\overline{\mathbf{e}}_4$	4	-4	-1	5	8	-21	29	35	-35	0

táblázatot kapjuk. Az együtthatómátrix rangja 3, a kibővített mátrixé szintén, mivel a $\overline{\mathbf{b}}$ vektorral nem tudjuk kiváltani $\overline{\mathbf{e}}_4$-et hiszen a $\overline{\mathbf{b}}$ vektor rá vonatkozó koordinátája éppen 0:

$$\overline{\mathbf{b}} = 2\overline{\mathbf{a}}_1 + \overline{\mathbf{a}}_2 - \overline{\mathbf{a}}_3 + 0\overline{\mathbf{e}}_4 = 2\overline{\mathbf{a}}_1 + \overline{\mathbf{a}}_2 - \overline{\mathbf{a}}_3.$$

Az ismeretlenek száma is 3, ezért ismét egyértelmű a megoldás:

$$\{\, x = 2\ ,\ y = 1\ ,\ z = -1\,\}\,. \tag{7.3.14}$$

A kapott eredmény úgy is felfogható, hogy (7.3.13) négy síkjának egy közös pontja van. □

7.3.12. Példa. Az

$$\left.\begin{array}{rcrcrcr} x_1 & + & -2x_2 & + & x_3 & = & 1 \\ -2x_1 & + & x_2 & + & x_3 & = & 3 \\ x_1 & + & x_2 & + & -2x_3 & = & 1 \end{array}\right\}$$

egyenletrendszer az alábbi táblázatra vezet.

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}$
$\overline{\mathbf{a}}_1 \not\in \overline{\mathbf{e}}_1$	[1]	-2	1	1	-2	1	1	-1	$-7/3$
$\overline{\mathbf{a}}_2 \not\in \overline{\mathbf{e}}_2$	-2	1	1	3	[-3]	3	5	-1	$-5/3$
$\overline{\mathbf{e}}_3$	1	1	-2	1	3	-3	0	0	5

Mint már korábban említettük, a generálóelemek választása önkényes, igyekszünk azokat a számolás kényelmessé tétele érdekében 1-nek vagy -1-nek választani.

Az eljárás során két elemi bázistranszformációt tudtunk végrehajtani. Tehár $r(\mathbf{A}) = 2$, az $\overline{\mathbf{a}}_3$ vektor már nem vihető be $\overline{\mathbf{e}}_3$ helyére, mert a generálóelem nem lehet nulla, ellenben $\overline{\mathbf{b}}$ bevihető, így $r(\mathbf{A}_b) = 3 > r(\mathbf{A})$. Eszerint az egyenletrendszernek nincs megoldása. Megjegyezzük, hogy ha más generálóelemeket választottunk volna az algoritmus során, akkor is csak 2 elemi bázitranszformációt tudtunk volna végrehajtani, hiszen az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3$ vektorrendszer rangja 2. □

7.3.13. Példa. Végül oldjuk meg a

$$\left.\begin{array}{rcrcrcrcrcr} 2x_1 & - & 2x_2 & + & x_3 & - & x_4 & + & x_5 & = & 1 \\ x_1 & + & 2x_2 & - & x_3 & + & x_4 & - & 2x_5 & = & 1 \\ 4x_1 & - & 10x_2 & + & 5x_3 & - & 5x_4 & + & 7x_5 & = & 1 \\ 2x_1 & - & 14x_2 & + & 7x_3 & - & 7x_4 & + & 11x_5 & = & -1 \end{array}\right\}$$

lineáris egyenletrendszert. Az $\overline{\mathbf{a}}_i$ $(i = 1, \ldots, 5)$ vektorok közül mindössze kettőt lehet becserélni a bázisba. A számításokat a következő táblázat mutatja.

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_5$	$\overline{\mathbf{b}}$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_5$	$\overline{\mathbf{b}}$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{a}}_5$	$\overline{\mathbf{b}}$
$\overline{\mathbf{a}}_3 \not\in \overline{\mathbf{e}}_1$	2	-2	1	-1	1	1	-6	[3]	-3	5	-1	-2	-1	$5/3$	$-1/3$
$\overline{\mathbf{a}}_1 \not\in \overline{\mathbf{e}}_2$	[1]	2	-1	1	-2	1	2	-1	1	-2	1	0	0	$-1/3$	$2/3$
$\overline{\mathbf{e}}_3$	4	-10	5	-5	7	1	-18	9	-9	15	-3	0	0	0	0
$\overline{\mathbf{e}}_4$	2	-14	7	-7	11	-1	-18	9	-9	15	-3	0	0	0	0

A táblázat utolsó részét megvizsgálva megállapíthatjuk, hogy $r(\mathbf{A}) = r(\mathbf{A}_b) = 2 < 5 = n$ mert két új és két régi bázisvektor van a végső bázisban, de a $\overline{\mathbf{b}}$ vektor előállításához nincs szükség a régiekre. Eszerint az egyenletrendszer megoldható, végtelen sok megoldása van, x_2, x_4, x_5 lesznek a szabad változók. Nehezebb kérdés az, hogy hogyan lehet megadni a kötött változókat leíró lineáris függvényeket. Induljunk ki a $\overline{\mathbf{b}}$ vektor végső bázisban történő felírásából. A táblázat szerint

$$\overline{\mathbf{b}} = \frac{2}{3}\overline{\mathbf{a}}_1 - \frac{1}{3}\overline{\mathbf{a}}_3. \tag{7.3.15}$$

Vizsgáljuk meg, hogy hogyan lehetne bevonni az $\overline{\mathbf{a}}_2, \overline{\mathbf{a}}_4, \overline{\mathbf{a}}_5$ vektorokat $\overline{\mathbf{b}}$ lineáris kombinációként való előállításába. Ehhez szükségünk lesz a táblázatból kiolvasható

$$\overline{\mathbf{a}}_2 = -2\overline{\mathbf{a}}_3 \quad , \quad \overline{\mathbf{a}}_4 = -\overline{\mathbf{a}}_3 \quad , \quad \overline{\mathbf{a}}_5 = -\frac{1}{3}\overline{\mathbf{a}}_1 + \frac{5}{3}\overline{\mathbf{a}}_3 \tag{7.3.16}$$

lineáris kombinációkra is. Mivel a szabad x_2, x_4, x_5 változók rendre $\overline{\mathbf{a}}_2, \overline{\mathbf{a}}_4, \overline{\mathbf{a}}_5$ együtthatói, meg kellene határozni, hogyan válasszuk meg az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_3$ vektorok x_1, x_3 együtthatóit, hogy a

$$\overline{\mathbf{b}} = (x_2\overline{\mathbf{a}}_2 + x_4\overline{\mathbf{a}}_4 + x_5\overline{\mathbf{a}}_5) + (x_1\overline{\mathbf{a}}_1 + x_3\overline{\mathbf{a}}_3) \tag{7.3.17}$$

ekvivalens legyen a (7.3.15) felírással. Helyettesítsük (7.3.16) alapján a (7.3.17)-beli $\overline{\mathbf{a}}_2, \overline{\mathbf{a}}_4, \overline{\mathbf{a}}_5$ vektorokat. Ezután összevonva a megfelelő együtthatókat, a

$$\overline{\mathbf{b}} = \left(-\frac{1}{3}x_5 + x_1\right)\overline{\mathbf{a}}_1 + \left(-2x_2 - x_4 + \frac{5}{3}x_5 + x_3\right)\overline{\mathbf{a}}_3$$

egyenlőséghez jutunk. A kapott eredményt összehasonlítva $\overline{\mathbf{b}}$ táblázatból kiolvasott alakjával

$$\frac{2}{3} = -\frac{1}{3}x_5 + x_1 \quad , \quad -\frac{1}{3} = -2x_2 - x_4 + \frac{5}{3}x_5 + x_3$$

adódik, ahonnan a kötött változókat kifejezve megkapjuk a kívánt

$$x_1 = \frac{2}{3} + \frac{1}{3}x_5 \quad , \quad x_3 = -\frac{1}{3} + 2x_2 + x_4 - \frac{5}{3}x_5 \tag{7.3.18}$$

kifejezéseket. Ha visszatekintünk a táblázat utolsó részében szereplő koordinátákra, rögtön szembetűnik, hogy pusztán a szabad változókhoz tartozó $\overline{\mathbf{a}}_i$ vektorok koordinátáinak (-1)-szeresével, valamint $\overline{\mathbf{b}}$ koordinátáival azonnal felírható (7.3.18). A végeredményt összefoglalva az alábbiakat mondhatjuk. Az egyenletrendszernek végtelen sok megoldása van, amelyek az alábbiak szerint karakterizálhatók.

