Értékünk az ember

Humánerőforrás-fejlesztési Operatív Program

Gáspár Csaba

Molnárka Győző

Miletics Edit

Lineáris algebra és többváltozós függvények

SZÉCHENYI ISTVÁN EGYETEM GYŐR

Magyarország célba ér

NEMZETI FEJLESZTÉSI TERV

Készült a HEFOP 3.3.1-P.-2004-09-0102/1.0 pályázat támogatásával

Szerzők: Gáspár Csaba
Molnárka Győző
Miletics Edit

Lektor: Hegedűs Csaba, egyetemi docens

Tartalomjegyzék

1. Bevezetés **5**

2. Vektorterek **7**

2.1. Vektorterek és altereik 7
2.2. Lineáris kombináció, lineáris függetlenség 11
2.3. Vektorterek bázisa, dimenziója 15
2.4. Skaláris szorzat és norma $\mathbf{R}^n$-ben 16
2.5. Ortogonalitás, ortogonális vetület 19
2.6. Feladatok 24

3. Vektorgeometria **28**

3.1. Síkvektorok, egyenesek a síkon 28
3.2. Térvektorok, egyenesek a térben 33
3.3. Vektoriális szorzat 36
3.4. Síkok a térben 38
3.5. Feladatok 43

4. Lineáris leképezések, mátrixok **49**

4.1. Lineáris leképezések 49
4.2. Mátrixok, műveletek mátrixokkal 51
4.3. Mátrixszorzás és lineáris leképezések 54
4.4. Mátrixok inverze és determinánsa 56
4.5. Lineáris egyenletrendszerek megoldhatósága 61
4.6. Megoldási algoritmus: a Gauss-elimináció 63
4.7. Sajátérték, sajátvektor 70
4.8. Önadjungált mátrixok 76
4.9. Néhány speciális mátrixosztály 80
4.10. Feladatok 85

5. Többváltozós függvények **94**

5.1. Többváltozós függvények bevezetése 94
5.2. Folytonosság 95
5.3. Többváltozós függvények differenciálhatósága 97
5.4. Parciális és iránymenti derivált, a második derivált mátrixa . 99
5.5. Többváltozós függvények lokális szélsőértékei 104
5.6. Feltételes szélsőérték feladatok 110
5.7. Néhány alkalmazás 115
5.8. Többszörös integrálok 118

5.9. Feladatok 130

1. Bevezetés

Ez a jegyzet a Széchenyi István Egyetem mérnöki BSc szakos hallgatói számára készült. A jegyzet a lineáris algebra és a többváltozós analízis bevezető fejezeteit tartalmazza, melyek a mérnöki BSc képzés keretében a második félévben kerülnek leadásra.

A jegyzet feltételezi az (első félévben leadott) Analízis c. tantárgy ismeretét.

A jegyzet négy fejezetből áll. Az első fejezetben bevezetjük a legfontosabb lineáris algebrai fogalmakat (vektortér, altér, lineáris kombináció, lineáris függetlenség, bázis, dimenzió stb.), és az ezekre vonatkozó alapvető tételeket. A második fejezet lényegében az első fejezetben felépített fogalom- és tételkör sík- és térgeometriai alkalmazása: itt egyenesek és síkok leírását végezzük el, azok jellemző problémaköreinek felvázolásával. A harmadik fejezet a lineáris leképezések és a mátrixok vizsgálatának van szentelve. Itt foglalkozunk a sokismeretlenes lineáris egyenletrendszerek megoldásának néhány algoritmusával is (Gauss- ill. Gauss-Jordan-elimináció). Végül a negyedik fejezet a többváltozós analízis alapjait mutatja be (többváltozós függvények bevezetése, differenciálásuk, integrálásuk, többváltozós szélsőérték problémák).

A jegyzet anyagát igyekeztünk *alkalmazáscentrikusan* felépíteni. Ennek megfelelően pl. elhagytuk a lineáris egyenletrendszerek megoldására vonatkozó klasszikus Cramer-szabályt (mely elméletileg nagy jelentőségű, de gyakorlati feladatmegoldásra – egészen kisméretű feladatoktól eltekintve – teljesen alkalmatlan). Helyette a Gauss-elimináció néhány változatának leírását építettük be a jegyzetbe. A többváltozós analízis tárgyalásakor (utolsó fejezet) külön hangsúlyt kaptak a szélsőérték feladatok (mind feltételes, mind feltétel nélküli megfogalmazásban). Itt találkozik legszembetűnőbben az analízis és a lineáris algebra addig különállónak látszó problémaköre. Ez a feladattípus különösen alkalmas a konstruktív problémamegoldás fejlesztésére, mivel a megoldandó matematikai probléma pontos megfogalmazása maga is a feladat része.

Mindegyik fejezet utolsó szakasza a fejezet témakörébe vágó feladatokat tartalmaz. Ugyanitt megtalálhatók a megoldások is, rövidebb-hosszabb levezetésekkel, útmutatásokkal együtt. Hangsúlyozzuk azonban, hogy ezek a feladatok az adott témakör lehetséges alkalmazásainak csak nagyon vékony szeletét jelentik, és semmiképp sem pótolják egy önálló feladatgyűjtemény használatát.

Köszönet illeti dr. Hegedüs Csabát, az Eötvös Loránd Tudományegyetem docensét a jegyzet gondos lektorálásáért és értékes megjegyzései-

ért.

Kérjük a tisztelt Olvasókat, hogy véleményüket, megjegyzéseiket, észrevételeiket küldjék el a

gasparcs@sze.hu

e-mail címre.
Eredményes felhasználást kívánnak a szerzők:

Dr. Gáspár Csaba
Dr. Molnárka Győző
Miletics Edit

2. Vektorterek

Ebben a fejezetben a geometriai vektorfogalmat ("irányított szakasz") általánosítjuk. Egymástól egészen különböző matematikai objektumokat is vektoroknak fogunk nevezni, ha definiálva vannak rajtuk bizonyos egyszerű műveletek, melyek ugyanolyan (alább specifikált) tulajdonságokkal rendelkeznek. Ily módon a bevezetésre kerülő vektorfogalom a közönséges összeadás és szorzás jól ismert tulajdonságait terjeszti ki a számoknál sokkal általánosabb struktúrákra.

2.1. Vektorterek és altereik

Vektortér:

Az X nemüres halmazt *valós vektortérnek* (vagy *lineáris térnek*) nevezzük, ha X elemei közt értelmezett egy *összeadás*, $\mathbf{R}$ és X elemei közt pedig egy *skalárral való szorzás* úgy, hogy a következő állítások (az ún. *vektortér-axiómák*) teljesülnek. Tetszőleges $x,y,z \in X$, $\lambda,\mu \in \mathbf{R}$ esetén:

- $x + y = y + x$ (az összeadás kommutatív)
- $x + (y + z) = x + (y + z)$ (az összeadás asszociatív)
- létezik X-ben egy $\mathbf{0}$ *zérusvektor*, melyre $x + \mathbf{0} = x$ teljesül minden $x \in X$ esetén;
- az összeadás *invertálható*, azaz bármely $x \in X$ vektorhoz van oly $x_{-1} \in X$ vektor, hogy összegük a zérusvektor: $x + x_{-1} = \mathbf{0}$
- $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$
- $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$
- $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$
- $1 \cdot x = x$

Könnyen látható, hogy az axiómák közt szereplő additív inverz, azaz az x_{-1} vektor épp $(-1) \cdot x$-szel egyezik, melyet a későbbiekben röviden csak $(-x)$-szel jelölünk. Ha nem okoz félreértést, a skalárral való szorzást jelző szorzópontot elhagyjuk.

Hasonlóan vezethető be a *komplex vektortér* fogalma is (valós helyett komplex $\lambda,\mu \in \mathbf{C}$ számokat szerepeltetve a vektortér-axiómákban). E jegyzet keretein belül valós vektorterekkel foglalkozunk, de megjegyezzük, hogy az eredmények jelentős része komplex vektorterek esetében is igaz marad.

Altér:
Legyen X vektortér. Az $X_0 \subset X$ részhalmazt *altérnek* nevezzük, ha maga is vektortér az X-beli műveletekre nézve.

Egy $X_0 \subset X$ részhalmaz altér voltának eldöntésekor, mivel a vektoroktól megkövetelt műveleti azonosságok nyilván X_0-ban is teljesülnek, elég csak azt ellenőrizni, hogy vajon minden $x,y \in X_0$, $\lambda \in \mathbf{R}$ esetén teljesül-e, hogy $x + y \in X_0$, és $\lambda x \in X_0$. Ezt a tulajdonságot nevezzük *műveleti zártságnak*. Ha ez teljesül, akkor X_0 altér X-ben. A fenti két vizsgálat egyesíthető: könnyen látható, hogy a műveleti zártság pontosan akkor teljesül, ha minden $x,y \in X_0$ vektorra és $\alpha,\beta \in \mathbf{R}$ számokra $\alpha x + \beta y \in X_0$.

Nyilvánvaló, hogy maga X és az egyelemű $\{\mathbf{0}\}$ halmaz alterek X-ben. Ezeket *triviális altereknek* nevezzük. Az is nyilvánvaló, hogy akárhány altér metszete is altér (az unióra ez nem áll!).

Alább példákat mutatunk vektorterekre: a vektortér-axiómák teljesülése könnyen ellenőrizhető.

2.1. Példa: A *valós számok* $\mathbf{R}$ halmaza egyúttal valós vektortér is az összeadásra és a szorzásra nézve.

2.2. Példa: A *komplex számok* $\mathbf{C}$ halmaza egyúttal valós vektortér is az összeadásra és valós számmal való szorzásra nézve.

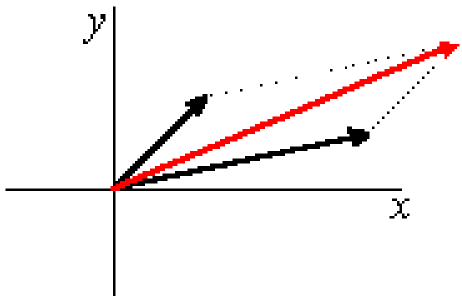
2.3. Példa: A *rendezett valós számpárok* $\mathbf{R}^2$ halmaza valós vektortér az összeadásra és a valós számmal való szorzásra nézve:

$$(x_1,x_2) + (y_1,y_2) := (x_1 + y_1,x_2 + y_2)$$

$$\lambda \cdot (x_1,x_2) := (\lambda x_1,\lambda x_2)$$

tetszőleges $(x_1,x_2),(y_1,y_2) \in \mathbf{R}^2$ és $\lambda \in \mathbf{R}$ esetén.

Az $\mathbf{R}^2$ vektortér elemei azonosíthatók a *geometriai sík* pontjaival. Rögzítve a síkon egy koordinátarendszert, minden pontnak megfelel egy és csakis egy valós, rendezett számpár, ti. a pont koordinátáiból képezett számpár. A sík pontjai viszont azonosíthatók a rögzített koordinátarendszer origójából az illető pontokba mutató irányított szakaszokkal (a pontok helyvektoraival). Ebben a megfeleltetésben az $\mathbf{R}^2$-ben fentebb definiált összeadás épp a jól ismert paralelogramma szabállyal meghatározott összegvektor, míg a skalárral való szorzás a nyújtásnak felel meg. Látjuk tehát, hogy ebben a speciális esetben visszakapjuk a geometriai síkvektor-fogalmat. Könnyen látható az is, hogy a sík nemtriviális alterei az origóra illeszkedő egyenesek (és csak azok).



2.1. ábra. *A geometriai sík mint vektortér*

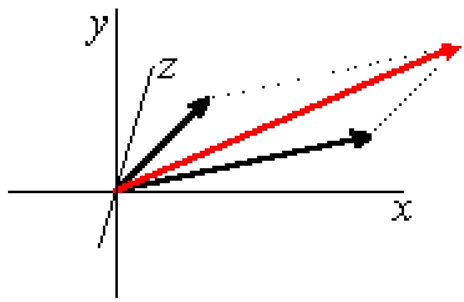
2.4. Példa: A *rendezett valós számhármasok* $\mathbf{R}^3$ halmaza valós vektortér az öszeadásra és a valós számmal való szorzásra nézve:

$$(x_1,x_2,x_3) + (y_1,y_2,y_3) := (x_1 + y_1,x_2 + y_2,x_3 + y_3)$$

$$\lambda \cdot (x_1,x_2,x_3) := (\lambda x_1,\lambda x_2,\lambda x_3)$$

tetszőleges $(x_1,x_2,x_3),(y_1,y_2,y_3) \in \mathbf{R}^3$ és $\lambda \in \mathbf{R}$ esetén.

Az $\mathbf{R}^3$ vektortér elemei azonosíthatók a *geometriai tér* pontjaival. Rögzítve a térben egy koordinátarendszert, minden pontnak megfelel egy és csakis egy valós, rendezett számhármas, ti. a pont koordinátáiból képezett számhármas. A tér pontjai pedig azonosíthatók a rögzített koordinátarendszer origójából az illető pontokba mutató irányított szakaszokkal (a pontok helyvektoraival). Így – az előző példával analóg módon – $\mathbf{R}^3$ elemei a geometriai térvektoroknak is felfoghatók, ahol az összeadást a paralelogramma szabállyal, a skalárral való szorzást a nyújtással definiáljuk. Könnyen látható az is, hogy a geometriai tér nemtriviális alterei az origóra illeszkedő egyenesek és síkok (és csak azok): így pl. az origó mint egyelemű halmaz kivételével, a geometriai tér semmilyen korlátos részhalmaza sem altér.



2.2. ábra. *A geometriai tér mint vektortér*

2.5. Példa: Általánosítva az előző két példát, a *rendezett valós szám-n-esek* $\mathbf{R}^n$ halmaza valós vektortér az alábbi összeadásra és szorzásra nézve:

$$(x_1,x_2,...,x_n) + (y_1,y_2,...,y_n) := (x_1 + y : 1,x_2 + y_2,...,x_n + y_n)$$

$$\lambda \cdot (x_1,x_2,...,x_n) := (\lambda x_1,\lambda x_2,...,\lambda x_n)$$

tetszőleges $(x_1,x_2,...,x_n),(y_1,y_2,...,y_n) \in \mathbf{R}^n$ és $\lambda \in \mathbf{R}$ esetén.

Itt már igen sokféle altér létezik, de ezek nem olyan szemléletesek, mint az előző két példában. Könnyen ellenőrizhető, hogy pl. mindazon rendezett valós szám-n-esek, melyek első komponense 0-val egyenlő, alteret alkotnak $\mathbf{R}^n$-ben. Hasonlóan, mindazon rendezett valós szám-n-esek, melyek komponenseinek összege 0-val egyenlő, szintén alteret alkotnak $\mathbf{R}^n$-ben. Nem alkotnak viszont alteret azon rendezett valós szám-n-esek, melyek komponenseinek összege adott, de 0-tól különböző szám.