- *szabad változók:* x_2, x_4, x_5;
- *kötött változók:*

$$\begin{aligned} x_1 &= \frac{2}{3} + \frac{1}{3}x_5, \\ x_3 &= -\frac{1}{3} + 2x_2 + x_4 - \frac{5}{3}x_5. \end{aligned}$$

□

A Cramer-szabály

A lineáris egyenletrendszerek egyik fontos speciális esete ha az egyenletek és az ismeretlenek száma megegyezik. Ez azt jelenti, hogy az

$$\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}} \tag{7.3.19}$$

egyenletrendszer $\mathbf{A} \in \mathbb{R}^{n\times n}$ együtthatómátrixa négyzetes. Ha még azt is feltesszük, hogy $\det(\mathbf{A}) \neq 0$ akkor (7.3.19) egyértelműen megoldható, továbbá a megoldások felírhatóak az ún. Cramer-szabállyal Mielőtt rátérnénk ennek kifejtésére, vezessünk be néhány jelölést. Legyen $D = \det(\mathbf{A})$, melyet a rendszer determinánsának nevezünk, továbbá D_i $(i = 1, 2, \ldots, n)$ legyen annak a mátrixnak a determinánsa, amely annyiban különbözik az $\mathbf{A}$ mátrixtól, hogy abban az i-edik oszlopot a $\overline{\mathbf{b}}$ vektorral helyettesítettük. Pl.

$$D_1 = \begin{vmatrix} b_1 & a_{12} & \ldots & a_{1n} \\ b_2 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

7.3.14. Tétel. *(Cramer-szabály) Ha az* $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$ *lineáris egyenletrendszerre* $D = \det(\mathbf{A}) \neq 0$ *(*$\mathbf{A} \in \mathbb{R}^{n\times n}$*), akkor az egyenletrendszer egyértelműen megoldható és a megoldásokra*

$$\left\{ x_1 = \frac{D_1}{D} \ , \ x_2 = \frac{D_2}{D} \ , \ \ldots \ , \ x_n = \frac{D_n}{D} \right\}$$

teljesül.

A tételt nem bizonyítjuk. A Cramer-szabály tehát csak speciális esetekben alkalmazható, és akkor is általában sokkal több a számolásigénye, mint az elemi bázistranszformációs módszernek. Ezért a most kimondott tétel inkább elméleti jelentőségű. Persze érdemes a Cramer-szabályt használni, ha a benne szereplő determinánsok könnyen meghatározhatók.

7.3.15. Példa. Tekintsük most a két egyenletből és ismeretlenből álló

$$\left.\begin{array}{rcrcl} 7x_1 & + & 2x_2 & = & 13 \\ -x_1 & + & 5x_2 & = & 14 \end{array}\right\}$$

lineáris egyenletrendszert. A rendszer determinánsa

$$D = \begin{vmatrix} 7 & 2 \\ -1 & 5 \end{vmatrix} = 37 \neq 0,$$

tehát alkalmazható a Cramer-szabály. Az egyetlen megoldás

$$\left\{ x_1 = \frac{\begin{vmatrix} 13 & 2 \\ 14 & 5 \end{vmatrix}}{37} = \frac{37}{37} = 1 \ , \ x_2 = \frac{\begin{vmatrix} 7 & 13 \\ -1 & 14 \end{vmatrix}}{37} = \frac{111}{37} = 3 \right\}.$$

□

7.3.5. Homogén lineáris egyenletrendszerek

Elevenítsük fel az alcímben levő fogalmat! Az $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$ lineáris egyenletrendszert homogénnek hívjuk, ha $\overline{\mathbf{b}} = \overline{\mathbf{o}}$, ahol az $\mathbf{A} \in \mathbb{R}^{k\times n}$ mátrix és $\overline{\mathbf{o}} \in \mathbb{R}^k$ nullvektor adottak, $\overline{\mathbf{x}} = [x_1, x_2, \ldots, x_n]^\top \in \mathbb{R}^n$ az ismeretleneket tartalmazó vektor.

Egy homogén lineáris egyenletrendszernek mindig van megoldása, az ún. $\overline{\mathbf{x}} = \overline{\mathbf{o}} \in \mathbb{R}^n$ triviális megoldás. Eszerint homogén egyenletrendszereknél – a megoldások számát tekintve – csak két eset fordulhat elő: vagy egyértelmű a megoldás, vagy végtelen sok megoldás van. Egyértelmű megoldás esetén az csak a triviális megoldás lehet.

7.3.16. Tétel. *A fenti jelöléseket felhasználva, az $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{o}}$ homogén lineáris egyenletrendszernek végtelen sok megoldása van*

1. *ha $k < n$;*
2. *$k = n$ esetén akkor és csak akkor, ha $\det(\mathbf{A}) = 0$.*

Bizonyítás. A tételben szereplő egyenletrendszer nyilvánvalóan megoldható.

1. $r(\mathbf{A}) \leq \min\{k, n\} = k < n$ miatt végtelen sok megoldás lesz $n - k$ szabad paraméterrel.
2. Az állítás a feltételek és a Cramer-szabályra vonatkozó tétel (*) közvetlen következménye. ■

Ellenőrző kérdések

1. Milyen kérdések merülnek fel az egyenletrendszerek kapcsán?

2. Melyik az a mátrixokkal kapcsolatos fogalom, amely fontos szerepet játszik a megoldhatóság és a megoldások számának vizsgálatánál?

3. A lineáris egyenletrendszerek megoldásszámát tekintve milyen alapesetek fordulhatnak elő?

4. Mi az elemi bázistranszformáción alapuló eljárás lényege?

5. Milyen feltételek mellett lehet a Cramer-szabályt alkalmazni?

A feladat megoldása

7.3.1.
Az $\mathbf{A}$ mátrix rangja nem változik, ha hozzávesszük a $\overline{\mathbf{b}} = \overline{\mathbf{o}}$ oszlopvektort. Tehát $r(\mathbf{A}) = r(\mathbf{A}_b)$ és a tétel miatt az egyenletrendszer megoldható.

7.4. Mátrixok sajátértékei, sajátvektorai

Ha egy kvadratikus mátrixszal megszorzunk egy oszlopvektort, akkor a vektortér egy másik oszlopvektorához jutunk. Úgy is fogalmazhatunk, hogy a szóban forgó mátrixszorzás segítségével egy függvényt adunk meg, melynek értelmezési tartománya az oszlopvektorok egy rögzített dimenziós tere. Az alkalmazások szempontjából gyakran fontos megkeresni azokat az oszlopvektorokat, melyek képének iránya megegyezik magának a kiindulási oszlopvektornak az irányával. Most e kérdéskört fogjuk megvizsgálni.

7.4.1. Definíció. Legyen $\mathbf{A} \in \mathbb{R}^{n\times n}$ egy rögzített négyzetes mátrix. A λ valós számot az $\mathbf{A}$ mátrix **sajátértékének** nevezzük, ha létezik olyan $\overline{\mathbf{x}} \in \mathbb{R}^n$, $\overline{\mathbf{x}} \neq \overline{\mathbf{o}}$ vektor melyre

$$\mathbf{A}\overline{\mathbf{x}} = \lambda\overline{\mathbf{x}}\,. \tag{7.4.1}$$

(7.4.1) teljesülése esetén $\overline{\mathbf{x}}$-et az $\mathbf{A}$ mártix (λ sajátértékéhez tartozó) **sajátvektorának** hívjuk. $\diamondsuit$

7.4.2. Példa. Az

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

mátrixnak $\lambda = -1$ egy sajátértéke, mert pl. $\overline{\mathbf{x}} = [-2, 2]^\top \neq \overline{\mathbf{o}}$ esetén

$$\mathbf{A}\overline{\mathbf{x}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix} = (-1)\begin{bmatrix} -2 \\ 2 \end{bmatrix} = \lambda\overline{\mathbf{x}}\,.$$

Ekkor $\overline{\mathbf{x}} = [-2, 2]^\top$ az $\mathbf{A}$ mátrix (-1) sajátértékhez tartozó sajátvektora. $\square$

A következőkben azt vizsgáljuk, hogyan lehet meghatározni egy kvadratikus mátrix sajátértékeit és sajátvektorait.

Induljunk ki az $\mathbf{A}\overline{\mathbf{x}} = \lambda\overline{\mathbf{x}}$ mátrixegyenletből, ahol $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ egy rögzített négyzetes mátrix, $\lambda \in \mathbb{R}$ és $\overline{\mathbf{x}} \in \mathbb{R}^n$ ismeretlenek. Ez a mátrixegyenlet ekvivalens az

$$\left.\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & \lambda x_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & \lambda x_2 \\ & & & & & & & \vdots & \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & a_{nn}x_n & = & \lambda x_n \end{array}\right\}$$

egyenletrendszerrel, ahol $\overline{\mathbf{x}} = [x_1, x_2, \ldots, x_n]^\top$. Átrendezés után az

$$\left.\begin{array}{ccccccccc} (a_{11}-\lambda)x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & (a_{22}-\lambda)x_2 & + & \ldots & + & a_{2n}x_n & = & 0 \\ & & & & & & & \vdots & \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & (a_{nn}-\lambda)x_n & = & 0 \end{array}\right\} \tag{7.4.2}$$

homogén lineáris egyenletrendszerhez jutunk, amely n egyenletet és n ismeretlent tartalmaz. Mivel nullvektor nem lehet sajátvektor, ezért az $\mathbf{A}$ mátrixnak akkor és csak akkor van sajátvektora, ha a kapott lineáris egyenletrendszernek van a triviálistól eltérő megoldása, azaz végtelen sok megoldása van. Ez akkor és csak akkor lehetséges, ha a rendszer determinánsa, melyet röviden $\det(\mathbf{A} - \lambda\mathbf{E}_n)$ alakban írhatunk, nulla lesz. Ezzel beláttuk az alábbi tételt.

7.4.3. Tétel. *$\lambda \in \mathbb{R}$ akkor és csak akkor sajátértéke az $\mathbf{A} \in \mathbb{R}^{n\times n}$ mátrixnak, ha*

$$\det(\mathbf{A}-\lambda\mathbf{E}_n) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0\,. \tag{7.4.3}$$

Könnyű látni, hogy $\det(\mathbf{A}-\lambda\mathbf{E}_n)$ egy n-ed fokú polinomja λ-nak, tehát az előző tétel szerint az $\mathbf{A}$ mátrix sajátértékeinek meghatározásához egy n-ed fokú polinomegyenletet kell λ-ra megoldani a valós számok halmazán. Erre léteznek jó közelítő módszerek, sőt $n = 2, 3, 4$ esetén pontosan megadhatók a megoldások. Itt jegyezzük meg, hogy a (7.4.3) egyenletnek – a multiplicitást figyelembe véve – pontosan n komplex gyöke van, de ebben a fejezetben csak valós gyökökkel foglalkozunk, mert a sajátértéket valós számként definiáltuk.

Térjünk most rá a sajátvektorok megadására. Ez alatt azt kell érteni, hogy minden egyes sajátértékhez külön-külön meghatározzuk a hozzá tartozó sajátvektorokat. Tegyük fel, hogy λ az $\mathbf{A}$ mátrix egy korábban már kiszámított sajátértéke. Ekkor a (7.4.2) egyenletrendszerben λ ismert, annak nemtriviális megoldásai szolgáltatják a λ-hoz tartozó sajátvektorokat.

7.4.4. Példa. Legyen

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 4 & 1 \\ -12 & -8 & -2 \end{bmatrix}.$$

A sajátértékek meghatározásához a

$$0 = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 4 & 4-\lambda & 1 \\ -12 & -8 & -2-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 4-\lambda & 1 \\ -8 & -2-\lambda \end{vmatrix} = (1-\lambda)(\lambda^2-2\lambda)$$

harmadfokú egyenletet kell megoldani. Ezt nagyon könnyű az $(1-\lambda)(\lambda-2)\lambda = 0$ alakra hozni, ahonnan $\mathbf{A}$ sajátértékei:

$$\lambda_1 = 1\ ,\ \ \lambda_2 = 2\ ,\ \ \lambda_3 = 0\,.$$

Most a sajátértékekhez tartozó sajátvektorok meghatározása következik.

$\lambda_1 = 1$ esetén a

$$\left.\begin{array}{rcrcrcl} 0x_1 & + & 0x_2 & + & 0x_3 & = & 0 \\ 4x_1 & + & 3x_2 & + & x_3 & = & 0 \\ -12x_1 & - & 8x_2 & - & 3x_3 & = & 0 \end{array}\right\}$$

homogén lineáris egyenletrendszert kapjuk. Ebből az első egyenlet automatikusan teljesül, ezért elhagyható. A redukált egyenletrendszert megoldva végtelen sok megoldás lesz egy szabad változóval: x_1 tetszőleges. A kötött változókra $x_2 = 0$, $x_3 = -4x_1$ adódik. Tehát a $\lambda_1 = 1$ sajátértékhez tartozó sajátvektorokat

$$\overline{\mathbf{x}} = [x_1, 0, -4x_1]^\top\ ,\ \ x_1 \in \mathbb{R}\setminus\{0\}$$

alakban írhatjuk.

$\lambda_2 = 2$ esetén a

$$\left.\begin{array}{rcrcrcl} -x_1 & & & & & = & 0 \\ 4x_1 & + & 2x_2 & + & x_3 & = & 0 \\ -12x_1 & - & 8x_2 & - & 4x_3 & = & 0 \end{array}\right\}$$

rendszerhez jutunk. Az első egyenlet egyből $x_1 = 0$-t kapunk, melyet felhasználva

$$\left.\begin{array}{rcrcl} 2x_2 & + & x_3 & = & 0 \\ -8x_2 & - & 4x_3 & = & 0 \end{array}\right\}$$

következik. Vegyük észre, hogy a jelenlegi második egyenlet az első (-4)-szerese, így az is elhagyható. Végül $x_3 = -2x_2$. A keresett sajátvektorok:

$$\overline{\mathbf{x}} = [0, x_2, -2x_2]^\top , \quad x_2 \in \mathbb{R} \setminus \{0\}.$$

7.4.1. Feladat. Határozzuk meg a $\lambda_3 = 0$ sajátértékhez tartozó sajátvektorokat!

□

Ellenőrző kérdések

1. Mit értünk egy kvadratikus mátrix sajátértékén és sajátvektorán?
2. Milyen tétel képezi az alapját a sajátértékek meghatározásának?
3. Melyiket tudjuk előbb meghatározni: egy négyzetes mátrix sajátvektorait vagy sajátértékeit?
4. Van-e olyan kétszer kettes mátrix melynek nincs (valós) sajátértéke?

A feladat megoldása

7.4.1.
$\overline{\mathbf{x}} = [0, x_2, -4x_2]^\top , \quad x_2 \in \mathbb{R} \setminus \{0\}.$

7.4.2. Feladat. Az alábbi táblázatok lineáris egyenletrendszerek megoldásakor keletkeztek. Adja meg, hogy hány egyenletet és ismeretlent tartalmaznak az egyes egyenletrendszerek! Ezután írja fel a megoldásokat!

a)

	$\overline{\mathbf{a}}_4$	$\overline{\mathbf{b}}$
$\overline{\mathbf{a}}_2$	1	0
$\overline{\mathbf{a}}_3$	1	-5
$\overline{\mathbf{a}}_1$	1	2
$\overline{\mathbf{e}}_4$	0	13

b)

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_5$	$\overline{\mathbf{a}}_7$	$\overline{\mathbf{b}}$
$\overline{\mathbf{a}}_3$	-2	0	0	4
$\overline{\mathbf{a}}_6$	-1	-1	1	0
$\overline{\mathbf{a}}_4$	0	0	0	-3
$\overline{\mathbf{a}}_2$	2	3	-5	-6

c)

	$\overline{\mathbf{b}}$
$\overline{\mathbf{a}}_2$	-5
$\overline{\mathbf{a}}_3$	2
$\overline{\mathbf{a}}_1$	13
$\overline{\mathbf{e}}_3$	0

7.4.2.

Mindhárom egyenletrendszer 4 egyenletből áll, az ismeretlenek száma rendre 4, 7, 3.