2.6. Példa: Tovább általánosítva az előzőeket, a *valós sorozatok* halmaza szintén valós vektortér a szokásos összeadásra és skalárral való szorzásra nézve:

$$(x_n) + (y_n) := (x_n + y_n)$$

$$\lambda \cdot (x_n) := (\lambda x_n)$$

tetszőleges $(x_n),(y_n) \subset \mathbf{R}$ és $\lambda \in \mathbf{R}$ esetén.

Ennek a vektortérnek egy nemtriviális altere pl. a korlátos valós sorozatok halmaza; további, még szűkebb altérek pl. a (valós) konvergens sorozatok és a zérussorozatok halmaza. Nem altér viszont a monoton sorozatok halmaza (két monoton sorozat összege, különbsége nem feltétlen monoton), és nem altér azon konvergens sorozatok halmaza, melyek egy adott, zérustól különböző számhoz tartanak.

2.7. Példa: Még további általánosítással már a "hagyományos" (geometriai) vektorfogalomtól egészen távoleső vektorterekhez jutunk. Legyen $[a,b] \subset \mathbf{R}$ egy adott intervallum: akkor az összes, $[a,b]$-n értelmezett *valós függvények* halmaza vektorteret alkot a szokásos függvényösszeadásra és számmal való szorzásra nézve. Ennek egy fontos altere az $[a,b]$-n *folytonos függvények* halmaza, melyet $C[a,b]$-vel jelölünk. Ennél szűkebb alteret alkotnak a k-szor *folytonosan differenciálható* függvények (k pozitív egész): ezt $C^k[a,b]$-vel jelöljük. További (még szűkebb) altér az $[a,b]$-n értelmezett *polinomok* (azaz az $x \to a_0 + a_1x + a_2x^2 + ... + a_nx^n$ alakú függvények) $\mathcal{P}[a,b]$ halmaza. Ennek egy altere a legfeljebb k-adfokú polinomok $\mathcal{P}_k[a,b]$ halmaza (k nemnegatív egész): ugyanakkor a *pontosan* k-adfokú polinomok halmaza nem alkot alteret (két k-adfokú polinom összege alacsonyabb fokszámú is lehet).

2.2. Lineáris kombináció, lineáris függetlenség

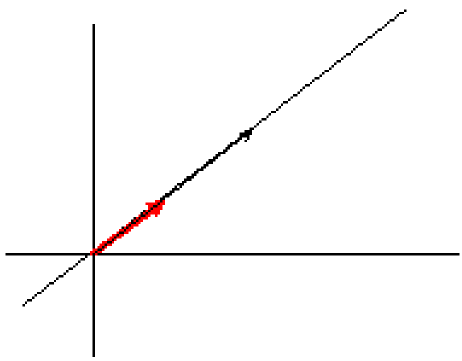
Elnevezés: Legyen X vektortér, $x_1,x_2,...,x_N \in X$ tetszőleges vektorok, $\lambda_1,\lambda_2,...,\lambda_N \in \mathbf{R}$ tetszőleges együtthatók (számok), ahol N tetszőleges pozitív egész. A $\lambda_1x_1 + \lambda_2x_2 + ... + \lambda_Nx_N \in X$ vektort a fenti vektorok egy *lineáris kombinációjának* nevezzük. Ha a lineáris kombináció mindegyik együtthatója zérus, akkor azt *triviális* lineáris kombinációnak, ha legalább egy együttható zérustól különbözik, akkor azt *nemtriviális* lineáris kombinációnak nevezzük.

A következő állítás azon múlik, hogy lineáris kombinációk lineáris kombinációja nyilvánvalóan maga is lineáris kombináció:

2.1. Állítás: Legyen X vektortér, $A \subset X$ tetszőleges részhalmaz. Akkor az A-beli vektorok összes lineáris kombinációinak halmaza alteret alkot X-ben. Ezt az A halmaz *lineáris burkának* vagy az A *halmaz által generált altérnek* nevezzük és $[A]$-val jelöljük.

2.1. Következmény: $[A]$ a *legszűkebb* olyan altér, mely A-t tartalmazza. Valóban, minden A-t tartalmazó altérnek tartalmaznia kell az összes A-beli vektorok lineáris kombinációját is, azaz az $[A]$ altér összes elemét.

2.8. Példa: (lineáris burok $\mathbf{R}^2$-ben): Egyetlen (az origótól különböző) pont lineáris burka a pontot az origóval összekötő egyenes. Két olyan pont lineáris burka, amelyik az origóval nem esik egy egyenesbe, megegyezik a teljes síkkal.



2.3. ábra. *Egypontú halmaz lineáris burba a síkon*

2.9. Példa: (lineáris burok $\mathbf{R}^5$-ben): Az (1,0,0,0,0), (0,3,0,0,0) és (2,2,0,0,0) vektorok lineáris burka az $\{(x,y,0,0,0) \in \mathbf{R}^5 : x,y \in \mathbf{R}\}$.

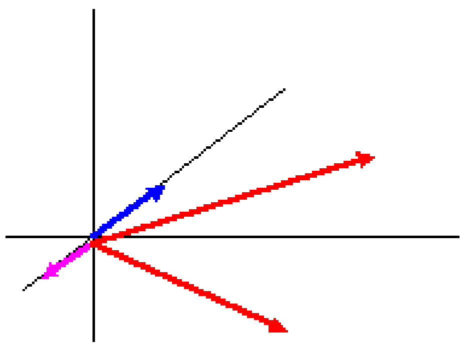
2.10. Példa: (lineáris burok a polinomok terében): A 2, $2x$ és a $-5x^2$ kifejezésekkel definiált polinomok lineáris burka a legfeljebb másodfokú polinomok alterével egyezik.

Lineáris függetlenség:

Legyen X vektortér. Azt mondjuk, hogy az $x \in X$ vektor *lineárisan függ az $x_1, x_2, \ldots, x_N \in X$ vektoroktól*, ha előáll azok valamilyen lineáris kombinációjaként. Azt mondjuk, hogy az X-beli vektorok egy véges rendszere *lineárisan összefüggő*, ha van köztük olyan vektor, mely lineárisan függ a többitől, ill. *lineárisan független*, ha közülük egyik vektor sem függ lineárisan a többitől.

Nyilvánvaló, hogy a **0** zérusvektor minden vektortól lineárisan függ. A következő példák állításai egyszerű meggondolásokkal adódnak:

2.11. Példa: (lineáris függetlenség $\mathbf{R}^2$-ben): Két olyan pont helyvektora, mely az origóval egy egyenesbe esik, lineárisan összefüggő, ellenkező esetben lineárisan független. Három vektor $\mathbf{R}^2$-ben mindig lineárisan összefüggő.



2.4. ábra. *Két síkvektor lineáris függetlensége ill. összefüggősége*

2.12. Példa: (lineáris függetlenség $\mathbf{R}^4$-ben): A (0,1,0,0) és a (0,0,0,3) vektorok lineárisan függetlenek. Ugyanakkor a (4,5,6,0), (1,1,3,0), és a (2,3,0,0) vektorok lineárisan összefüggők, mert $(4,5,6,0) = 2 \cdot (1,1,3,0) + (2,3,0,0)$.

2.13. Példa: (lineáris függetlenség a polinomok terében): Az $x^3 + 2x^5$ ötödfokú polinom nem függ lineárisan másodfokú polinomok semmilyen rendszerétől. Az $1 + x$, $x + x^2$, $-1 + x^2$ polinomok lineárisan összefüggők, de az $1 + x$, $x + x^2$, $-1 + x^3$ polinomok már lineárisan függetlenek.

Látni kell azonban, hogy bonyolultabb esetekben a lineáris függetlenség és összefüggőség ellenőrzése nagyon fáradságos lehet. Ezt egyszerűsíti le a következő tétel, mely szerint elég a zérusvektort megpróbálni előállítani a szóbanforgó vektorok lineáris kombinációjaként:

2.1. Tétel: : Legyen X vektortér. Az $x_1,x_2,...,x_N \in X$ vektorok pontosan akkor lineárisan összefüggők, ha létezik olyan nemtriviális lineáris kombinációjuk, mely a zérusvektorral egyenlő. Másszóval, a fenti vektorok pontosan akkor lineárisan függetlenek, ha *csak* a triviális lineáris kombinációjuk egyenlő a zérusvektorral, azaz $\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_N x_N = \mathbf{0}$ csak úgy lehetséges, ha $\lambda_1 = \lambda_2 = ... = \lambda_N = 0$.

Bizonyítás: Legyenek az $x_1,x_2,...,x_N \in X$ vektorok lineárisan összefüggők: az általánosság csorbítása nélkül feltehető, hogy épp x_1 fejezhető ki a többi lineáris kombinációjaként: $x_1 = \lambda_2 x_2 + ... + \lambda_N x_N$. Ekkor a zérusvektor előáll $x_1,x_2,...,x_N$ nemtriviális lineáris kombinációjaként, hiszen $x_1 - \lambda_2 x_2 - ... - \lambda_N x_N = \mathbf{0}$. Megfordítva, tegyük fel, hogy $\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_N x_N = \mathbf{0}$ egy nemtriviális lineáris kombináció. Akkor valamelyik λ_k biztosan nem zérus, feltehető, hogy épp $\lambda_1 \neq 0$. Akkor x_1 kifejezhető a többi vektor lineáris kombinációjaként, mert $x_1 = -\frac{\lambda_2}{\lambda_1} x_2 - ... - \frac{\lambda_N}{\lambda_1} x_N$.

A tétel értelmében tehát az $x_1,x_2,...,x_N \in X$ vektorok lineáris függetlenségének eldöntése esetén elegendő a $\lambda_1,\lambda_2,...,\lambda_N$ együtthatókra felírt $\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_N x_N = \mathbf{0}$ egyenletet vizsgálni. Ha találunk olyan megoldást is, hogy valamelyik együttható zérustól különbözik, akkor a szóbanforgó vektorok lineárisan összefüggők, ha ilyen nincs, akkor lineárisan függetlenek. Ily módon a lineáris függetlenség kérdését egy speciális egyenlet megoldhatóságának problémájára vezettük vissza: ilyen problémákkal a következő fejezetben részletesen is foglalkozunk.

2.3. Vektorterek bázisa, dimenziója

Bázis, dimenzió:

Legyen X vektortér. Az $A \subset X$ részhalmazt az X vektortér egy *bázisának* nevezzük, ha lineárisan független, és az egész teret generálja, azaz [A] = X. A bázis számosságát az X vektortér *dimenziójának* nevezzük. Megállapodunk abban, hogy a triviális $\{\mathbf{0}\}$ alteret 0-dimenziósnak nevezzük.

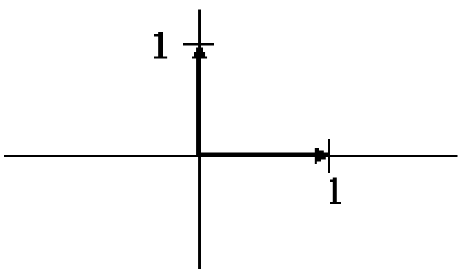
Egyáltalán nem nyilvánvaló, hogy ilyen halmaz létezik. A következő, igen mély tétel azonban pozitívan válaszol erre:

2.2. Tétel: Minden, a triviális $\{\mathbf{0}\}$-tól különböző vektortérnek létezik bázisa. Egy adott vektortér minden bázisa azonos számosságú, azaz a dimenzió egyértelműen meghatározott.

A tételből nyomban adódik, hogy véges, pl. n-dimenziós vektortérben bármely $N > n$ számú vektor lineárisan összefüggő. Ellenkező esetben ui. volna egy N-dimenziós altere, így az egész tér dimenziója is legalább N volna.

2.14. Példa: **R** egydimenziós, bármely nemnulla szám mint egyelemű halmaz, bázist alkot. **C** mint valós vektortér, kétdimenziós. Egy bázisa pl: $\{1,i\}$. Egy másik bázisa: $\{1+i, 1-i\}$

2.15. Példa: $\mathbf{R}^2$ kétdimenziós, egy bázisa pl. $\{(1,0),(0,1)\}$ (*standard bázis*. Általában, bármely két pont, mely az origóval nem esik egy egyenesbe, bázist alkot a síkon.



2.5. ábra. *A sík standard bázisa*

2.16. Példa: $\mathbf{R}^n$ n-dimenziós, egy bázisa pl. a $(1,0,0,...,0)$, $(0,1,0,...,0)$, ..., $(0,0,...,0,1)$ vektorrendszer (ezt *standard bázisnak* nevezzük).

2.17. Példa: A polinomok tere végtelen dimenziós. A legfeljebb k-adfokú polinomok altere $(k+1)$-dimenziós, egy bázisát az $1,x,x^2,...,x^k$ alappolinomok alkotják.

2.4. Skaláris szorzat és norma $\mathrm{R^n}$-ben

Skaláris szorzat:

Az $x := (x_1,x_2,...,x_n) \in \mathbf{R}^n$, $y := (y_1,y_2,...,y_n) \in \mathbf{R}^n$ vektorok *skaláris szorzatának* a következő számot nevezzük:

$$\langle x,y\rangle := \sum_{k=1}^{n} x_k y_k = x_1y_1 + x_2y_2 + ... + x_ny_n$$

Norma:

Az $x \in \mathbf{R}^n$ vektor *normája* vagy *abszolút értéke*:

$$||x|| := \sqrt{\langle x,x\rangle} = \sqrt{\sum_{k=1}^{n} x_k^2} = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$$

Távolság:

Az $x,y \in \mathbf{R}^n$ vektorok *távolságának* pedig az $||x-y||$ számot nevezzük.

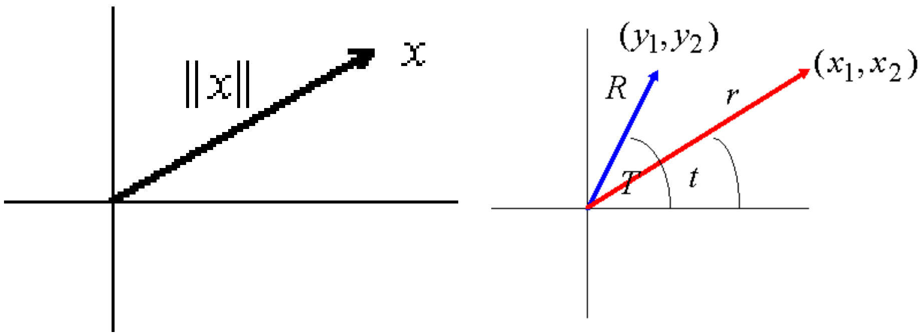
Következésképp $||x||$ az x vektor távolsága a $\mathbf{0}$ zérusvektortól.

Ezekkel a fogalmakkal a már ismert Cauchy-egyenlőtlenség az alábbi tömör alakba írható:

$$|\langle x,y\rangle| \leq ||x|| \cdot ||y||$$

A skaláris szorzatnak és a normának síkbeli vektorok esetén szemléletes jelentése van. Legyenek $x = (x_1,x_2)$, $y = (y_1,y_2) \in \mathbf{R}^2$ tetszőleges vektorok. Kifejezve a vektorok koordinátáit a vektorok hosszával és irányszögével:

$$x_1 = r\cos t,\ x_2 = r\sin t, \quad y_1 = R\cos T,\ y_2 = R\sin T,$$



2.6. ábra. *A norma és a skaláris szorzat szemléltetése a síkon*

a két vektor skaláris szorzatára a következő kifejezés adódik:

$$\langle x,y\rangle = Rr\cos T\cos t + Rr\sin T\sin t =$$

$$= Rr\cos(T-t) = ||x||\cdot||y||\cdot\cos\phi,$$

ahol ϕ a két vektor által bezárt szög.

Az alábbiakban összefoglaljuk a skaláris szorzat alapvető tulajdonságait. Ezek a definícióból azonnal adódnak, és azt mutatják, hogy a skaláris szorzattal a közönséges szorzáshoz hasonló módon számolhatunk:

2.2. Állítás: Tetszőleges $x,y,z \in \mathbf{R}^n$ vektorok és $\lambda \in \mathbf{R}$ esetén:
$\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$
$\langle x,y\rangle = \langle y,x\rangle$
$\langle \lambda x,y\rangle = \lambda\cdot\langle x,y\rangle$
$\langle x,x\rangle \geq 0$, és $\langle x,x\rangle = 0$ pontosan akkor teljesül, ha $x = \mathbf{0}$.

A norma legfontosabb tulajdonságai pedig a következők:

2.3. Állítás: Tetszőleges $x,y \in \mathbf{R}^n$ vektorok és $\lambda \in \mathbf{R}$ esetén:
$||x|| \geq 0$, és $||x|| = 0$ pontosan akkor teljesül, ha $x = \mathbf{0}$
$||\lambda x|| = |\lambda|\cdot||x||$
$||x+y|| \leq ||x|| + ||y||$ (*háromszög-egyenlőtlenség*)

Ezek közül csak a háromszög-egyenlőtlenség nem nyilvánvaló. Használva a skaláris szorzat előzőekben összefoglalt azonosságait:

$$||x+y||^2 = \langle x+y,x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle =$$

$$= ||x||^2 + 2\langle x,y\rangle + ||y||^2$$

A jobb oldal a Cauchy-egyenlőtlenséggel becsülhető felülről, innen:

$$||x+y||^2 \le ||x||^2 + 2||x|| \cdot ||y|| + ||y||^2,$$

ahonnan a háromszög-egyenlőtlenség már adódik.

Érdemes külön is megjegyezni a fenti bizonyításban levezetett azonosságot:

2.4. Állítás: Tetszőleges $x,y \in \mathbf{R}^n$ vektorok esetén:

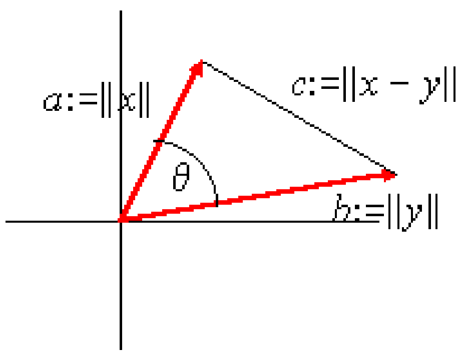
$$||x+y||^2 = ||x||^2 + 2\langle x,y\rangle + ||y||^2,$$

melyhez teljesen hasonlóan adódik az is, hogy:

$$||x-y||^2 = ||x||^2 - 2\langle x,y\rangle + ||y||^2,$$

Ez utóbbi azonosság síkbeli vektorok esetén a jól ismert *koszinusztételt* adja (ld. a 2. ábrát). Az ábra jelöléseivel, a háromszög a,b ill. c oldalának hossza $||x||,||y||$ ill. $||x-y||$. Mivel pedig $\langle x,y\rangle = ||x|| \cdot ||y|| \cdot \cos\theta$, azért innen:

$$c^2 = a^2 - 2ab\cos\theta + b^2.$$



2.7. ábra. *Jelölések a koszinusztételhez*

A későbbiekben túlnyomórészt a skaláris szorzatnak és a normának nem a definícióját használjuk, hanem az előző állításokban összefoglalt tulajdonságokat. Ilyen tulajdonságú $\langle .,. \rangle$ és $||.||$ függvények bevezetése más vektorterekben is lehetséges. Így pl. láttuk, hogy az $[a,b]$ korlátos és zárt intervallumon folytonos függvények $C[a,b]$ halmaza vektortér a szokásos függvényműveletekre nézve: itt az $\langle f,g \rangle := \int_a^b fg$ előírással skaláris szorzatot definiálunk, mely rendelkezik a **??**. Állításban összefoglalt tulajdonságokkal. Így jutunk az *euklideszi terek* fogalmához (skaláris szorzattal ellátott vektorterek). Egy euklideszi térben az $||x|| := \sqrt{\langle x,x \rangle}$ definícióval mindig lehet normát definiálni, mely rendelkezik a **??**. Állításban összefoglalt tulajdonságokkal. Azonban normát egyéb módon is lehet bevezetni, skaláris szorzat nélkül. Így pl. az említett $C[a,b]$ függvénytérben az $||f|| := \max\{|f(x)| : x \in [a,b]\}$ előírás is könnyen ellenőrizhetően normát ad, melyről viszont megmutatható, hogy nem származtatható semmilyen skaláris szorzatból. A normával ellátott vektortereket *normált tereknek* nevezzük. Az euklideszi és normált terek tanulmányozása meghaladja a jegyzet kereteit: a továbbiakban e terek közül egyedül a véges dimenziós $\mathbf{R}^n$ terekkel foglalkozunk.

2.5. Ortogonalitás, ortogonális vetület

Ortogonális vektorok:

Azt mondjuk, hogy az $x,y \in \mathbf{R}^n$, vektorok *merőlegesek* vagy *ortogonálisak*, ha $\langle x,y \rangle = 0$. Ezt a tényt így jelöljük: $x \perp y$.

Az ortogonalitás a geometriai merőlegességfogalom általánosítása, mert sík- ill. térbeli vektorok esetén az előző szakasz szerint két nemzérus vektor skaláris szorzata pontosan akkor zérus, ha az általuk bezárt szög koszinusza 0, azaz, ha a két vektor merőleges.

Nyilvánvaló, hogy egyedül a zérusvektor ortogonális saját magára, mert egy vektor önmagával vett skaláris szorzata megegyezik saját normájának négyzetével.

A **??**. Állítás azonnali következményeképp:

2.5. Állítás: (Pitagorász tétele): Tetszőleges $x,y \in \mathbf{R}^n$, $x \perp y$ vektorok esetén:

$$||x+y||^2 = ||x||^2 + ||y||^2$$

Az állítás könnyen általánosítható kettőnél több vektorra:

2.2. Következmény: Tetszőleges $x_1,x_2,...,x_m \in \mathbf{R}^n$, $(m \leq n)$ páronként ortogonális vektor esetén:

$$||\sum_{j=1}^{m} x_j||^2 = \sum_{j=1}^{m} ||x_j||^2$$

Bizonyítás: Valóban,

$$||\sum_{j=1}^{m} x_j||^2 = \langle \sum_{j=1}^{m} x_j, \sum_{k=1}^{m} x_k \rangle =$$

$$= \sum_{j=1}^{m} \sum_{k=1}^{m} \langle x_j, x_k \rangle = \sum_{k=1}^{m} \langle x_k, x_k \rangle = \sum_{k=1}^{m} ||x_k||^2.$$

Az ortogonális vektoroknak – csakúgy mint a sík- ill. a térvektorok esetében – kitüntetett szerepük van. A következő állításokban ezeket foglaljuk össze.

2.6. Állítás: Minden $x_1,x_2,...,x_m \in \mathbf{R}^n$ $(m \leq n)$ páronként ortogonális nemzérus vektorokból álló vektorrendszer lineárisan független.

Bizonyítás: Ha ui. $\sum_{j=1}^{m} \lambda_j x_j = \mathbf{0}$, akkor ezen egyenlőség mindkét oldalát x_k-val skalárisan szorozva és a páronkénti ortogonalitást felhasználva kapjuk, hogy $\lambda_k ||x_k||^2 = 0$, ahonnan $\lambda_k = 0$, $k = 1,2,...,m$-re. Tehát $x_1,x_2,...,x_m$ valóban lineárisan függetlenek.

2.7. Állítás: Ha egy $x \in \mathbf{R}^n$ vektor ortogonális egy $e_1,e_2,...,e_m \in \mathbf{R}^n$ generátorrendszer minden elemére, akkor szükségképp $x = \mathbf{0}$.

Bizonyítás: Ha ui. $x = \sum_{j=1}^{n} \lambda_j e_j$, alakú, akkor az egyenlőség mindkét oldalát skalárisan szorozva x-szel, kapjuk, hogy $||x||^2 = 0$, azaz $x = \mathbf{0}$.

Használva a skaláris szorzatnak a **??**. Állításban összefoglalt tulajdonságait, nyomban adódik, hogy:

2.8. Állítás: Egy tetszőleges $M \subset X$ halmaz összes elemére ortogonális vektorok alteret alkotnak X-ben.

Ezt az alteret M halmaz *ortogonális kiegészítő alterének* nevezzük, és az $M^\perp$ szimbólummal jelöljük.

Ezzel a fogalommal a **??**. Állítás röviden úgy fogalmazható meg, hogy egy tetszőleges $e_1,e_2,...,e_m \in \mathbf{R}^n$ generátorrendszer ortogonális kiegészítő altere a triviális $\{\mathbf{0}\}$ altér.

Ortogonális bázis, ortonormált bázis:

Az $e_1,e_2,...,e_n \in \mathbf{R}^n$ bázist *ortogonális bázisnak* nevezzük, ha $e_k \perp e_j$ minden $k \neq j$ esetén. Az ortogonális bázist *ortonormáltnak* nevezzük, ha még $||e_k|| = 1$ is teljesül minden $k = 1,2,...,n$-re.

Nyilvánvaló, hogy pl. a standard bázis egyúttal ortonormált bázis is $\mathbf{R}^n$-ben.

2.3. Tétel: Bármely $\{\mathbf{0}\} \neq X_0 \subset \mathbf{R}^n$ altérnek van ortonormált bázisa.

Bizonyítás: Legyen $e_1 \in X_0$ egy tetszőleges, 1 normájú vektor. Ezekután válasszunk egy 1 normájú e_2 vektort az $\{e_1\}^\perp \cap X_0$ altérből, majd egy 1 normájú e_3 vektort az $\{e_1,e_2\}^\perp \cap X_0$ altérből, és így tovább. Ily módon páronként ortogonális, ezért lineárisan független X_0-beli vektorok rendszeréhez jutunk. Az eljárás véget ér, ha az $\{e_1,e_2,...e_m\}^\perp \cap X_0$ altér már nem tartalmaz 1 normájú vektort, azaz a triviális $\mathbf{0}$ altérrel egyenlő. Ekkor $\{e_1,e_2,...e_m\}$ generálja is az X_0 alteret. Valóban, ha valamely $x \in X_0$ vektor nem lenne előállítható $e_1,...,e_m$ lineáris kombinációjaként, akkor az $x - \sum_{j=1}^m \langle x,e_j\rangle e_j$ vektor egy nemzérus eleme lenne az $\{e_1,e_2,...e_m\}^\perp \cap X_0$ altérnek, mivel mindegyik e_k-ra ortogonális ($k = 1,2,...,m$):

$$\langle x - \sum_{j=1}^m \langle x,e_j\rangle e_j, e_k\rangle = \langle x,e_k\rangle - \sum_{j=1}^m \langle x,e_j\rangle\langle e_j,e_k\rangle =$$

$$= \langle x,e_k\rangle - \langle x,e_k\rangle = 0$$

Tehát $\{e_1,e_2,...e_m\}$ egy generátorrendszer X_0-ban, és mivel páronként ortogonális, azért lineárisan független is, azaz bázist alkot X_0-ban. A konstrukció miatt pedig e bázis elemei mind 1 normájúak, tehát a bázis ortonormált.

Az ortonormált bázisok kitüntetett szerepét világítja meg a következő példa. Legyen $e_1,e_2,...,e_n \in \mathbf{R}^n$ egy tetszőleges (nem feltétlen ortogonális) bázis, és $x \in \mathbf{R}^n$ tetszőleges vektor. Ha x-et elő akarjuk állítani az $e_1,e_2,...,e_n$ bázisvektorok lineáris kombinációjaként:

$$x = \lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_n e_n,$$

akkor ez a $\lambda_1,...,\lambda_n$ együtthatókra egy n-ismeretlenes algebrai egyenletrendszer megoldását jelenti. A helyzet lényegesen egyszerűsödik, ha az $e_1,e_2,...,e_n$ bázis ortonormált. Ekkor ui. érvényes a következő tétel:

2.4. Tétel: Legyen $e_1,e_2,...,e_n \in \mathbf{R}^n$ egy ortonormált bázis, és $x \in \mathbf{R}^n$ tetszőleges vektor, akkor:

$$x = \sum_{j=1}^n \langle x,e_j\rangle e_j$$

Bizonyítás: Jelölje y a jobb oldali vektort: $y := \sum_{j=1}^{n} \langle x,e_j \rangle e_j$. Meg kell mutatnunk, hogy $x = y$. Tetszőleges $k = 1,2,...,n$ indexre:

$$\langle x - y,e_k \rangle = \langle x,e_k \rangle - \sum_{j=1}^{n} \langle x,e_j \rangle \langle e_j,e_k \rangle =$$

$$= \langle x,e_k \rangle - \langle x,e_k \rangle \cdot \langle e_k,e_k \rangle = 0,$$

Tehát az $x - y$ vektor ortogonális az $e_1,e_2,...,e_n$ bázis minden elemére, ezért szükségképp $x - y = \mathbf{0}$ (**??**. Állítás).

Következésképp az együtthatók egyenletmegoldás nélkül, egy-egy skaláris szorzat kiszámításával adódnak.

A fenti előállítás lehetővé teszi az elemi geometriából jól ismert *merőleges vetület* fogalmának általánosítását:

2.5. Tétel: Legyen $X_0 \subset \mathbf{R}^n$ egy tetszőleges altér. Akkor minden $x \in \mathbf{R}^n$ vektor egyértelműen előáll $x = x_0 + x_0^\perp$ alakban, ahol $x_0 \in X_0$ és $x_0^\perp \in X_0^\perp$. Ezt az x_0 vektort az x vektornak az X_0 altérre vett *ortogonális vetületének* nevezzük.