Az első esetben nincs megoldás, mert a kibővített mátrix rangja 4, eggyel nagyobb mint a rendszer mátrixáé.

A második esetben végtelen sok megoldás van 3 szabad változóval: x_1, x_5, x_7. A kötött változók: $x_2 = -6 - 2x_1 - 3x_5 + 5x_7$, $x_3 = 4 + 2x_1$, $x_4 = -3$, $x_6 = x_1 + x_5 - x_7$.

Az utolsó esetben egyértelmű a megoldás, mert a rendszer mátrixának és a kibővített mátrixnak is 3 a rangja, amely megegyezik az ismeretlenek számával. a megoldás: $x_1 = 13$, $x_2 = -5$, $x_3 = 2$.

7.5. Mátrixok inverzének kiszámítása

Kvadratikus mátrixok inverzének fogalma már korábban bevezetésre került. Most szükséges és elégséges feltételt adunk létezésére, és megvizsgáljuk, hogy hogyan lehet meghatározni elemi bázistranszformáció segítségével.

Térjünk vissza egy időre a valós számok halmazához, melyben minden 0-tól különböző s számhoz egyértelműen létezik olyan t szám, hogy a kettőjük szorzata 1-et ad. A t számot s multiplikatív inverzének hívjuk. Nyilvanvalóan, $t = 1/s$, azaz t az s szám reciproka. Maga az 1 valós szám azért kitüntetett eleme $\mathbb{R}$-nek, mert bármely $s \in \mathbb{R}$ esetén $s \cdot 1 = 1 \cdot s = s$ következik. A most elmondottakra szeretnénk analógiát találni rögzített típusú kvadratikus mátrixok esetén.

Az $n \times n$-es mátrixok körében az $\mathbf{E}_n$ egységmátrix felel meg az $\mathbb{R}$-beli 1-nek, az invertálható mátrixok pedig a nullától különböző valós számoknak. Be fogjuk látni, hogy $\mathbf{A} \in \mathbb{R}^{n\times n}$ esetén $\mathbf{A}$-nak pontosan akkor létezik (multiplikatív) inverze ha $\det(\mathbf{A}) \neq 0$. Ez egyben azt is jelenti, hogy – a valós számoktól eltérően – végtelen sok $n \times n$ típusú mátrix van, amelynek nincs multiplikatív inverze. Ha egy négyzetes mátrixnak van inverze, akkor az egyértelmű. Az inverz mátrix kiszámolására az elemi bázistranszformációkon alapuló módszert fogjuk használni. Ennek az algoritmusnak az egyik előnye, hogy alkalmazásának nem előfeltétele az invertálni kívánt mátrix előzetes vizsgálata, ugyanis az eljárás során mindig kiderül, hogy $\mathbf{A}$ invertálható vagy nem. Ha igen, akkor az elemi bázistranszformáció táblázatából leolvasható az $\mathbf{A}^{-1}$ inverz.

Egy $\mathbf{A} = [a_{ij}]_{n\times n}$ mátrix inverzének megkeresése az

$$\mathbf{A}\mathbf{X} = \mathbf{E}_n \tag{7.5.1}$$

mátrixegyenlet megoldását jelenti, ahol az $\mathbf{X} = [x_{ij}]_{n\times n}$ ismeretlen mátrix összesen n^2 darab x_{ij} ismeretlen skalárt tartalmaz. Rögzítsük le a j $(1 \leq j \leq n)$ indexet. Ekkor a mátrixszorzás értelmezése alapján

$$\mathbf{A}\overline{\mathbf{x}}_j = \overline{\mathbf{e}}_j \tag{7.5.2}$$

teljesül, ahol az $\overline{\mathbf{x}}_j = [x_{1j}, x_{2j}, \ldots, x_{nj}]^\top$ vektor az $\mathbf{X}$ mátrix j-edik oszlopvektora.

Könnyű látni, hogy (7.5.2) egy n egyenletből álló n ismeretlenes egyenletrendszer, melyben $\mathbf{A}$ az együtthatómátrix és $\overline{\mathbf{x}}_j$ tartalmazza az ismeretleneket. Ha a (7.5.2) egyenletrendszert egyszerre felírjuk az összes lehetséges j $(1 \leq j \leq n)$ indexre, akkor n darab n ismeretlenes n egyenletből álló lineáris egyenletrendszert kapunk (lásd (7.5.3)), melyeknek fontos közös tulajdonságuk, hogy mindegyikük együtthatómátrixa $\mathbf{A}$. Ezen egyenletrendszerek jobb oldalán rendre az $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \ldots, \overline{\mathbf{e}}_n$ vektorok állnak. Hogy $\mathbf{A}$-nak létezzen inverze, ahhoz az

$$\begin{aligned} \mathbf{A}\overline{\mathbf{x}}_1 &= \overline{\mathbf{e}}_1 \\ \mathbf{A}\overline{\mathbf{x}}_2 &= \overline{\mathbf{e}}_2 \\ &\vdots \\ \mathbf{A}\overline{\mathbf{x}}_n &= \overline{\mathbf{e}}_n \end{aligned} \tag{7.5.3}$$

egyenletrendszerek mindegyikének megoldhatónak kell lennie. Ez azt jelenti, hogy $\mathbf{A}$ oszlopvektorai által generált térben benne kell lennie az $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \ldots, \overline{\mathbf{e}}_n$ vektorok mindegyikének. Következésképpen $\mathbf{A}$ oszlopvektorai $\mathbb{R}^n$-t generálják (hiszen $L(\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \ldots, \overline{\mathbf{e}}_n) = \mathbb{R}^n$), azaz $\mathbf{A}$ n darab oszlopvektora lineárisan független. Ebből következik, hogy $\det(\mathbf{A}) \neq 0$ (lásd 7.2.1. tétel).

Ha most feltesszük, hogy $\det(\mathbf{A}) \neq 0$, tehát $\mathbf{A}$ oszlopvektorai lineárisan függetlenek, akkor $\mathbf{A}$ oszlopvektorai $\mathbb{R}^n$ egy bázisát adják, és lineáris kombinációjukként egyértelműen felírhatók az $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \ldots, \overline{\mathbf{e}}_n$ vektorok.

7.5.1. Tétel. *Az $\mathbf{A} \in \mathbb{R}^{n\times n}$ mátrix akkor és csak akkor invertálható, ha* $\det(\mathbf{A}) \neq 0$.

Bizonyítás. A bizonyítás a tétel megfogalmazása előtt található. ■

7.5.1. Feladat. Vizsgáljuk meg, hogy az alábbi mátrixoknak van-e inverze?

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad , \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}. \tag{7.5.4}$$

A 7.2.1. tétel alapján az inverz létezésének szükséges és elégséges feltételét a következő módon is megfogalmazhatjuk. Egy $\mathbf{A} \in \mathbb{R}^{n\times n}$ négyzetes mátrix akkor és csak akkor invertálható, ha rangja n.

A fenti bizonyítás egyben útmutatást is ad az inverz mátrix meghatározására. Összesen n darab lineáris egyenletrendszert kell megoldani, de az együtthatómátrix minden egyenletrendszer esetén ugyanaz. Ez a tény lehetőséget nyújt az egyenletrendszerek párhuzamos, egyszerre történő megoldására elemi bázistranszformációval.

Ha az $\mathbf{A}$ mátrix összes oszlopvektorát sikerül bevinni a bázisba az $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \ldots, \overline{\mathbf{e}}_n$ kiindulási bázisvektorok helyére, akkor $\mathbf{A}$-nak van inverze, amely a táblázatból leolvasható. Ennek a technikáját már az egyenletrendszerek megoldása során ismertettük. Amennyiben az $\mathbf{A}$ mátrixnak marad(nak) oszlopvektora(i) amely(ek) már nem vihető(k) be az aktuális bázisba, akkor $\mathbf{A}$ nem invertálható. Végezetül tekintsünk két mintafeladatot.

7.5.2. Példa. Amennyiben létezik, határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix} \tag{7.5.5}$$

mátrix inverzét!

Az alábbi táblázat első részébe bemásoltuk az $\mathbf{A}$ mátrix elemeit.