Bizonyítás: Legyen $e_1,e_2,...,e_m$ egy ortonormált bázis X_0-ban (ilyen van, ld. a **??**. Tételt), és jelölje:

$$x_0 := \sum_{j=1}^{m} \langle x,e_j \rangle e_j$$

Akkor az $x^\perp := x - x_0$ valóban ortogonális X_0-ra, mert mindegyik e_k bázisvektorra ortogonális:

$$\langle x - x_0,e_k \rangle = \langle x - \sum_{j=1}^{m} \langle x,e_j \rangle e_j,e_k \rangle =$$

$$= \langle x,e_k \rangle - \sum_{j=1}^{m} \langle x,e_j \rangle \langle e_j,e_k \rangle = \langle x,e_k \rangle - \langle x,e_k \rangle = 0,$$

amivel a kívánt előállítás létezését igazoltuk. Már csak az egyértelműséget kell belátni. Ha $x = x_0 + x_0^\perp$ és $x = y_0 + y_0^\perp$ két olyan felbontás, hogy $x_0,y_0 \in X_0$ és $x_0^\perp,y_0^\perp \in X_0^\perp$, akkor innen $x_0 - y_0 = -(x_0^\perp - y_0^\perp)$ következik. Ámde a bal oldali vektor X_0-beli, míg a jobb oldali $X_0^\perp$-beli, azaz egymásra ortogonálisak. Ezért csak úgy lehetnek egyenlők, ha mindketten a $\mathbf{0}$ zérusvektorral egyenlők, azaz $x_0 = y_0$ és $x_0^\perp = y_0^\perp$. Tehát a tételben szereplő ortogonális felbontás valóban egyértelmű.

Speciálisan, ha X_0 egydimenziós, és egy $\mathbf{0} \neq \mathbf{e} \in \mathbf{R^n}$ vektor generálja, akkor a tételből adódik, hogy egy tetszőleges $x \in \mathbf{R}^n$ vektor e irányú ortogonális vetülete az $\frac{\langle x,e \rangle e}{||e||^2}$ vektor (ui. a $\frac{e}{||e||}$ vektor normája épp 1).

A tétel másik következménye, hogy tetszőleges $X_0 \subset \mathbf{R}^n$ altér esetén X_0 és az $X_0^\perp$ alterek dimenzióinak összege éppen n. Valóban, vegyünk fel mindkét altérben egy $e_1,...,e_m \in X_0$ ill. $f_1,...,f_k \in X_0^\perp$ ortonormált bázist, akkor X_0 m-dimenziós, és $X_0^\perp$ k-dimenziós. Ezek egyesítése, azaz az $e_1,...,e_m,f_1,...,f_k$ vektorrendszer továbbra is páronként ortogonális (ezért lineárisan független) vektorokból áll, továbbá a **??**. Tétel értelmében generálják is az $\mathbf{R}^n$ teret, így bázist alkotnak $\mathbf{R}^n$-ben. Ezért e bázis elemszáma épp n, tehát valóban, $n = m + k$.

Végezetül megmutatjuk, hogy az ortogonális vetület rendelkezik egyfajta minimumtulajdonsággal, mely a két- és háromdimenziós terekben az elemi geometriából már jól ismert:

2.6. Tétel: Legyen $X_0 \subset \mathbf{R}^n$ egy tetszőleges altér, $x \in \mathbf{R}^n$ pedig egy tetszőleges vektor. Jelölje x_0 az x vektor X_0-ra vett ortogonális vetületét. Akkor x_0 az x vektorhoz legközelebb eső X_0-beli vektor, azaz minden $y \in X_0$ esetén $||x - x_0|| \leq ||x - y||$.

Bizonyítás: Tetszőleges $y \in X_0$ esetén nyilván $x - y = (x_0 - y) + (x - x_0)$. Az első zárójeles tag X_0-beli, míg a második a **??**. Tétel értelmében $X_0^\perp$-beli. A Pitagorász-tétel miatt ezért $||x - y||^2 = ||x_0 - y||^2 + ||x - x_0||^2$. A jobb oldal akkor minimális, ha $y = x_0$, innen az állítás adódik.

2.6. Feladatok

1. Vektorteret alkotnak-e a szokásos függényműveletekre nézve az alábbi $\mathbf{R} \to \mathbf{R}$ típusú függvények?

(a) a korlátos függvények

(b) a monoton függvények

(c) a 2π-periodikus függvények

(d) a trigonometrikus polinomok, azaz az $x \to a_0 + a_1 \cos x + a_2 \cos 2x + \ldots + a_n \cos nx + b_1 \sin x + b_2 \sin 2x + \ldots + b_n \sin nx$ alakú függvények ($a_0,\ldots,a_n,b_1,\ldots,b_n \in \mathbf{R}$, $n \in \mathbf{N}$)

(e) azon folytonos függvények, melyek egy adott x_0 helyen zérus értéket vesznek fel

(f) a nemnegatív függvények

2. A valós sorozatok vektorterében alteret alkotnak-e az alábbi sorozatok?

(a) a felülről korlátos sorozatok

(b) az alulról korlátos sorozatok

(c) a Cauchy-sorozatok

(d) a $+\infty$-be vagy a $-\infty$-be tartó sorozatok

3. Igazoljuk, hogy a geometriai tér semmilyen korlátos részhalmaza nem lehet altér (az origó mint egy pontból álló triviális altér kivételével).

4. Lineárisan függetlenek-e az alábbi formulákkal megadott függvények?

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

(b) e^x, sh x, ch x, 1, e^{-2x}

(c) $\log(1 + e^x)$, x, 1, $\log \sqrt{\frac{e^{-x}+1}{e^x+1}}$

5. Állítsuk elő az $x \to \sin^3 x$ függvényt $\sin x$, $\sin 2x$ és $\sin 3x$ lineáris kombinációjaként.

6. Lineárisan függetlenek-e az $\mathbf{a} := (1,-2,0,5)$, $\mathbf{b} := (1,3,-1,0)$, $\mathbf{c} :=$

$(1,-1,1,-1)$, $\mathbf{d} := (0,7,1,-17)$ $\mathbf{R}^4$-beli vektorok? Ha igen, bizonyítsuk ezt be. Ha nem, állítsuk elő valamelyiket a többi lineáris kombinációjaként.

7. Határozzuk meg adott $n \in \mathbf{N}$ mellett a trigonometrikus polinomok, azaz az $x \to a_0 + a_1 \cos x + a_2 \cos 2x + \ldots + a_n \cos nx + b_1 \sin x + b_2 \sin 2x + \ldots + b_n \sin nx$ alakú függvények alkotta vektortér dimenzióját, és egy bázisát.

8. Tekintsük $\mathbf{R}^n$-nek ($n > 2$) azon vektorok alkotta alterét, melyek ortogonálisak az $(1,-1,0,0,\ldots,0)$, $(-1,2,0,0,\ldots,0)$, $(1,1,0,0,\ldots,0)$ vektorok mindegyikére. Határozzuk meg ennek az altérnek a dimenzióját.

9. Határozzuk meg az $a := (1,-1,1,-1,1,-1,\ldots) \in \mathbf{R}^n$ vektornak az $e := (1,1,1,\ldots.,1) \in \mathbf{R}^n$ irányú ortogonális vetületét.

10. Mutassuk meg, hogy ha $x,y \in \mathbf{R}^n$ egyenlő hosszúságú vektorok, akkor $x + y$ és $x - y$ szükségképp ortogonálisak.

11. Igazoljuk, hogy tetszőleges $x = (x_1,x_2,\ldots,x_n) \in \mathbf{R}^n$ vektor esetén:

$$\sum_{k=1}^{n} x_k \leq \sqrt{n} \cdot ||x||.$$

Megoldások

1. (a) igen, (b) nem (két monoton függvény összege nem feltétlen monoton), (c) igen, (d) igen, (e) igen, (f) nem (két nemnegatív függvény különbsége nem feltétlen nemnegatív).

2. (a) nem, (b) nem, (c) igen, (d) nem.

3. Ha $A \subset \mathbf{R}^3$ korlátos, akkor befoglalható egy origó középű, elég nagy R sugarú gömbbe. Ha pedig $\mathbf{x} \in A$, $\mathbf{x} \neq \mathbf{0}$ tetszőleges, akkor $\alpha \cdot \mathbf{x}$ biztosan nincs A-ban, ha $\alpha > \frac{2R}{|\mathbf{x}|}$, így A nem lehet altér.

4. (a) igen, (b) nem (e^x kifejezhető $\operatorname{sh} x$ és $\operatorname{ch} x$ lineáris kombinációjaként), (c) nem (az utolsó kifejezés – egyszerűsítés után – $\left(-\frac{1}{2} \cdot x\right)$-szel egyenlő).

5.
$$\begin{aligned}\sin^3 x = \sin x \sin^2 x = \sin x \frac{1-\cos 2x}{2} &= \\ = \frac{1}{2}\sin x - \frac{1}{2}\sin x \cos 2x &= \\ = \frac{1}{2}\sin x - \frac{1}{4}(\sin 3x + \sin(-x)) &= \\ = \frac{3}{4}\sin x - \frac{1}{4}\sin 3x&\end{aligned}$$
ahol felhasználtuk a $\sin\alpha\cos\beta = \frac{1}{2}(\sin(\alpha+\beta) + \sin(\alpha-\beta))$ azonosságot.

6. Az ilyen típusú feladatok általában úgy oldhatók meg, hogy megpróbáljuk előállítani a zérusvektort a szóban forgó vektorok lineáris kombinációjaként. Ez a lineáris kombináció együtthatóira egy homogén lineáris egyenletrendszert jelent, a kérdés pedig az, hogy van-e ennek nemtriviális megoldása: ha igen, a vektorok lineárisan összefüggők, ha nincs, akkor pedig függetlenek.

Jelen feladat egyszerűbben is kezelhető: vegyük észre, hogy $\mathbf{b} - \mathbf{a} = (0{,}5, -1, -5)$ és $\mathbf{c} - \mathbf{a} = (0{,}1{,}1, -6)$. Ez utóbbi kétszeresét az előbbihez adva épp a $\mathbf{d}$ vektort kapjuk: $\mathbf{d} = 2 \cdot (\mathbf{c} - \mathbf{a}) + \mathbf{b} - \mathbf{a} = -3\mathbf{a} + \mathbf{b} + 2\mathbf{c}$. Tehát a vektorok lineárisan összefüggők.

7. Az 1, $\cos x$, $\cos 2x$, ..., $\cos nx$, $\sin x$, $\sin 2x$, ..., $\sin nx$ kifejezésekkel értelmezett függvények rendszere nyilván generálja a trigonometrikus polinomokat. Lineáris függetlenségük a következőképp látható: ha

$$a_0 + a_1 \cos x + a_2 \cos 2x + ... + a_n \cos nx +$$

$$+ b_1 \sin x + b_2 \sin 2x + ... + b_n \sin nx \equiv 0,$$

akkor ezt az egyenlőséget $\cos kx$-szel szorozva és a $(0,2\pi)$ intervallumon integrálva a bal oldalon a $a_k \cos kx$ tag kivételével minden tag eltűnik, innen szükségképp $a_k = 0$, és hasonlóan, $\sin jx$-szel szorozva és a $(0,2\pi)$ intervallumon integrálva a bal oldalon a $b_j \sin jx$ tag kivételével minden tag eltűnik, innen szükségképp $b_j = 0$ ($k = 0,1,...,n$, $j = 1,2,...,n$). Tehát a fenti függvényrendszer bázis, így a szóban forgó tér dimenziója $(2n + 1)$.

8. Az első két vektor lineárisan független, de a harmadik már lineárisan függ az első kettőtől. Így e három vektor egy *kétdimenziós* alteret generál $\mathbf{R}^n$-ben. A szóbanforgó altér ennek ortogonális kiegészítő altere, így $(n - 2)$-dimenziós.

9.
$$a_e = \frac{\langle a,e\rangle}{||e||^2} e = \frac{1 - 1 + 1 - 1 + 1 ...}{1^2 + 1^2 + ... + 1^2} e = \begin{cases} 0, & \text{ha } n \text{ páros} \\ \frac{1}{n} e, & \text{ha } n \text{ páratlan} \end{cases}$$

10. A skaláris szorzat tulajdonságait használva:

$$\langle x + y, x - y\rangle = \langle x,x\rangle + \langle y,x\rangle - \langle x,y\rangle - \langle y,y\rangle = ||x||^2 - ||y||^2 = 0,$$

azaz $(x + y) \perp (x - y)$.

11. Jelölje $e := (1,1,...,1) \in \mathbf{R}^n$. A Cauchy-egyenlőtlenséget használva:

$$\sum_{k=1}^{n} x_k = \langle x,e\rangle \le ||x|| \cdot ||e|| = \sqrt{n} \cdot ||x||.$$

3. Vektorgeometria

Az előző fejezetben bevezetett fogalmakat és tételeket most speciálisan a két- és háromdimenziós térben alkalmazzuk. Emlékeztetünk rá, hogy a háromdimenziós geometriai térben a nemtriviális alterek az origóra illeszkedő egyenesek és síkok, így a lineáris algebra a geometriai térben az egyenesek és síkok kényelmes leírását teszi lehetővé.

Ebben a fejezetben – történeti okokból – az $\mathbf{R}^2$ és $\mathbf{R}^3$ vektortér elemeit *pontoknak* fogjuk nevezni, míg "vektor" alatt az origóból egy pontba húzott irányított szakaszt értünk (ami az illető pont helyvektora, bár ezek kölcsönösen egyértelműen megfeleltethetők egymásnak). Ezért pl. két *pont* távolságáról ugyanakkor két *vektor* által bezárt szögről fogunk beszélni. Szokásos még – és ebben a fejezetben mi is így teszünk – a pontokat, vektorokat álló, félkövér betűkkel, míg a skalár (szám) paramétereket dőlt, nem vastagított betűkkel jelölni.

A következő szakaszokban mindig feltételezzük, hogy a geometriai síkon ill. térben adott egy rögzített (merőleges) koordinátarendszer, így a sík (tér) pontjai kölcsönösen egyértelműen megfeleltethetők mind $\mathbf{R}^2$ (ill. $\mathbf{R}^3$) elemeinek, mind pedig az illető pontok helyvektorainak. Így pl. egy $\mathbf{x} \in \mathbf{R}^3$ elem alatt egyszerre értünk: (a) egy rendezett számhármast, (b) egy térbeli pontot, ill. (c) egy térbeli helyvektort, ahogy épp a legkényelmesebb.

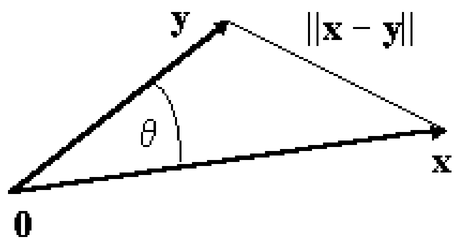
3.1. Síkvektorok, egyenesek a síkon

Az alábbi eredmények minden további meggondolás nélkül adódnak az előző fejezet általános eredményeiből. Mindazonáltal javasoljuk az Olvasónak annak átgondolását, hogy a most következő állítások mely korábbi állításokon múlnak!

3.1. Állítás: Az $\mathbf{x},\mathbf{y} \in \mathbf{R}^2$ pontok távolsága $||\mathbf{x}-\mathbf{y}||$, míg az $\mathbf{x}$ és $\mathbf{y}$ vektorok által bezárt θ szögre teljesül a

$$\cos\theta = \frac{\langle \mathbf{x},\mathbf{y}\rangle}{||\mathbf{x}|| \cdot ||\mathbf{y}||}$$

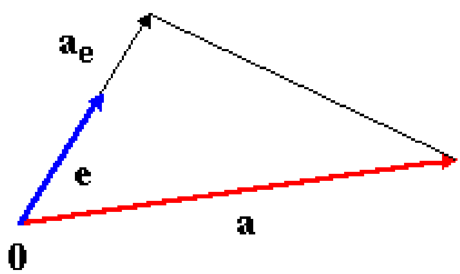
egyenlőség, feltéve, hogy a egyik vektor hossza sem 0 (ld. az ábrát).



3.1. ábra. *Pontok távolsága, vektorok szöge*

3.2. Állítás: Legyenek $\mathbf{a},\mathbf{e} \in \mathbf{R}^2$, $\mathbf{e} \neq \mathbf{0}$. Az $\mathbf{a}$ vektor $\mathbf{e}$ irányú ortogonális vetülete (ld. az ábrát):

$$\mathbf{a_e} = \frac{\langle \mathbf{a},\mathbf{e} \rangle}{||\mathbf{e}||^2} \cdot \mathbf{e}$$



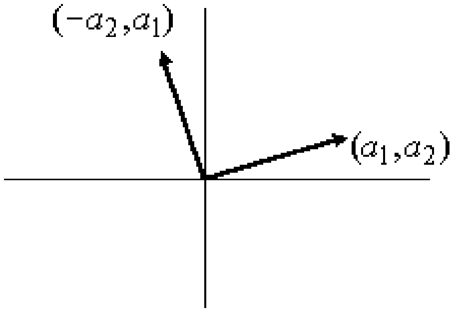
3.2. ábra. *Adott irányú ortogonális vetületvektor*

Amennyiben az $\mathbf{e}$ irányvektor egységnyi hosszúságú, akkor a formula leegyszerűsödik: $\mathbf{a_e} = \langle \mathbf{a},\mathbf{e} \rangle \cdot \mathbf{e}$.

3.3. Állítás: Legyen $\mathbf{a} = (a_1,a_2) \in \mathbf{R}^2$ tetszőleges vektor. Az $\mathbf{a}$ vektor 90^{O}-kal való elforgatottja: $\mathbf{a}^{\perp} = (-a_2,a_1)$.

Következésképp, ha $\mathbf{0} \neq \mathbf{e} \in \mathbf{R}^2$, akkor egy tetszőleges $\mathbf{x} \in \mathbf{R}^2$ vektor az alábbi módon bontható fel egy $\mathbf{e}$ irányú és egy arra merőleges vektor összegeként:

$$\mathbf{x} = \frac{\langle \mathbf{x},\mathbf{e} \rangle}{||\mathbf{e}||^2} \cdot \mathbf{e} + \frac{\langle \mathbf{x},\mathbf{e}^{\perp} \rangle}{||\mathbf{e}||^2} \cdot \mathbf{e}^{\perp}$$



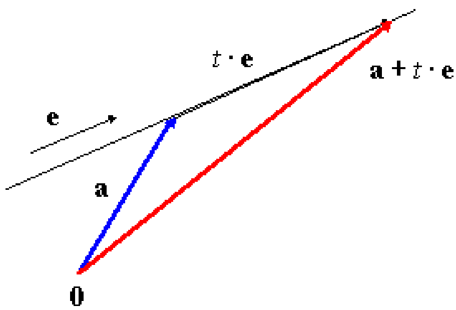
3.3. ábra. *Síkvektor* 90^o*-os elforgatottja*

Amennyiben az $\mathbf{e}$ irányvektor egységnyi hosszúságú (ekkor $\mathbf{e}^\perp$ is az), a formula leegyszerűsödik: $\mathbf{x} = \langle \mathbf{x},\mathbf{e}\rangle \cdot \mathbf{e} + \langle \mathbf{x},\mathbf{e}^\perp\rangle \cdot \mathbf{e}^\perp$.

3.4. Állítás: Legyenek $\mathbf{a},\mathbf{e} \in \mathbf{R}^2$, $\mathbf{e} \neq \mathbf{0}$ tetszőleges vektorok. Akkor az

$$\mathbf{x} = \mathbf{a} + t \cdot \mathbf{e} \qquad (t \in \mathbf{R})$$

pontok egy egyenest alkotnak, mely illeszkedik az $\mathbf{a}$ pontra, és párhuzamos az $\mathbf{e}$ vektorral.



3.4. ábra. *Az egyenes paraméteres vektoregyenletéhez*

Az $\mathbf{e}$ vektort az egyenes *irányvektorának*, a fenti formulát pedig az egyenes *paraméteres vektoregyenletének* nevezzük. Nyilvánvaló, hogy ez az egyenlet nem egyértelműen meghatározott: ugyanannak az egyenesnek $\mathbf{a}$ és $\mathbf{e}$ választásától függően több, különböző alakú egyenlete is lehet.

Ha $\mathbf{a} = (a_1,a_2)$, $\mathbf{e} = (e_1,e_2)$, akkor az egyenes paraméteres vektoregyenlete nyilván ekvivalens az alábbi (skalár) egyenletrendszerrel:

$$x_1 = a_1 + t \cdot e_1$$

$$x_2 = a_2 + t \cdot e_2$$

ahonnan a t paraméter kiküszöbölésével az

$$\frac{x_1 - a_1}{e_1} = \frac{x_2 - a_2}{e_2}$$

nemparaméteres egyenletet nyerjük (feltéve, hogy $e_1,e_2 \neq 0$). Innen x_2-t kifejezve, az egyenes jól ismert iránytényezős egyenletéhez jutunk:

$$x_2 = a_2 + m \cdot (x_1 - a_1),$$

ahol $m = \frac{e_2}{e_1}$, az egyenes iránytényezője (irányszögének tangense).

A jelöléstechnika e téren nem egységes. Egyik fajta jelölési konvenció szerint a pontok koordinátáit indexekkel jelöljük, mint idáig is tettük: $\mathbf{x} = (x_1,x_2)$, $\mathbf{e} = (e_1,e_1)$, s.í.t. A másik elterjedt konvenció, hogy az egyes komponenseket betűkkel különböztetjük meg: az egyenes egy tetszőleges pontjának koordinátái (x,y), egy pontja (a_x,a_y), irányvektora (e_x,e_y). Ezzel a jelöléstechnikával az egyenes paraméteres egyenletrendszerének alakja:

$$x = a_x + t \cdot e_x$$

$$y = a_y + t \cdot e_y$$

Jegyezzük még meg, hogy az egyenes paraméteres egyenletrendszeréből azonnal leolvashatók az egyenes egy pontjának koordinátái és az irányvektor koordinátái is. Ugyanez vonatkozik a fentebb leírt nemparaméteres alakra is.

3.1. Példa: Az $(1,2)$ ponton átmenő, $(3,-1)$ irányú egyenes paraméteres egyenletrendszere:

$$x = 1 + 3t$$

$$y = 2 - t$$

3.2. Példa: Határozzuk meg azon egyenes egyenletét, mely illeszkedik a $(-1,-3)$ pontra, és merőleges az

$$x = 5 - 2t$$

$$y = -2 + 3t$$

egyenesre!

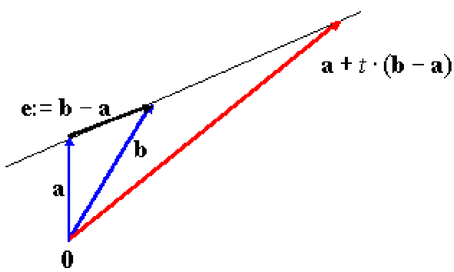
Megoldás: A adott egyenes irányvektora $(-2{,}3)$. Egy erre merőleges vektor: $(3{,}2)$, ez jó lesz a keresett egyenes irányvektorának. Innen a keresett egyenes egyenletrendszere:

$$x = -1 + 3t$$
$$y = -3 + 2t$$

A gyakorlatban sokszor előfordul, hogy az egyenes irányvektora explicite nem adott, viszont két különböző pl. **a** és **b** pontja is ismert. Ekkor irányvektor gyanánt az $\mathbf{e} := \mathbf{b} - \mathbf{a}$ vektor nyilván megfelelő, innen nyerjük a két ponton átmenő egyenes egyenletét:

3.5. Állítás: Legyenek **a**,**b** $\in \mathbf{R}^2$ tetszőleges síkbeli pontok. Az **a** és **b** pontokon átmenő egyenes paraméteres vektoregyenlete:

$$\mathbf{x} = \mathbf{a} + t \cdot (\mathbf{b} - \mathbf{a}) \qquad (t \in \mathbf{R}).$$



3.5. ábra. *Két pont általmeghatározott egyenes*

Megjegyezzük, hogy ha a t paraméter csak a $[0{,}1]$ intervallumot futja be, akkor a fenti formulával definiált **x** pontok épp az **a** és **b** végpontú szakasz pontjait futják be.

3.3. Példa: Az $(1, -2)$ és a $(6,5)$ pontokon átmenő egyenes paraméteres egyenletrendszere:

$$x = 1 + 5t$$

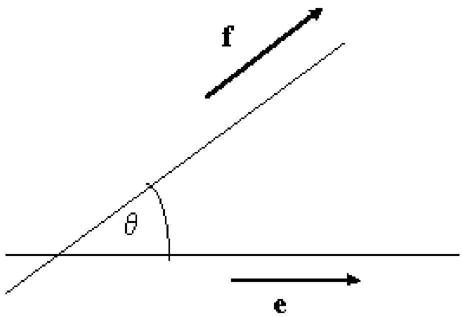
$$y = -2 + 7t,$$

de ugyancsak ezt az egyenest határozza meg az alábbi egyenletrendszer is:

$$x = 6 + 5t$$

$$y = 5 + 7t.$$

Két egyenes szögén az irányvektoraik által bezárt θ hegyesszöget értjük. Ha a két egyenes paraméteres vektoregyenlete $\mathbf{x} = \mathbf{a} + t \cdot \mathbf{e}$ ill. $\mathbf{x} = \mathbf{b} + \tau \cdot \mathbf{f}$, akkor nyilván: $\cos\theta = \frac{|\langle \mathbf{e},\mathbf{f}\rangle|}{||\mathbf{e}||\cdot||\mathbf{f}||}$.



3.6. ábra. *Két egyenes által bezárt szög*

3.2. Térvektorok, egyenesek a térben

Az előző szakasz meggondolásai (a vektor elforgatását kivéve) a háromdimenziós geometriai térben is minden további nélkül alkalmazhatók. Így jutunk a térbeli vektorokra ill. térbeli egyenesekre vonatkozó alapvető állításokhoz, melyet az alábbiakban foglalunk össze. Megjegyezzük, hogy a fejezet hátralevő részében a skaláris szorzat ill. a norma értelemszerűen az $\mathbf{R}^3$-beli skaláris szorzatot és normát jelenti. Szokásos még a standard bázis elemeinek alábbi jelölése: $\mathbf{i} := (1,0,0)$, $\mathbf{j} := (0,1,0)$, $\mathbf{k} := (0,0,1)$.

3.6. Állítás: Az $\mathbf{x},\mathbf{y} \in \mathbf{R}^3$ pontok távolsága $||\mathbf{x}-\mathbf{y}||$, míg az $\mathbf{x}$ és $\mathbf{y}$ vektorok által bezárt θ szögre teljesül a

$$\cos\theta = \frac{\langle \mathbf{x},\mathbf{y} \rangle}{||\mathbf{x}|| \cdot ||\mathbf{y}||}$$

egyenlőség, feltéve, hogy a egyik vektor hossza sem 0.

3.7. Állítás: Legyenek $\mathbf{a},\mathbf{e} \in \mathbf{R}^3$, $\mathbf{e} \neq \mathbf{0}$. Az $\mathbf{a}$ vektor $\mathbf{e}$ irányú ortogonális vetülete:

$$\mathbf{a}_\mathbf{e} = \frac{\langle \mathbf{a},\mathbf{e} \rangle}{||\mathbf{e}||^2} \cdot \mathbf{e}$$

Amennyiben az $\mathbf{e}$ irányvektor egységnyi hosszúságú, akkor a formula leegyszerűsödik: $\mathbf{a}_\mathbf{e} = \langle \mathbf{a},\mathbf{e} \rangle \cdot \mathbf{e}$.

3.8. Állítás: Legyenek $\mathbf{a},\mathbf{e} \in \mathbf{R}^3$, $\mathbf{e} \neq \mathbf{0}$ tetszőleges vektorok. Akkor az

$$\mathbf{x} = \mathbf{a} + t \cdot \mathbf{e} \qquad (t \in \mathbf{R})$$

pontok egy egyenest alkotnak, mely illeszkedik az $\mathbf{a}$ pontra, és párhuzamos az $\mathbf{e}$ vektorral. Az $\mathbf{e}$ vektort az egyenes *irányvektorának*, a fenti formulát pedig az egyenes *paraméteres vektoregyenletének* nevezzük.

Ha $\mathbf{a} = (a_1,a_2,a_3)$, $\mathbf{e} = (e_1,e_2,e_3)$, akkor az egyenes paraméteres vektoregyenlete nyilván ekvivalens az alábbi (skalár) egyenletrendszerrel:

$$x_1 = a_1 + t \cdot e_1$$

$$x_2 = a_2 + t \cdot e_2$$

$$x_3 = a_3 + t \cdot e_3$$

ahonnan a t paraméter kiküszöbölésével az

$$\frac{x_1 - a_1}{e_1} = \frac{x_2 - a_2}{e_2} = \frac{x_3 - a_3}{e_3}$$

(nemparaméteres) egyenletpárt nyerjük (feltéve, hogy $e_1,e_2,e_3 \neq 0$).

A síkbeli egyenesek esetéhez hasonlóan, itt is kétféle jelöléstechnika terjedt el: a pontok (vektorok) koordinátáit 1-től 3-ig terjedő indexekkel különböztethetjük meg, mint a fenti paraméteres egyenletrendszerben, de szokásos az egyes koordináták különböző betűkkel (pl. x, y, z-vel) való jelölése is. Ez utóbbi esetben az egyenes paraméteres egyenletrendszerének alakja:

$$x = a_x + t \cdot e_x$$

$$y = a_y + t \cdot e_y$$

$$z = a_z + t \cdot e_z$$

ahol $\mathbf{a} = (a_x,a_y,a_z)$, $\mathbf{e} = (e_x,e_y,e_z)$. Jegyezzük még meg, hogy az egyenes paraméteres egyenletrendszeréből azonnal leolvashatók az egyenes egy pontjának koordinátái és az irányvektor koordinátái is.