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{e}}_3$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{e}}_3$	$\overline{\mathbf{e}}_1$	$\overline{\mathbf{a}}_3$
$\overline{\mathbf{a}}_2$ $\not\overline{\mathbf{e}}_1$	1	0	1	-1	[-1]	2	1	-1	-2
$\overline{\mathbf{e}}_2$	2	1	0	-2	-1	2	-1	-1	0
$\overline{\mathbf{a}}_1$ $\not\overline{\mathbf{e}}_3$	[1]	1	-1	1	1	-1	0	1	1

Válasszuk először generálóelemnek pl. az [1]-et. Ezzel $\overline{\mathbf{a}}_1$ került be $\overline{\mathbf{e}}_3$ helyére a bázisba. Korábban már részleteztük, hogy most egyszerre $n = 3$ egyenletrendszert oldunk meg, ahol azok jobb oldalán rendre az $\overline{\mathbf{e}}_1, \overline{\mathbf{e}}_2, \overline{\mathbf{e}}_3$ vektorok állnak. Mivel az $\overline{\mathbf{a}}_1$ vektor bázisba való bevitelével $\overline{\mathbf{e}}_3$ kikerült onnan – és szükségünk van $\overline{\mathbf{e}}_3$ koordinátáira –, így a táblázat következő részétől fogva $\overline{\mathbf{e}}_3$ foglalja el az $\overline{\mathbf{a}}_1$ vektor helyét a fejlécben. Az már korábban szerepelt, hogyan kell meghatározni a bázisból kikerülő vektor új koordinátáit.

Második lépésként [-1] legyen a generálóelem, amely az $\overline{\mathbf{a}}_2$ és $\overline{\mathbf{e}}_1$ vektorok szerepcseréjével jár. Az eljárás ezután félbeszakad, mert $\overline{\mathbf{a}}_3$ – az utolsó beviendő vektor – $\overline{\mathbf{e}}_2$-re (az utolsó régi bázisvektorra) vonatkozó koordinátája 0. Nem sikerült bejuttatni a bázisba az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3$ vektorok mindegyikét, tehát az $\mathbf{A}$ mátrix nem invertálható. □

7.5.3. Példa. Határozzuk meg a

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \\ 0 & 2 & 1 \end{bmatrix} \tag{7.5.6}$$

mátrix inverzét, ha létezik!

A számításokat a következő táblázat mutatja.

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{e}}_2$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{e}}_2$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{e}}_3$	$\overline{\mathbf{e}}_2$	$\overline{\mathbf{e}}_1$	$\overline{\mathbf{e}}_3$
$\overline{\mathbf{a}}_2$ $\not\overline{\mathbf{e}}_1$	1	0	1	-1	-2	-2	-1	$\boxed{2}$	2	-1/2	1/2	1
$\overline{\mathbf{a}}_1$ $\not\overline{\mathbf{e}}_2$	$\boxed{1}$	2	3	1	2	3	1	-4	-3	-1	2	1
$\overline{\mathbf{a}}_3$ $\not\overline{\mathbf{e}}_3$	0	2	1	0	2	$\boxed{1}$	0	2	1	1	-1	-1

Mivel az $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3$ vektorok mindegyikét sikerült bevinni a bázisba, ezért $\mathbf{A}$ invertálható. Az egyenletrendszerek megoldásának leolvasási technikája szerint az $\overline{\mathbf{a}}_i$ végső bázisvektorokat ($i = 1, 2, 3$) az indexük szerint növekvő sorrendben kell egymás alá írni, hogy a változók sorrendjében kapjuk meg a megoldást. Az $\overline{\mathbf{e}}_2, \overline{\mathbf{e}}_1, \overline{\mathbf{e}}_3$ (oszlop)vektorokat is az eredeti sorrendbe kell állítani ahhoz, hogy a 3 egyenletrendszer megoldását a felírásuk sorrendjében lássuk. Tehát $\mathbf{B}$ inverzének megállapításához a táblázat utolsó részében látható 3×3-as mátrixban sor- ill. oszlopcseréket kell végrehajtani. Először felcseréljük az első és második sort, majd a kapott újabb mátrixban az első és második oszlopot. Ezzel helyreálltak az előbb előírt vektorsorrendek, megkaptuk $\mathbf{B}$ inverzét:

$$\mathbf{B}^{-1} = \begin{bmatrix} 2 & -1 & 1 \\ 1/2 & -1/2 & 1 \\ -1 & 1 & -1 \end{bmatrix} . \tag{7.5.7}$$

7.5.2. Feladat. Egy $\mathbf{C} \in \mathbb{R}^{2\times 2}$ mátrix inverzének meghatározása során végül az alábbi táblázathoz jutottunk.

	$\overline{\mathbf{e}}_2$	$\overline{\mathbf{e}}_1$
$\overline{\mathbf{a}}_2$	5	-8
$\overline{\mathbf{a}}_1$	-8	13

Olvassuk le a táblázatból a $\mathbf{C}$ mátrix inverzét, majd határozzuk meg $\mathbf{C}$-t is!

□

Bár nem tartozik az inverz mátrix problémakörhöz, de filozófiájában és a megoldás menetében nagyon hasonló az alábbi feladat. Tekintsük az

$$\mathbf{AX} = \mathbf{B} \tag{7.5.8}$$

mátrixegyenletet, ahol $\mathbf{A}$ és $\mathbf{B}$ megegyező számú sorral rendelkező mátrixok (azaz $\mathbf{A}$ és $\mathbf{B}$ oszlopvektorai azonos vektortérből valók). Keressük azon $\mathbf{X}$ mátrixokat, melyek kielégítik a (7.5.8) egyenletet. Vegyük észre, hogy ha $\mathbf{A}$ egy adott $n \times n$-es mátrix, továbbá $\mathbf{B} = \mathbf{E}_n$, akkor a most vizsgált probléma egy speciális eseteként az inverz mátrix meghatározásához jutunk.

A (7.5.8) egyenletrendszer – hasonlóan a (7.5.3) rendszerhez – a $\mathbf{B}$ mátrix oszlopvektorai számával megegyező számú egyenletrendszerre esik szét, és ezen egyenletrendszerek együtthatómátrixai azonosak. Emiatt párhuzamosan meg tudjuk oldani őket, és akkor

kapunk megoldást, ha $\mathbf{A}$ oszlopvektorainak terében benne lesz $\mathbf{B}$ összes oszlopvektora. Szemléltessük a most elhangzottakat egy egyszerű példával!

7.5.4. Példa. Oldjuk meg az alábbi mátrixegyenletet!

$$\begin{bmatrix} -4 & -5 & -6 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ -1 & 1 \end{bmatrix}. \tag{7.5.9}$$

Jelölje rendre $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3$ ill. $\overline{\mathbf{b}}_1, \overline{\mathbf{b}}_2$ az adott mátrixok oszlopvektorait. Elvégezve két elemi bázistranszformációt, a

	$\overline{\mathbf{a}}_1$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}_1$	$\overline{\mathbf{b}}_2$	$\overline{\mathbf{a}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}_1$	$\overline{\mathbf{b}}_2$	$\overline{\mathbf{a}}_3$	$\overline{\mathbf{b}}_1$	$\overline{\mathbf{b}}_2$
$\overline{\mathbf{a}}_2 \not\in \overline{\mathbf{a}}_1$	-4	-5	-6	2	-3	3	6	-2	1	2	-2/3	1/3
$\overline{\mathbf{a}}_1 \not\in \overline{\mathbf{a}}_2$	1	2	3	-1	1	2	3	-1	1	-1	1/3	1/3

táblázathoz juthatunk, amelyből leolvasható, hogy $\overline{\mathbf{a}}_3, \overline{\mathbf{b}}_1, \overline{\mathbf{b}}_2 \in L(\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2)$. Tehát a két egyenletrendszer mindegyike megoldható, mindegyikben egy szabad változó van, a megoldások száma végtelen. Ügyelni kell arra, hogy a táblázat utolsó két sora a generálóelemek választása miatt rossz sorrendben van a leolvasáshoz. Az ismeretlen mátrix két szabad paraméterrel rendelkezik (x_{31} és x_{32}):

$$\mathbf{X} = \begin{bmatrix} 1/3 + x_{31} & 1/3 + x_{32} \\ -2/3 - 2x_{31} & 1/3 - 2x_{32} \\ x_{31} & x_{32} \end{bmatrix} \quad . \tag{7.5.10}$$

□

Ellenőrző kérdések

1. Milyen típusú mátrixoknak értelmeztük az inverzét?
2. Mi az inverz mátrix definíciója?
3. Milyen szükséges és elégséges feltételt tanultunk az $\mathbf{A} \in \mathbb{R}^{n \times n}$ mátrix inverzének létezésére?
4. Milyen módszert tanultunk az inverz mátrix meghatározására?
5. Milyen kapcsolat van az inverz mátrix meghatározása és az egyenletrendszerek megoldása között?
6. Hogyan ellenőrizhető le egyszerűen, hogy két mátrix egymás inverze?