A gyakorlatban sokszor előfordul, hogy az egyenes irányvektora explicite nem adott, viszont két különböző pl. $\mathbf{a}$ és $\mathbf{b}$ pontja is ismert. Ekkor irányvektor gyanánt az $\mathbf{e} := \mathbf{b} - \mathbf{a}$ vektor nyilván megfelelő, innen nyerjük a két ponton átmenő egyenes egyenletét:

3.9. Állítás: Legyenek $\mathbf{a},\mathbf{b} \in \mathbf{R}^3$ tetszőleges térbeli pontok. Az $\mathbf{a}$ és $\mathbf{b}$ pontokon átmenő egyenes paraméteres vektoregyenlete:

$$\mathbf{x} = \mathbf{a} + t \cdot (\mathbf{b} - \mathbf{a}) \quad (t \in \mathbf{R}).$$

Itt is igaz marad, hogy ha a t paraméter csak a $[0,1]$ intervallumot futja be, akkor a fenti formulával definiált $\mathbf{x}$ pontok épp az $\mathbf{a}$ és $\mathbf{b}$ végpontú szakasz pontjait futják be.

Két egyenes szögén most is az irányvektoraik által bezárt θ hegyesszöget értjük. Ha a két egyenes paraméteres vektoregyenlete $\mathbf{x} = \mathbf{a} + t \cdot \mathbf{e}$ ill. $\mathbf{x} = \mathbf{b} + \tau \cdot \mathbf{f}$, akkor nyilván: $\cos\theta = \frac{|\langle \mathbf{e},\mathbf{f}\rangle|}{||\mathbf{e}||\cdot||\mathbf{f}||}$. Azonban a síkbeli esettől eltérően, e két egyenes nem feltétlen metszi egymást (akkor sem, ha nem párhuzamosak), azok lehetnek kitérők is. Két egyenes metsző ill. kitérő voltának eldöntése legkézenfekvőbb módon úgy lehetséges, hogy megpróbáljuk a két egyenes közös pontját megkeresni. Legyen a két egyenes paraméteres vektoregyenlete $\mathbf{x} = \mathbf{a}+t\cdot\mathbf{e}$ ill. $\mathbf{x} = \mathbf{b}+\tau\cdot\mathbf{f}$, akkor a közös pont koordinátái mindkét vektoregyenletet kielégítik, azaz $\mathbf{a} + t \cdot \mathbf{e} = \mathbf{b} + \tau \cdot \mathbf{f}$.

A probléma tehát annak eldöntésére redukálódott, hogy a kétismeretlenes, de három egyenletből álló

$$t \cdot e_x - \tau \cdot f_x = b_x - a_x$$

$$t \cdot e_y - \tau \cdot f_y = b_y - a_y$$

$$t \cdot e_z - \tau \cdot f_z = b_z - a_z$$

rendszernek van-e megoldása. Bizonyos kivételes esetektől eltekintve, célhoz vezet, ha megoldjuk pl. az első két egyenlet alkotta rendszert, és behelyettesítéssel megnézzük, hogy a megoldás kielégíti-e a harmadik egyenletet. Ennél a megközelítésnél sokkal gépiesebb eljárást is fogunk mutatni a következő szakaszban bevezetésre kerülő *vektoriális szorzat* fogalmának használatával.

3.3. Vektoriális szorzat

Vektoriális szorzat:

Az $\mathbf{a} = (a_1,a_2,a_3),\mathbf{b} = (b_1,b_2,b_3) \in \mathbf{R}^3$ vektorok *vektoriális szorzatán* az

$$\mathbf{a} \times \mathbf{b} := (a_2b_3 - a_3b_2,a_3b_1 - a_1b_3,a_1b_2 - a_2b_1) \in \mathbf{R}^3$$

vektort értjük.

3.4. Példa: : $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, $\mathbf{k} \times \mathbf{i} = \mathbf{j}$.

A definíció azonnali következménye, hogy bármely vektor önmagával vett vektoriális szorzata a zérusvektorral egyenlő: $\mathbf{a} \times \mathbf{a} = \mathbf{0}$.

Könnyen látható, az is, hogy ha $\mathbf{a}$ és $\mathbf{b}$ harmadik komponense 0, azaz $\mathbf{a} = (a_1,a_2,0)$, $\mathbf{b} = (b_1,b_2,0)$, akkor a vektoriális szorzatvektornak csak a harmadik komponense lehet zérustól különböző, mert ekkor $\mathbf{a} \times \mathbf{b} = (0,0,a_1b_2 - a_2b_1)$. Ekkor pedig nyilván $||\mathbf{a} \times \mathbf{b}|| = |a_1b_2 - a_2b_1| = |\langle(a_1,a_2),(-b_2,b_1)\rangle| = |\langle(a_1,a_2),(b_1,b_2)^{\perp}\rangle|$ Ha pedig az (a_1,a_2) és (b_1,b_2) vektorok θ szöget zárnak be, akkor az (a_1,a_2) és az elforgatott $(b_1,b_2)^{\perp}$ vektor által bezárt szög nyilván $\frac{\pi}{2} - \theta$, innen a skaláris szorzat tulajdonságai alapján: $||\mathbf{a} \times \mathbf{b}|| := ||\mathbf{a}|| \cdot ||\mathbf{b}|| \cdot |\cos(\frac{\pi}{2} - \theta)| = ||\mathbf{a}|| \cdot ||\mathbf{b}|| \cdot |\sin\theta|$. Mivel pedig a vektorok hossza és egymással bezárt szöge szempontjából a koordinátarendszer megválasztása közömbös, azért a következő, most már tetszőleges térbeli vektorokra vonatkozó állítást kaptuk:

3.10. Állítás: Tetszőleges $\mathbf{a},\mathbf{b} \in \mathbf{R}^3$ esetén

$$||\mathbf{a} \times \mathbf{b}|| := ||\mathbf{a}|| \cdot ||\mathbf{b}|| \cdot |\sin\theta|,$$

ahol θ az $\mathbf{a}$ és $\mathbf{b}$ vektorok szögét jelöli.

A következő tétel a vektoriális szorzatvektornak a vektorgeometria szempontjából legfontosabb tulajdonságát fejezi ki:

3.1. Tétel: A vektoriális szorzatvektor mindkét tényezőjére ortogonális, azaz tetszőleges $\mathbf{a},\mathbf{b} \in \mathbf{R}^3$ esetén $\langle \mathbf{a} \times \mathbf{b},\mathbf{a} \rangle = 0$ és $\langle \mathbf{a} \times \mathbf{b},\mathbf{b} \rangle = 0$.

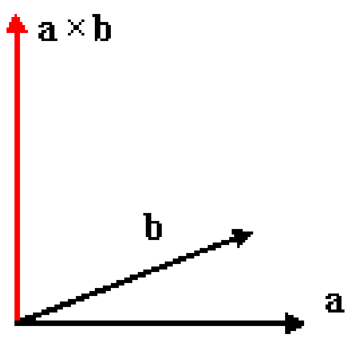
Bizonyítás: Csak az első egyenlőséget igazoljuk, a másik hasonlóan végezhető el:

$$\langle \mathbf{a} \times \mathbf{b},\mathbf{a} \rangle =$$

$$= \langle (a_2b_3 - a_3b_2,a_3b_1 - a_1b_3,a_1b_2 - a_2b_1),(a_1,a_2,a_3) \rangle =$$

$$= a_1a_2b_3 - a_1a_3b_2 + a_2a_3b_1 - a_1a_2b_3 + a_1a_3b_2 - a_2a_3b_1 = 0.$$

Két vektor vektoriális szorzata tehát automatikusan merőleges mindkét vektorra. Irányítása az $\mathbf{i} \times \mathbf{j} = \mathbf{k}$ egyenlőség alapján könnyen megjegyezhető a következő módon: jobb kezünk első három ujját nyújtsuk ki egymásra merőlegesen úgy, hogy hüvelykujjunk az első, mutatóujjunk a második vektor irányába mutasson, ekkor középső ujjunk a vektoriális szorzatvektor irányát mutatja. Ezt a (fizikából átvett) szabályszerűséget *jobbkézszabálynak* nevezzük.



3.7. ábra. *A vektoriális szorzatvektor iránya*

Most pedig összefoglaljuk a vektoriális szorzatra vonatkozó azonosságokat. Ezek teljesülését a definíció alapján az Olvasó könnyen ellenőrizheti:

3.11. Állítás: Tetszőleges $\mathbf{a},\mathbf{b},\mathbf{c} \in \mathbf{R}^3$ és $\lambda \in \mathbf{R}$ esetén:
$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
$(\lambda\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda\mathbf{b}) = \lambda \cdot (\mathbf{a} \times \mathbf{b})$
$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$, és $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$

Vagyis a számokra vonatkozó jól ismert azonosságok érvényesek maradnak azzal a jelentős különbséggel, hogy a vektoriális szorzat nem kommutatív (a tényezők felcserélésével előjelet vált), és nem is asszociatív.

Most már egyszerűen megoldhatjuk az előző szakaszban felvetett problémát, azaz annak eldöntését, hogy két adott, $\mathbf{x} = \mathbf{a} + t \cdot \mathbf{e}$ és $\mathbf{x} = \mathbf{b} + \tau \cdot \mathbf{f}$ vektoregyenletű egyenes kitérő-e vagy sem. A két egyenes párhuzamossága könnyen ellenőrizhető, így feltehető, hogy az egyenesek vagy metszik egymást, vagy kitérőek. Az $\mathbf{n} := \mathbf{e} \times \mathbf{f}$ vektor mindkét egyenesre merőleges. Tekintsük most a $(\mathbf{b} - \mathbf{a})$ vektort. Ha a két egyenes metszi egymást, akkor $(\mathbf{b} - \mathbf{a})$ is közös síkjukban van, így $\mathbf{n}$ ortogonális $(\mathbf{b} - \mathbf{a})$-ra. Ha a két egyenes kitérő, akkor $(\mathbf{b} - \mathbf{a})$-nak $\mathbf{n}$ irányú merőleges vetülete nem 0, éspedig épp a normáltranszverzális szakasz hossza (azaz a két egyenest összekötő, mindkét egyenesre merőleges szakasz hossza). Tehát elég a $\langle \mathbf{e} \times \mathbf{f}, \mathbf{b} - \mathbf{a} \rangle$ kifejezést kiértékelni: ha ez nem 0, akkor a két egyenes kitérő, egyébként pedig egy síkban vannak.

Az előző gondolatmenetben fellépő $\langle \mathbf{e} \times \mathbf{f}, \mathbf{g} \rangle$ alakú kifejezést az $\mathbf{e}$, $\mathbf{f}$, $\mathbf{g}$ vektorok *vegyesszorzatának* is nevezik, és egyszerűen $\mathbf{efg}$-vel jelölik. Értéke egy szám, melynek abszolút értéke – elemi geometriai meggondolások alapján – az $\mathbf{e}$, $\mathbf{f}$, $\mathbf{g}$ vektorok által kifeszített paralellepipedon (paralellogramma alapú ferde hasáb) térfogatával egyezik, ennélfogva pontosan akkor 0, ha a három vektor egy síkba esik.

3.4. Síkok a térben

A térbeli egyenesek mintájára a síkok is leírhatók paraméteres vektoregyenlettel. Legyen $\mathbf{e},\mathbf{f} \in \mathbf{R}^3$ két lineárisan független (tehát nem egy egyenesbe eső) vektor, $\mathbf{a} \in \mathbf{R}^3$ pedig adott pont. Akkor az

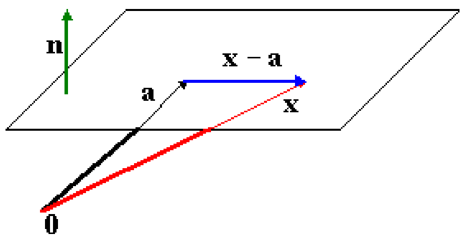
$$\mathbf{x} := \mathbf{a} + u \cdot \mathbf{e} + v \cdot \mathbf{f} \qquad (u,v \in \mathbf{R})$$

pontok egy síkot alkotnak, mely illeszkedik az $\mathbf{a}$ pontra és párhuzamos az $\mathbf{e}$ és az $\mathbf{f}$ vektorokkal. Sokkal egyszerűbb és elterjedtebb azonban a síkokat egy, az adott síkra merőleges ún. normálvektor segítségével leírni. Jelöljön $\mathbf{n} \neq \mathbf{0}$ egy ilyen normálvektort, és legyen $\mathbf{a} \in \mathbf{R}^3$ a sík egy tetszőleges pontja. Nyilvánvaló, hogy egy $\mathbf{x} \in \mathbf{R}^3$ pont akkor és csakis van rajta a síkon, ha az $\mathbf{x} - \mathbf{a}$ különbségvektor párhuzamos a síkkal, azaz, ha $\mathbf{x} - \mathbf{a}$ merőleges az $\mathbf{n}$ normálvektorra. Innen nyerjük az adott pontra illeszkedő, adott normálvektorú sík ún *normálegyenletét* (ld. az ábrát):

$$\langle \mathbf{x} - \mathbf{a}, \mathbf{n} \rangle = 0,$$

melyet kifejtve kapjuk az alábbi egyenletet:

$$(x_1 - a_1) \cdot n_1 + (x_2 - a_2) \cdot n_2 + (x_3 - a_3) \cdot n_3 = 0$$



3.8. ábra. *Egy pontjával és normálvektorával adott sík*

Csakúgy mint az egyeneseknél, a jelölési konvenciók itt sem egységesek. Sokszor szokás a sík egy pontját (x,y,z)-vel jelölni: ekkor a sík normálegyenlete $(x - a_x) \cdot n_x + (y - a_y) \cdot n_y + (z - a_z) \cdot n_z = 0$ alakú,. ahol most (a_x,a_y,a_z) a sík egy pontja, és (n_x,n_y,n_z) a normálvektor.

3.5. Példa: Az origóra illeszkedő, az $x = 2 - 3t$, $y = t$, $z = 1 + 4t$ egyenesre merőleges sík normálegyenlete:

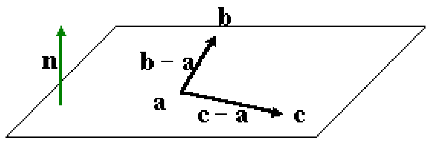
$$-3x + y + 4z = 0,$$

mivel az egyenes $(-3,1,4)$ irányvektora egyúttal a sík normálvektora is.