A feladatok megoldásai

7.5.1.
Mivel $\det(\mathbf{A}) = -2 \neq 0$, $\det(\mathbf{B}) = 0$ ezért $\mathbf{A}$-nak létezik inverze, $\mathbf{B}$ nem invertálható.

7.5.2.
A

$$\mathbf{C}^{-1} = \begin{bmatrix} 13 & -8 \\ -8 & 5 \end{bmatrix} \tag{7.5.11}$$

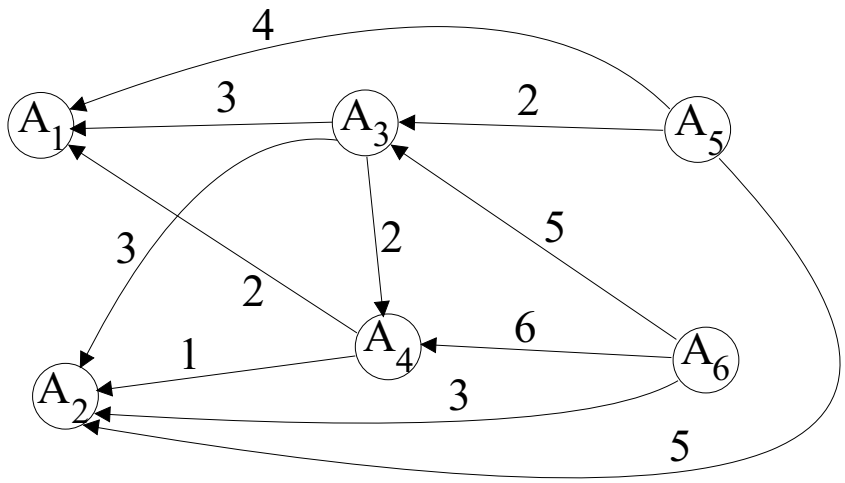
mátrixhoz a táblázatban látható mátrix két sorának ill. oszlopának cseréjével jutunk. Az eredeti $\mathbf{C}$ mátrix meghatározásához használjuk a $\mathbf{C} = (\mathbf{C}^{-1})^{-1}$ összefüggést. (Ha a feladat táblázatában található (-8) elemek helyett 8-at veszünk, akkor éppen $\mathbf{C}$-t kapjuk.)

7.6. Egy gazdasági feladat megoldása

Ebben a fejezetben a mátrixok egy egyszerű gazdasági alkalmazását adjuk meg. Egy gyártási folyamat során a közvetlen ráfordítások ismeretében meg fogjuk határozni a teljes ráfordításokat.

Egy üzemben konyhabútorokat készítenek. A tervezés során feljegyzik, hogy az egyes bútordarabokhoz pontosan mire van szükségük. Ez a terv pl. a következő módon nézhet ki. Egy polc elkészítéséhez szükség van két deszkára és egy doboz lakkra. Egy oldallaphoz három deszka kell és két doboz lakk. Ugyanígy meg lehet adni az alsó lap, felső lap és hátsó lap anyagszükségletét. Egy szekrényhez szükség van két polcra, két oldallapra, egy alsó lapra, tetőre és hátsó borításra. Ezen kívül még 16 szögre, melyekkel a hátsó lapot felszögelik a szekrényre ill. 24 facsapra a polcok és oldallapok összeillesztéséhez. Ezzel persze a konyhabútornak csak az egyik elemét állították elő, de hasonlóan meg lehet adni a többi elem anyagszükségletét is. Természetesen egy üzemnek nem csak az egyes résztermékek előállításához szükséges anyagok mennyiségét kell ismernie, hanem ennél sokkal fontosabb számára, hogy hány deszkát, hány doboz lakkot, hány szöget ill. facsapot kell vennie ahhoz, hogy azokból az adott konyhabútor elkészüljön. Ebben a fejezetben ennek a problémának a matematikai modelljét adjuk meg.

Tekintsük az alábbi ábrát, amely valamilyen gyártási folyamat során a közvetlen ráfordításokat szemlélteti hat termék $(A_1, \ldots, A_6)$ között. Az A_6 termékből 5 egység-



7.6.1. ábra. *A közvetlen ráfordítások szemléltetése.*

nyi kell egy egységnyi A_3 termék előállításához. Ezt úgy szemléltettük, hogy az A_6 terméktől egy A_3-ba mutató nyilat húztunk, és a nyílra ráírtuk az 5-ös számjegyet. Az A_1 és A_2 termékekből nem indul ki nyíl, azaz ezeket már nem használjuk újabb termékek előállítására. Ezeket a termékeket ezért végtermékeknek nevezzük. Az A_5 és A_6 termékekbe nem érkezik nyíl, ezeket alaptermékeknek (nyersanyagoknak hívjuk) hívjuk.

A közvetlen ráfordításokat szemléltető ábrából felírhatunk egy $\mathbf{K} = [k_{ij}]_{n\times n}$ mátrixot (n a feladatban szereplő termékek száma), melyet a közvetlen ráfordítások mátrixának fogunk hívni. Ezen mátrix k_{ij} eleme azt adja meg, hogy az A_i termékből hány egységnyire van szükség közvetlenül egy egységnyi A_j termék előállításához. A $\mathbf{K}$ mátrix néhány tulajdonsága közvetlenül adódik az értelmezéséből.

a) A $\mathbf{K}$ mátrix végtermékekhez tartozó sorai csak nullákat tartalmaznak.

b) A $\mathbf{K}$ mátrix alaptermékekhez tartozó oszlopai csak nullákat tartalmaznak.

c) A $\mathbf{K}$ mátrix főátlója csak nullákat tartalmaz.

d) A k_{ij} és k_{ji} elemek közül legfeljebb az egyik lehet csak nullától különböző, hiszen ha az A_i termékből kell az A_j termékhez valahány, akkor az A_j-ből nem kellhet már az A_i-hez és fordítva.

A 7.6.1. ábrához tartozó $\mathbf{K}$ mátrix

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 3 & 0 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 \\ 4 & 5 & 2 & 0 & 0 & 0 \\ 0 & 3 & 5 & 6 & 0 & 0 \end{bmatrix}.$$

Hasonló módon definiálhatunk egy $\mathbf{T} = [t_{ij}]_{n\times n}$ mátrixot is, amelyet a teljes ráfordítások mátrixának fogunk hívni. Ennek t_{ij} eleme azt adja meg, hogy az A_i termékből hány egységre van szükség összesen egy darab A_j termék előállításához. A feladat az, hogy adott $\mathbf{K}$ közvetlen ráfordítási mátrixból határozzuk meg a teljes ráfordítások mátrixát.

7.6.1. Tétel. *Egy adott $\mathbf{K}$ közvetlen ráfordítási mátrixból a $\mathbf{T} = (\mathbf{E} - \mathbf{K})^{-1}$ képlettel határozható meg a teljes ráfordítások mátrixa.*

Bizonyítás. Vizsgáljuk meg, hogy mit adnak meg a $\mathbf{K}$ mátrix hatványainak elemei. A $\mathbf{K}^2$ mátrix elemei a

$$(\mathbf{K}^2)_{ij} = k_{i1}k_{1j} + k_{i2}k_{2j} + \ldots + k_{in}k_{nj} \tag{7.6.1}$$

alakban írhatók. Itt a $k_{i1}k_{1j}$ első tag első tényezője azt mutatja meg, hogy az A_i termékből mennyi kell az A_1-hez, a második pedig azt, hogy az A_1-ből mennyi kell az A_j-hez. A szorzat tehát azt adja meg, hogy az A_i termékből mennyi kell az A_j termékhez úgy, hogy az az A_1 termékbe építve kerül A_j-be. A $\mathbf{K}^2$ mátrix $(\mathbf{K}^2)_{ij}$ eleme tehát azt fogja megadni, hogy az A_i termékből mennyi kell A_j-hez úgy, hogy azok egy közbülső termékbe építve - azaz két lépésben - kerülnek A_j-be. A $\mathbf{K}$ mátrix többi hatványát is hasonlóan megvizsgálva azt kapjuk, hogy a $\mathbf{K}^s$ $(s \in \mathbb{N})$ mátrix $(\mathbf{K}^s)_{ij}$ eleme azt adja meg, hogy hány A_i termék kell egy A_j termékhez úgy, hogy azok $s-1$ közbülső termékbe épülve – s lépésben – kerülnek A_j-be. A fentiek miatt nyilvánvaló, hogy $\mathbf{K}^s = \mathbf{0}$, ha $s \geq n$.