Gyakran előfordul, hogy a normálvektor maga nem adott, viszont a síknak három pontját ($\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$) ismerjük: ha e három pont nem esik egy egyenesbe, akkor meghatározza a síkot. Ekkor első lépésben normálvektort kell keresnünk. Nyilvánvaló, hogy a $(\mathbf{b}-\mathbf{a})$ és a $(\mathbf{c}-\mathbf{a})$ különbségvektorok mindketten párhuzamosak a síkkal, és lineárisan függetlenek, mivel nem esnek egy egyenesbe. Következésképp ezek vektoriális szorzata, az $\mathbf{n} := (\mathbf{b}-\mathbf{a}) \times (\mathbf{c}-\mathbf{a})$ vektor merőleges az egész síkra, így választható normálvektornak. Így nyerjük a három pontra illeszkedő sík normálegyenletét:

3.12. Állítás: Az $\mathbf{a}$,$\mathbf{b}$,$\mathbf{c} \in \mathbf{R}^3$ (nem egy egyenesbe eső) pontok által meghatározott sík normálegyenlete:

$$\langle \mathbf{x}-\mathbf{a}, (\mathbf{b}-\mathbf{a}) \times (\mathbf{c}-\mathbf{a}) \rangle = 0,$$



3.9. ábra. *Három pontjával adott sík*

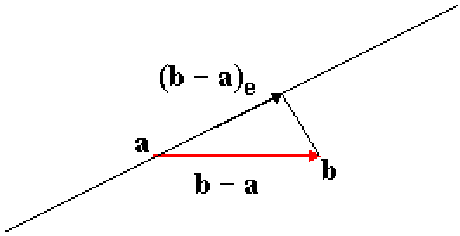
3.6. Példa: Határozzuk meg az (1,0,0), (0,1,0), (0,0,1) pontokra (azaz az $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ vektorok végpontjaira) illeszkedő sík normálegyenletét.

Megoldás: A normálvektor: $\mathbf{n} = (\mathbf{k}-\mathbf{i}) \times (\mathbf{j}-\mathbf{i}) = \mathbf{k} \times \mathbf{j} - \mathbf{i} \times \mathbf{j} - \mathbf{k} \times \mathbf{i} + \mathbf{i} \times \mathbf{i} = -\mathbf{j} \times \mathbf{k} - \mathbf{i} \times \mathbf{j} - \mathbf{k} \times \mathbf{i} = -\mathbf{i} - \mathbf{k} - \mathbf{j} = (-1, -1, -1)$, innen a sík egyenlete: $-(x-1) - y - z = 0$, azaz $x + y + z = 1$.

Végezetül összefoglalunk néhány, a térbeli pontokkal, egyenesekkel és síkokkal kapcsolatos néhány típusproblémát és azok egy-egy lehetséges megoldási módszerét.

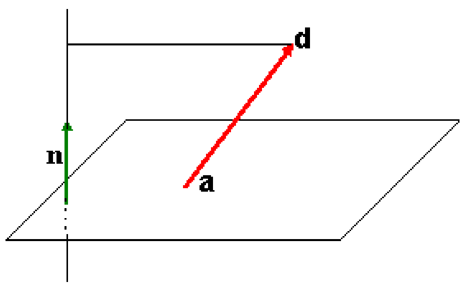
Pont és egyenes távolsága: A $\mathbf{b} \in \mathbf{R}^3$ pontnak az $\mathbf{x} = \mathbf{a} + t \cdot \mathbf{e}$ egyenestől mért távolsága épp az alábbi vektor hossza: $(\mathbf{b}-\mathbf{a}) - (\mathbf{b}-\mathbf{a})_{\mathbf{e}}$, ahol $(\mathbf{b}-\mathbf{a})_{\mathbf{e}}$ a $(\mathbf{b}-\mathbf{a})$ különbségvektornak $\mathbf{e}$ irányú ortogonális vetülete. A szóban forgó vektor hosszát Pitagorász tétele alapján számíthatjuk:

$$\sqrt{||\mathbf{b}-\mathbf{a}||^2 - \left(\frac{|\langle \mathbf{b}-\mathbf{a},\mathbf{e}\rangle|}{||\mathbf{e}||}\right)^2}$$



3.10. ábra. *Pont és egyenes távolsága*

Pont és sík távolsága: A $\mathbf{d} \in \mathbf{R}^3$ pontnak az $\langle \mathbf{x}-\mathbf{a},\mathbf{n}\rangle = 0$ síktól mért távolsága épp a $(\mathbf{d}-\mathbf{a})$ különbségvektor $\mathbf{n}$ irányú ortogonális vetületvektorának hossza (ld. az ábrát), azaz $\frac{|\langle \mathbf{d}-\mathbf{a},\mathbf{n}\rangle|}{||\mathbf{n}||}$ Következésképp



3.11. ábra. *Pont és sík távolsága*

az $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d} \in \mathbf{R}^3$ pontok akkor és csakis akkor vannak egy síkon, ha $\langle \mathbf{d}-\mathbf{a},(\mathbf{b}-\mathbf{a})\times(\mathbf{c}-\mathbf{a})\rangle = 0$. (Gondoljuk át a kivételes eseteket is!) Ez az eredmény újabb megoldási módszerét adja az egyenesek kitérő jellegének eldöntésére: nyilvánvaló, hogy két egyenes akkor és csakis akkor esik egy síkba, ha két-két különböző pontja egy síkon van.

Adott pontra és adott egyenesre illeszkedő sík. Az $\mathbf{x} = \mathbf{a} + t \cdot \mathbf{e}$ egyenesre és az azon kívül eső $\mathbf{b} \in \mathbf{R}^3$ pontra illeszkedő sík nyilván párhuzamos mind a $(\mathbf{b} - \mathbf{a})$ különbségvektorral, mind pedig az $\mathbf{e}$ irányvektorral, így normálvektora $\mathbf{n} := (\mathbf{b} - \mathbf{a}) \times \mathbf{e}$-nek választható.

Két adott síkkal párhuzamos egyenes: Az $\langle \mathbf{x} - \mathbf{a}, \mathbf{n} \rangle = 0$ és az $\langle \mathbf{x} - \mathbf{b}, \mathbf{m} \rangle = 0$ síkokkal párhuzamos egyenes merőleges mindkét sík normálvektorára, így irányvektora $\mathbf{e} := \mathbf{n} \times \mathbf{m}$-nek választható.

Egyenes és sík döféspontja. Az $\mathbf{x} = \mathbf{a} + t \cdot \mathbf{e}$ egyenes és az $\langle \mathbf{x} - \mathbf{b}, \mathbf{n} \rangle = 0$ sík döféspontja az az $\mathbf{x} \in \mathbf{R}^3$ pont, mely kielégíti mindkét egyenletet. Az ezt jellemző t paraméter így az $\langle \mathbf{a} + t \cdot \mathbf{e} - \mathbf{b}, \mathbf{n} \rangle = 0$ egyenletből határozható meg, ahonnan $t = \frac{\langle \mathbf{b} - \mathbf{a}, \mathbf{n} \rangle}{\langle \mathbf{e}, \mathbf{n} \rangle}$. Ezt a t számot behelyettesítve az egyenes egyenletébe, a döféspont már adódik.

3.5. Feladatok

1. Határozzuk meg annak az egyenesnek az egyenletét, mely illeszkedik a $(2,3,-1)$ pontra, és párhuzamos az $x = 1$, $y = -t$, $z = 1+t$ egyenessel.

2. Határozzuk meg annak az egyenesnek az egyenletét, mely illeszkedik az $(1,0,1)$ pontra, és merőleges az $x + y - 5z = 50$ síkra. Határozzuk meg a döféspont koordinátáit is.

3. Párhuzamos-e az $x = 3 + 4t$, $y = -1 - 2t$, $z = 7t$ egyenes a $3x - y - 2z = 23$ síkkal?

4. Mi a $(11,11,11)$ pont merőleges vetülete az $x + y + z = 1$ síkon?

5. Adottak az $A_1 := (x_1,y_1), A_2 := (x_2,y_2), A_3 := (x_3,y_3)$, $\in \mathbf{R}^2$ síkbeli pontok, melyek nem egy egyenesre illeszkednek. Hogyan dönthetjük el, hogy egy $A := (x,y) \in \mathbf{R}^2$ pont a fenti 3 pont meghatározta háromszög belsejében fekszik-e vagy sem?

6. Legyenek $\mathbf{a} := (1,10,1000)$ és $\mathbf{b} := (-1,0.1,-0.001)$. Határozzuk meg az $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{b} \times \mathbf{a})$ vektort.

7. Legyenek $\mathbf{a} := (-1,2,1)$ és $\mathbf{b} := (3,1,1)$. Bázist alkotnak-e $\mathbf{R}$-ban a $\mathbf{b}$, $\mathbf{b} \times \mathbf{a}$ és a $\mathbf{b} \times (\mathbf{b} \times \mathbf{a})$ vektorok?

8. Az $x + 2y + 3z = 4$ és a $3x + 2y + z = -4$ síkok metszésvonala merőleges-e az $(1,1,1)$ pont helyvektorára?

9. Melyik az az egyenes, mely párhuzamos az $x - y + 2z = 1$ és az $x + y - 2z = 1$ síkokkal, és illeszkedik az origóra?

10. Írjuk fel az $A_1 := (-2,-3,5)$, $A_2 := (-3,5,-2)$ és az $A_3 := (5,-2,-3)$ pontokra illeszkedő sík egyenletét. Altere-e ez a sík $\mathbf{R}^3$-nak?

11. Írjuk fel annak a síknak az egyenletét, mely illeszkedik az origóra és az $x = t$, $y = 2t + 1$, $z = 3t + 2$ egyenletű egyenesre.

12. Határozzuk meg annak az egyenesnek az egyenletét, mely illeszkedik az $(1,1,1)$ pontra, és merőleges az $x = 1 + t$, $y = t$, $z = -1 - t$ és az $x = 1 - t$, $y = t$, $z = -1 + t$ egyenletű egyenesekre.

13. Határozzuk meg az $x - y + 4z = 0$ és az $x + y + 4z = 0$ síkok metszésvonalát.

14. Egy síkon vannak-e az $(1{,}10{,}100)$, $(100{,}1{,}10)$, $(10{,}100{,}1)$ és a $(37{,}37{,}37)$ pontok?

15. Egy síkban vannak-e az $x = 1 + 2t$, $y = 1 - 3t$, $z = 1 + 4t$ egyenes valamint a $P := (1{,}2{,}3)$ és a $Q := (5, -4{,}11)$ pontok?

16. Metszik-e egymást az $x = 1 + 3t$, $y = 2 + 2t$, $z = 3 + t$ és az $x = 3 + t$, $y = 2 + 2t$, $z = 1 + 3t$ egyenesek? Ha igen, mi a metszéspont? Ha nem, miért nem?

Megoldások

1. Az adott egyenes irányvektora az egyenletéből kiolvasható: $\mathbf{e} = (0,-1,1)$, ahonnan a keresett egyenes egyenlete: $x = 2$, $y = 3 - t$, $z = -1 + t$.

2. Az egyenes irányvektora párhuzamos kell, hogy legyen a sík normálvektorával, azaz az $e := (1,1,-5)$ választás megfelel. Innen az egyenes egyenlete: $x = 1 + t$, $y = t$, $z = -1 - 5t$. A döféspont az egyenesnek olyan (egyelőre ismeretlen t paraméternek megfelelő) pontja, mely a sík egyenletét is kielégíti, azaz, melyre $(1 + t) + t - 5(1 - 5t) = 50$ teljesül. Innen t számítható: t-re $t = 2$ adódik. Következésképp a döféspont koordinátái: $(1 + 2,2,1 - 5 \cdot 2) = (3,2,-9)$.

3. Az egyenes és a sík párhuzamossága azzal egyenértékű, hogy az egyenes irányvektora és a sík normálvektora merőlegesek. Jelen esetben $\mathbf{e} = (4,-2,7)$ és $\mathbf{n} = (3,-1,-2)$. Mivel pedig $\langle \mathbf{e},\mathbf{n} \rangle = 0$, azért $\mathbf{e} \perp \mathbf{n}$, tehát az egyenes párhuzamos a síkkal.

4. A vetületi pont egyezik az (11,11,11) pontra illeszkedő, a sík normálvektorával párhuzamos egyenes és a sík döféspontjával. A sík normálvektora (1,1,1), így az egyenes egyenlete: $x = 11 + t$, $y = 11 + t$, $z = 11 + t$. A döféspont koordinátái kielégítik a sík egyenletét, azaz $(11 + t) + (11 + t) + (11 + t) = 1$, innen $t = -\frac{32}{3}$. Következésképp a döféspont koordinátái: $(11 - \frac{32}{3},11 - \frac{32}{3},11 - \frac{32}{3}) = (\frac{1}{3},\frac{1}{3},\frac{1}{3})$.

5. Könnyen látható, hogy a $t \cdot A_1 + (1 - t) \cdot A_2$ pontok akkor és csak akkor vannak rajta az A_1A_2 szakaszon, ha $0 \le t \le 1$. Másszóval, a $t_1 \cdot A_1 + t_2 \cdot A_2$ pontok akkor és csak akkor vannak rajta az A_1A_2 szakaszon, ha $t_1,t_2 \ge 0$ és $t_1 + t_2 = 1$. Innen könnyen látható, hogy a $t_1 \cdot A_1 + t_2 \cdot A_2 + t_3A_3$ pontok akkor és csak akkor esnek az $A_1A_2A_3$ háromszögbe, ha $t_1,t_2,t_3 \ge 0$ és $t_1 + t_2 + t_3 = 1$. Ezen észrevétel után az algoritmus a következő lehet: megkíséreljük előállítani az A pontot, illetve annak x,y koordinátáit a következő alakban:

$$\begin{aligned} t_1x_1 + t_2x_2 + t_3x_3 &= x \\ t_1y_1 + t_2y_2 + t_3y_3 &= y \\ t_1 \quad + t_2 \quad + t_3 \quad &= 1 \end{aligned}$$

azaz megoldjuk a fenti 3-ismeretlenes egyenletrendszert az ismeretlen t_1,t_2,t_3 együtthatókra. (Megoldás mindig van, mert mivel A_1,A_2,A_3 nem esnek egy egyenesbe, azért az $(A_2 - A_1)$ és az $(A_3 - A_1)$ vektorok lineárisan függetlenek, így bázist alkotnak $\mathbf{R}^2$-ben, tehát minden vektor, így (x,y) is előáll ezek lineáris kombinációjaként: $x = \alpha(x_2 - x_1) + \beta(x_3 - x_1)$, $y = \alpha(y_2 - y_1) + \beta(y_3 - y_1)$, ahonnan $t_1 = -\alpha - \beta$, $t_2 = \alpha$, $t_3 = \beta$.) Ha az így kapott t_1,t_2,t_3 együtthatók mindegyike pozitív, akkor az A pont a háromszög belsejében van, egyébként nem.

A fenti t_1,t_2,t_3 számokat az A pont *baricentrikus koordinátáinak* nevezzük. Az algoritmus értelemszerűen általánosítható tetszőleges, de *konvex* sokszögekre. A feladatban megadott probléma egyébként gyakran előfordul pl. a számítógépes grafikában.

6. Az eredmény (számolás nélkül) a zérusvektor, mivel $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{b} \times \mathbf{a}) = -(\mathbf{a} \times \mathbf{b}) \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$.

7. Mivel $\mathbf{a}$ és $\mathbf{b}$ nem párhuzamosak, így egy kétdimenziós alteret (síkot) generálnak: $\mathbf{b} \times \mathbf{a}$ merőleges erre a síkra, $\mathbf{b} \times (\mathbf{b} \times \mathbf{a})$ pedig mind $\mathbf{a}$-ra, mind $\mathbf{b} \times \mathbf{a}$-ra merőleges. A három vektor tehát páronként ortogonális, egyik sem zérusvektor, így lineárisan függetlenek. Mivel pedig $\mathbf{R}^3$ háromdimenziós, azért e három vektor bázist alkot $\mathbf{R}^3$-ban.