A $\mathbf{T}$ mátrixot tehát a $\mathbf{T} = \mathbf{E} + \mathbf{K} + \ldots + \mathbf{K}^{n-1}$ képlettel lehet kiszámítani. A képlet egyszerűsítése érdekében tekintsük a következő egyenlőséget.

$$(\mathbf{E} - \mathbf{K})\mathbf{T} = (\mathbf{E} - \mathbf{K})(\mathbf{E} + \mathbf{K} + \mathbf{K}^2 + \ldots + \mathbf{K}^{n-1}) = \mathbf{E} - \mathbf{K}^n = \mathbf{E}.$$

Ebből látható, hogy a $\mathbf{T}$ mátrix éppen az $\mathbf{E}-\mathbf{K}$ mátrix inverze, hiszen összeszorozva őket az egységmátrixot kapjuk. Ezt akartuk megmutatni. ■

A tétel egyik következménye, hogy a $\mathbf{T}$ mátrixból is meg lehet határozni $\mathbf{K}$-t. A $\mathbf{T} = (\mathbf{E} - \mathbf{K})^{-1}$ egyenlőségből egyszerű átrendezéssel adódik, hogy $\mathbf{K} = \mathbf{E} - \mathbf{T}^{-1}$

Az általunk vizsgált gyártási folyamat esetén a $\mathbf{K}$ mátrix egyes hatványai a következők.

$$\mathbf{K}^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 6 & 0 & 4 & 0 & 0 \\ 27 & 21 & 0 & 10 & 0 & 0 \end{bmatrix}, \quad \mathbf{K}^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 8 & 4 & 0 & 0 & 0 & 0 \\ 20 & 10 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{K}^4 = \mathbf{0}.$$

A $\mathbf{K}^2$ mátrix hatodik sorának első eleme 27. Ez azt jelenti, hogy egy A_1 termékbe 27 A_6 termék kerül két lépésben. Valóban, egy A_1-hez kettő A_4 kell, egy A_4-hez pedig hat A_6. Azaz összesen 12 A_6 kell egy A_1-hez az A_4 terméken keresztül. Hasonló módon az A_1 termékhez 15 darab kell A_6-ból az A_3 terméken keresztül. Így kapjuk, hogy $12+15=27$.

A $\mathbf{K}^3$ mátrix hatodik sorának első eleme 20. Ez azt jelenti, hogy egy A_1 termékbe 20 A_6 termék kerül három lépésben. Egy A_1-hez kettő A_4 kell, ahhoz kettő A_3 és ahhoz öt A_6. Ez összesen $2 \cdot 2 \cdot 5 = 20$ A_6 beépülését jelenti A_1-be három lépésben.

A $\mathbf{T}$ mátrix így a $\mathbf{T} = \mathbf{E} + \mathbf{K} + \mathbf{K}^2 + \mathbf{K}^3$ képlettel adódik, azaz

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 7 & 5 & 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 \\ 18 & 15 & 2 & 4 & 1 & 0 \\ 47 & 34 & 5 & 16 & 0 & 1 \end{bmatrix}.$$

Ugyanezt a $\mathbf{T}$ mátrixot kapjuk a $\mathbf{T} = (\mathbf{E} - \mathbf{K})^{-1}$ képlet segítségével. A $\mathbf{T}$ mátrix hatodik sorának első eleme 47, ami azt jelenti, hogy egy darab A_1 termék előállításához összesen 47 A_6 termékre van szükség.

Folytassuk tovább a gyártási folyamat tervezését. Tegyük fel, hogy az A_1 végtermékből 25, az A_2 végtermékből 40 darabot szeretnénk gyártani és az A_3-ból 10-et, az A_4-ből pedig 15 darabot szeretnénk a raktárba tenni. Arra vagyunk kíváncsiak, hogy ebben az esetben, az egyes termékekből mennyit kell gyártani.

Vizsgáljuk a problémát általánosan, azaz tegyük fel, hogy a gyártási folyamat során az A_1 termékből p_1, az A_2 termékből p_2, stb., az A_n termékből p_n darabot kell előállítanunk. Ez azt jelenti, hogy a végtermékekből és a közbülső termékekből az adott mennyiséget le kell gyártani, az alaptermékekből pedig be kell szerezni az előírt mennyiséget (az alap és az elkészült közbülső termékeket elraktározzuk későbbi felhasználásra). A szükséges mennyiségekkel definiáljuk a $\overline{\mathbf{p}} = [p_1, \ldots, p_n]^\top$ ún. programvektort, Most megmutatjuk, hogy az adott program megvalósításához szükséges termékdarabszámokat a $\mathbf{T}\overline{\mathbf{p}}$ vektor elemei fogják megadni. Írjuk fel a $\mathbf{T}\overline{\mathbf{p}}$ vektor i-edik elemét.

$$(\mathbf{T}\overline{\mathbf{p}})_i = t_{i1} \cdot p_1 + \ldots + t_{in} \cdot p_n.$$

Az első tag ($t_{i1} \cdot p_1$) első tényezője azt adja meg, hogy az A_i termékből mennyire van szükség összesen egy darab A_1 termék előállításához. A második tényező pedig azt mutatja meg hogy A_1-ből mennyit kell előállítanunk. Így a szorzat azt fogja megmutatni, hogy a kívánt mennyiségű A_1-hez összesen hány A_i-re van szükség. Ezek szerint tehát $(\mathbf{T}\overline{\mathbf{p}})_i$ valóban azt mutatja meg, hogy A_i-ből mennyire van szükség az adott program megvalósításához.

Konkrét feladatunk esetén $\overline{\mathbf{p}} = [25, 40, 10, 15, 0, 0]^\top$ és

$$\mathbf{T}\overline{\mathbf{p}} = [25,\ 40,\ 415,\ 105,\ 1130,\ 2825]^\top.$$

Itt tehát a szereplő 2825-ös szám azt mutatja, hogy az adott program megvalósításához összesen 2825 darab kell az A_6 alaptermékből.

Végül feltehetjük a legfontosabb kérdést, azt, hogy mennyibe fog kerülni a $\overline{\mathbf{p}}$ programvektorral leírt gyártási folyamat. Ehhez definiálunk egy $\overline{\mathbf{a}} = [a_1, \ldots, a_n]^\top$ vektort, melyet árvektornak fogunk hívni és a_i eleme azt fogja megadni forintban, hogy egy darab A_i termék előállítása vagy megvásárlása (alaptermék esetén) mennyibe kerül. Könnyű látni, hogy egy $\overline{\mathbf{p}}$ program árát ekkor az $\overline{\mathbf{a}}^\top\mathbf{T}\overline{\mathbf{p}}$ 1×1-es mátrix egyetlen eleme fogja megadni, szintén forintban.

Konkrét példánk esetén, ha mondjuk $\overline{\mathbf{a}} = [2000, 1500, 500, 700, 50, 60]^\top$, akkor a $\overline{\mathbf{p}} = [25, 40, 10, 15, 0, 0]^\top$ program végrehajtása 617000 Ft-ba kerül.

Ellenőrző kérdések

1. Milyen gazdasági problémát vizsgált ez a fejezet?
2. Hogyan definiáltuk a közvetlen ráfordítások mátrixát?
3. Hogyan lehet kiszámítani a teljes ráfordítások mátrixát?
4. Mi az a programvektor, és hogyan lehet kiszámítani egy adott program megvalósításához szükséges termékmennyiségeket?

Tartalomjegyzék

1. A határozatlan integrál (Horváth Róbert) 3
1.1. Primitív függvények és a határozatlan integrál . . . 3
1.2. Integrálás helyettesítéssel . . . 7
1.3. Parciális integrálás . . . 11
1.4. Nevezetes függvények integrálása . . . 14

2. A határozott integrál (Horváth Róbert) 23
2.1. Egyváltozós függvények integrálhatósága és a határozott integrál . . . 23
2.2. A határozott integrál tulajdonságai . . . 29
2.3. A Newton-Leibniz formula . . . 30
2.4. A helyettesítéses és parciális integrálás határozott integrálokra . . . 33
2.5. Az improprius integrál . . . 34
2.6. A határozott integrál alkalmazásai . . . 38

3. Kétváltozós függvények integrálása (Horváth Róbert) 43
3.1. A kétváltozós integrál értelmezése . . . 43
3.2. A kétváltozós integrál kiszámítása . . . 49
3.2.1. Integrálás téglalapon . . . 49
3.2.2. Integrálás normáltartományon . . . 51
3.2.3. Integrálás helyettesítéssel . . . 53

4. Differenciálegyenletek (Szalay László) 60
4.1. Alapfogalmak . . . 60
4.2. Elsőrendű differenciálegyenletek . . . 66
4.2.1. Szétválasztható változójú differenciálegyenletek . . . 66
4.2.2. Szétválasztható változójúra visszavezethető differenciálegyenletek . 69
4.2.3. Lineáris differenciálegyenletek . . . 73
4.3. Másodrendű differenciálegyenletek . . . 78
4.3.1. Állandó együtthatós lineáris differenciálegyenletek . . . 79

5. Mátrixok (Horváth Róbert) 85
5.1. Nevezetes mátrixok, mátrixműveletek . . . 85
5.2. Mátrixok determinánsa . . . 99

6. Vektorterek (Horváth Róbert) 109
6.1. Vektortér fogalma, nevezetes vektorterek . . . 109
6.2. Altér, lineáris kombináció, generátorrendszer . . . 114
6.3. Lineáris függetlenség, összefüggőség . . . 119
6.4. Bázis, dimenzió . . . 123
6.5. Koordináták, elemi bázistranszformáció . . . 129

7. Alkalmazások **136**
7.1. Vektorrendszerek és mátrixok rangja (Horváth Róbert) 136
7.2. Mátrixok determinánsának kiszámítása (Horváth Róbert) 138
7.3. Lineáris egyenletrendszerek megoldása (Szalay László) 141
7.3.1. Lineáris egyenletrendszerek megoldhatósága 143
7.3.2. Lineáris egyenletrendszerek megoldásainak száma 144
7.3.3. Lineáris egyenletrendszerek megoldásainak leírása 144
7.3.4. Lineáris egyenletrendszerek megoldási módszerei 145
7.3.5. Homogén lineáris egyenletrendszerek 150
7.4. Mátrixok sajátértékei, sajátvektorai (Szalay László) 152
7.5. Mátrixok inverzének kiszámítása (Szalay László) 156
7.6. Egy gazdasági feladat megoldása (Horváth Róbert) 160

Tárgymutató

Additív inverz, 110
Alapintegrálok, 4
Alaptermék, 160
Aldeterminány, 105
Altér, 114
    generált, 116
    triviális, 114
Antiszimmetrikus mátrix, 91
Árvektor, 162

Bázis, 123
Bázistranszformáció, 129
Blokkmátrix, 94

Cauchy-Bunyakovszkij-Schwarz
        egyenlőtlenség, 94
CBS egyenlőtlenség, 94
Cramer-szabály, 149

det(**A**), 100
Determináns, 100
    al∼, 105
    előjeles al∼, 105
    kifejtése, 100, 105
    Sarrus-szabály, 101
Diagonális mátrix, 89
Differenciálegyenlet, 61
    általános megoldása, 63
    homogén fokszámú, 71
    partikuláris megoldása, 63
    szinguláris megoldása, 63
dim(V), 125
Dimenzió, 125
    véges, 123
    végtelen, 123
Dirichlet függvény, 25

Együtthatómátrix, 143
Egyenletrendszer, 141
    együtthatómátrixa, 143
    homogén, 142
    inhomogén, 142
    megoldása elemi
        bázistranszformációval, 145
    megoldásainak leírása, 144
    megoldásainak száma, 144
    megoldhatósága, 143
Egységmátrix, 89
Egységnégyzet, 49
Elemi bázistranszformáció, 129
    táblázata, 131
Előjeles aldeterminány, 105

Függvény
    n-ed rendben homogén, 70
Felosztás
    finomsága, 23
Ferdén szimmetrikus mátrix, 91
Főátló, 88
Független vektorok, 119

Generálóelem, 131
Generált altér, 116
Generátorrendszer, 116

Háromszögmátrix, 103
Határozott integrál, 23
    tulajdonságai, 29
Határozatlan integrál, 3
Helyettesítéses integrálás, 7, 53

Improprius integrál, 34
Integrál
    határozott, 23
Integrálás
    helyettesítéssel, 7, 53
    irracionális függvényeké, 19
    korlátos halmazon, 47
    nevezetes függvényeké, 14
    normáltartományon, 51
    parciális, 11
    racionális törtfüggvényeké, 14
    téglalapon, 49
    trigonmetrikus függvényeké, 17
Integrálhatóság

korlátos halmazon, 47
Integrálközelítő összeg, 23
kétváltozós, 44
Integrálkritérium, 39
Inverz mátrix, 89
meghatározása elemi bázistranszformációval, 156
Irracionális függvények integrálása, 19

Jacobi-determináns, 53

Közvetlen ráfordítások mátrixa, 160
Kétváltozós függvények integrálja, 47
Kifejtési tétel, 105
Kommutáló mátrixok, 88
Koordináták, 129

$L(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, 115
Lineáris összefüggőség, 119
Lineáris egyenletrendszer, 141
együtthatómátrixa, 143
homogén, 142
inhomogén, 142
megoldásaelemi bázistranszformációval, 145
megoldásainak leírása, 144
megoldásainak száma, 144
megoldhatósága, 143
Lineáris függetlenség, 119
Lineáris kombináció, 114
Lineáris tér, 110

Mátrix, 85, 86
-ok szorzása, 87
antiszimmetrikus, 91
blokk ∼, 94
determinánsa, 100
diagonális, 89
elemei, 85
főátlója, 88
ferdén szimmetrikus, 91
invertálható, 89
inverze, 89
közvetlen ráfordítások ∼a, 160
kvadratikus, 86
négyzetes, 86
rangja, 137
reguláris, 89
szimmetrikus, 91
teljes ráfordítások ∼a, 161
transzponált, 90
Mátrixegyenlet
megoldása elemi bázistranszformációval, 158
Mátrix
sajátértéke, 152
sajátvektora, 152

Nevezetes függvények integrálása, 14
Newton-Leibniz formula, 30
Normáltartomány
síkbeli, 51
Norma, 93
Nullmátrix, 87

Oszlopvektor, 86
Összefüggő vektorok, 119

Parciális integrálás, 11
Parciális törtekre bontás, 15
Permutációk, 100
száma, 100
Primitív függvény, 3
Programvektor, 162

$\mathbb{R}^n$, 86
$\mathbb{R}^{k\times n}$, 85
Racionális törtfüggvények integrálása, 14
Rang
mátrixok ∼ja, 137
vektorrendszer ∼ja, 136
Reguláris mátrix, 89
Rendszer determinánsa, 149

Sajátérték, 152
Sajátvektor, 152
Sakktábla szabály, 105
Sarrus-szabály, 101
Síkvektorok, 109
Skaláris szorzás, 92
Sorvektor, 91
Szabad paraméter, 144
Szabad változó, 144
Szimmetrikus mátrix, 91
Szinguláris mátrix, 89

Téglalap, 43
Teljes négyzetté alakítás, 19
Teljes ráfordítások mátrixa, 161
Transzponált mátrix, 90

Trigonometrikus függvények integrálása, 17
Triviális előállítás, 119

Végtermék, 160
Vektor hossza, 93
Vektorok szöge, 93
Vektorrendszer rangja, 136
Vektortér, 110
  bázisa, 123
  dimenziója, 125
  függvények ∼e, 111
  generálása, 116
  mátrixok ∼e, 111
  példák, 111
  polinomok ∼e, 111
  síkvektorok ∼e, 109
  tulajdonságai, 110
  véges dimenziós, 123
  végtelen dimenziós, 123

Wronski determináns, 78