8. A metszésvonal mindkét sík normálvektorára merőleges, így az irányvektornak az $\mathbf{e} := (1,2,3) \times (3,2,1) = (-4,8,-4)$ választás megfelel. Ez pedig merőleges az $(1,1,1)$ vektorra, mert $\langle(-4,8,-4),(1,1,1)\rangle = 0$.

9. Az egyenes merőleges mindkét sík normálvektorára, így az irányvektornak az $\mathbf{e} := (1,-1,2) \times (1,1,-2) = (0,4,2)$ választás megfelel. A keresett egyenes az origóra illeszkedik, így egyenlete: $x = 0$, $y = 4t$, $z = 2t$. Ugyanennek az egyenesnek egy másik egyenlete: $x = 0$, $y = 2t$, $z = t$.

10. Az $A_2 - A_1 = (-1,8,-7)$ és az $A_3 - A_1 = (7,1,-8)$ vektorok a síkkal párhuzamosak: a sík normálvektora tehát kettőjük vektoriális szorzatának vehető: $(-1,8,-7) \times (7,1,-8) = (-57,-57,-57)$. Ehelyett célszerűbb a vele párhuzamos $\mathbf{n} := (1,1,1)$ vektort venni. A sík illeszkedik pl. A_1-re, így egyenlete $(x+1) + (y-8) + (z+7) = 0$, azaz $x + y + z = 0$. Innen látható, hogy a sík illeszkedik az origóra is, így altér $\mathbf{R}^3$-ban.

11. A sík tartalmazza az origót és pl. a $(0,1,2)$ pontot (az egyenes $t = 0$ paraméterű pontját), így párhuzamos a $(0,1,2)$ vektorral, továbbá pár-

huzamos az egyenes $(1{,}2{,}3)$ irányvektorával is. Normálvektora tehát e két vektor vektoriális szorzatának vehető: $\mathbf{n} := (0{,}1{,}2) \times (1{,}2{,}3) = (-1{,}2,-1)$. A sík illeszkedik az origóra is, így egyenlete: $-x + 2y - z = 0$.

12. A sík normálvektora mindkét egyenes irányvektorára merőleges, így e kettő vektoriális szorzatának vehető: $(1{,}1,-1) \times (-1{,}1{,}1) = (2{,}0{,}2)$. Egyszerűbb ennek $\frac{1}{2}$-szeresét választani: $\mathbf{n} := (1{,}0{,}1)$. A sík illeszkedik az $(1{,}1{,}1)$ pontra is, így egyenlete: $(x-1) + (z-1) = 0$, azaz $x + z = 2$.

13. A metszésvonal irányvektora mindkét sík normálvektorára merőleges, így kettőjük vektoriális szorzatának vehető: $(1,-1{,}4) \times (1{,}1{,}4) = (-8{,}0{,}2)$. Egyszerűbb ennek $\frac{1}{2}$-szeresét választani: $\mathbf{e} := (-4{,}0{,}1)$. Mivel mindkét sík illeszkedik az origóra, azért a metszésvonal is. Így a metszésvonal egyenlete: $x = -4t$, $y = 0$, $z = -t$.

14. A feladat gépies megoldása: a $(99,-9,-90)$ és a $(9{,}90,-99)$ különbségvektorok párhuzamosak az első 3 pont által meghatározott síkkal, így annak normálvektora e két vektor vektoriális szorzatának vehető: $(99,-9,-90) \times (9{,}90,-99) = 9 \cdot 9 \cdot (111{,}111{,}111)$. Egyszerűbb ennek $\frac{1}{81\cdot 111}$-szeresét választani: $\mathbf{n} := (1{,}1{,}1)$. A sík illeszkedik pl. az $(1{,}10{,}100)$ pontra így egyenlete $(x-1) + (y-100) + (z-100) = 0$, azaz $x + y + z = 111$. Ezt pedig a negyedik pont koordinátái kielégítik, tehát a négy pont egy síkon van.

Jóval gyorsabban célhoz érünk, ha észrevesszük, hogy a 4. pont épp az első három pont által meghatározott háromszög súlypontja, így szükségképp velük egy síkon van.

15. Az egyenesre illeszkedik pl. az $A := (1{,}1{,}1)$ pont is ($t = 0$ mellett). Így az egyenes, az A és a P pont olyan síkra illeszkednek, melynek normálvektora merőleges az egyenes $(2,-3{,}4)$ irányvektorára és a $P - A = (0{,}1{,}2)$ különbségvektorra is, így kettőjük vektoriális szorzatának vehető: $(2,-3{,}4) \times (0{,}1{,}2) = (-10,-4{,}2)$. Célszerű ennek $\left(-\frac{1}{2}\right)$-ét venni: $\mathbf{n} := (5{,}2,-1)$. A sík illeszkedik pl. az $(1{,}1{,}1)$ pontra is, így egyenlete: $5(x-1) + 2(y-1) - (z-1) = 0$, azaz $5x + 2y - z = 6$. Ezt pedig a Q pont koordinátái kielégítik, tehát az egyenes és a két pont egy síkon van.

16. A mindkét egyenessel párhuzamos sík normálvektora az irányvektorok vektoriális szorzatának vehető: $(3{,}2{,}1) \times (1{,}2{,}3) = (4,-8{,}4)$. Célszerű ennek $\frac{1}{4}$-ét venni: $\mathbf{n} := (1,-2{,}1)$. Ha a síkra illeszkedik pl. az első

egyenes $(1{,}2{,}3)$ pontja is, akkor a sík egyenlete: $(x-1)-2(y-2)+(z-3)=0$, azaz $x-2y+z=0$. Ezt pedig mindkét egyenes minden pontja kielégíti, tehát a két egyenes valóban egy síkon van, és mivel irányvektoraik nem párhuzamosak, azért metszők. A metszéspont koordinátái kielégítik mindkét egyenes egyenletét:

$$\begin{aligned} 1+3t_1 &= 3+t_2 \\ 2+2t_1 &= 2+2t_2 \\ 3+t_1 &= 1+3t_2 \end{aligned}$$

Az egyenletrendszer megoldható, megoldása könnyen láthatóan $t_1=t_2=1$. Innen a metszéspont koordinátái: $(4{,}4{,}4)$.

Voltaképpen felesleges volt a közös sík egyenletét meghatározni. A metszéspont létezése ui. egyúttal azt jelenti, hogy a két egyenes egy síkon van.

4. Lineáris leképezések, mátrixok

Ebben a fejezetben vektorterek között ható speciális leképezéseket, ún. *lineáris leképezéseket* tanulmányozunk. Bevezetjük a téglalap alakú számtáblázatokat (mátrixokat), és műveleteket értelmezünk közöttük. Kiderül, hogy minden mátrix egyúttal egy-egy speciális lineáris leképezésként is felfogható. Az elmélet kitűnően alkalmas a sokismeretlenes lineáris egyenletrendszerek vizsgálatára és megoldására, de természetes módon felbukkan többváltozós analízisben is (ezt majd a következő fejezetben látjuk), és egy sereg más problémakör vizsgálatában (pl. nemlineáris egyenletrendszerek közelítő megoldása; egyenlőtlenség-rendszerek; differenciálegyenlet-rendszerek stb).

4.1. Lineáris leképezések

Legyenek X, Y vektorterek.

Lineáris leképezés:

Az $A : X \to Y$ függvényt *lineáris leképezésnek* nevezzük, ha

- a $\mathcal{D}_A$ értelmezési tartomány altér X-ben;
- $A(x + y) = A(x) + A(y)$ minden $x,y \in \mathcal{D}_A$ esetén;
- $A(\lambda x) = \lambda \cdot A(x)$ minden $x \in \mathcal{D}_A$, $\lambda \in \mathbf{R}$ esetén.

A továbbiakban, ha A lineáris leképezés, akkor $A(x)$ helyett egyszerűen csak Ax-et írunk, amennyiben ez nem okoz félreértést.

A definíció azonnali következményeként:

4.1. Állítás: Ha $A : X \to Y$ lineáris leképezés, akkor
(a) tetszőleges $x = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_n x_n$ lineáris kombináció esetén $Ax = \lambda_1 Ax_1 + \lambda_2 Ax_2 + ... + \lambda_n Ax_n$;
(b) az $\mathcal{R}_A$ képtér is altér (Y-ban).

Az állítás (a) pontja miatt nyilván mindig teljesül, hogy $A\mathbf{0} = \mathbf{0}$. (Itt $\mathbf{0}$ alatt a bal oldalon értelemszerűen az X tér zérusvektora, míg a jobb oldalon az Y tér zérusvektora értendő.)

Az alábbiakban példákat mutatunk lineáris leképezésekre. Ezekből világosan látszik, hogy – bizonyos értelemben – a lineáris leképezések a "lehető legegyszerűbb" függvények.

1. Legyen $A : \mathbf{R} \to \mathbf{R}$, $Ax := a \cdot x$ (ahol $a \in \mathbf{R}$ adott konstans). Akkor A lineáris leképezés. Könnyen látható továbbá, hogy *minden* $\mathbf{R} \to \mathbf{R}$ lineáris leképezés ilyen alakú ($a := A(1)$ választással). Így pl. az $A(x) := a \cdot x + b$ előírással értelmezett leképezés $b \neq 0$ esetén már nem lineáris leképezés.

2. Tekintsük az $[a,b]$ intervallumon folytonos függvények $C[a,b]$ halmazát, és az ugyanitt folytonosan differenciálható függvények $C^1[a,b]$ halmazát. Már láttuk, hogy ezek vektorteret alkotnak a szokásos függvényműveletekre nézve. Legyen $D : C^1[a,b] \to C[a,b]$ a differenciálás operátora, azaz $Df := f'$ minden $f \in C^1[a,b]$-re. Akkor D lineáris leképezés.

3. Legyen $I : C[a,b] \to \mathbf{R}$, $If := \int_a^b f(x)dx$. Akkor I lineáris leképezés.
Elnevezés: A számértékű lineáris leképezéseket – mint a fenti I leképezést is – gyakran *lineáris funkcionáloknak* nevezzük.

4 Legyen $\delta : C(\mathbf{R}) \to \mathbf{R}$, $\delta f := f(0)$, azaz rendeljük hozzá f-hez a 0-beli helyettesítési értékét. Akkor δ lineáris leképezés. (*Dirac-féle δ-funkcionál*).

5. Az $\mathbf{R}^2$ sík minden pontjához rendeljük hozzá az origó körüli, adott t szögű elforgatás útján kapott pontot. Ezzel egy $\mathbf{R}^2 \to \mathbf{R}^2$ lineáris leképezést definiáltunk. Ugyanakkor az origótól különböző pont körüli forgatás már nem eredményez lineáris leképezést.

6. Az $\mathbf{R}^3$ tér minden pontjához rendeljük hozzá a pontnak egy adott, az origóra illeszkedő síkra vett ortogonális vetületét. Ezzel egy $\mathbf{R}^3 \to \mathbf{R}^3$ lineáris leképezést definiáltunk. Ha viszont a sík nem illeszkedik az origóra, akkor az így definiált leképezés már nem lineáris.

7. Legyen $\mathbf{a} \in \mathbf{R}^3$ adott vektor és minden $\mathbf{x} \in \mathbf{R}^3$ esetén jelölje $A\mathbf{x} := \mathbf{x} \times \mathbf{a}$. Az így definiált $A : \mathbf{R}^3 \to \mathbf{R}^3$ leképezés lineáris leképezés.

4.2. Mátrixok, műveletek mátrixokkal

Mátrix:
Az n sorból és m oszlopból álló téglalap alakú

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

számtáblázatot $n \times m$ -es mátrixnak nevezzük.

A mátrix elemeire két indexszel hivatkozunk, az elsőt sor-, a másodikat oszlopindexnek nevezzük.

Elnevezések: Ha $n = m$, akkor a mátrixot *négyzetes mátrixnak*, ha $n = 1$, akkor *sorvektornak*, ha $m = 1$, akkor *oszlopvektornak* nevezzük. Az A mátrix *zérusmátrix*, ha minden eleme zérus (jele **0**). $n \times n$-es mátrixok esetén az n számot a mátrix *rendjének* nevezzük. Az A négyzetes mátrix *diagonálmátrix*, ha csak a bal felső saroktól a jobb alsó sarokig tartó, ún. főüdiagonálisban lévő elemek különbözhetnek 0-tól, azaz, ha a következő alakú:

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Az $n \times n$-es

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

diagonálmátrixot $n \times n$-es *egységmátrixnak* nevezzük és I-vel fogjuk jelölni.

Jelölések: Az $n \times m$ -es mátrixok halmazát $\mathbf{M}_{n\times m}$-mel jelöljük. Egy $A \in \mathbf{M}_{n\times m}$ mátrixot gyakran az elemekre utaló $A = [a_{kj}]$ szimbólummal jelölünk ($k = 1,2,\dots,n$, $j = 1,2,\dots,m$).

Mátrixok összege és konstansszorosa:

Legyenek $A = [a_{kj}]$,$B = [b_{kj}] \in \mathbf{M}_{n\times m}$ tetszőleges mátrixok, $\lambda \in \mathbf{R}$ tetszőleges szám. Az A és B mátrixok összegén az

$$A + B := \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2m}+b_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \dots & a_{nm}+b_{nm} \end{pmatrix} \in \mathbf{M}_{n\times m}$$

mátrixot, az A mátrix λ-szorosán pedig a

$$\lambda \cdot A := \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \dots & \dots & \dots & \dots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix} \in \mathbf{M}_{n\times m}$$

mátrixot értjük.

A definíció azonnali következményeként:

4.2. Állítás: Az $n \times m$-es mátrixok $\mathbf{M}_{n\times m}$ halmaza az összeadásra és a skalárral való szorzásra nézve vektorteret alkot. Ennek nulleleme a zérusmátrix, dimenziója nm. Egy bázisát (a standard bázist) mindazon $n \times m$-es mátrixok alkotják, melyek elemei közt egyetlenegy 1-es van, a többi elemük zérus.

A most következő definícióban a szorzást az eddigi komponensenkénti műveletektől teljesen eltérően definiáljuk. Ennek célja, hogy a mátrixszorzás a *függvénykompozícióval* legyen összhangban: ezt a témakört majd a következő szakaszban részletezzük.

Mátrixszorzás:

Legyenek $A = [a_{kj}] \in \mathbf{M}_{n\times m}$, $B = [b_{kj}] \in \mathbf{M}_{m\times r}$ tetszőleges mátrixok. Az A és B mátrixok $A \cdot B$ szorzatán azt a $C = [c_{kj}] \in \mathbf{M}_{n\times r}$ mátrixot értjük, melynek kj-edik elemét az alábbi összeg definiálja:

$$c_{kj} := \sum_{i=1}^{m} a_{ki} b_{ij},$$

azaz c_{kj} definíció szerint az A mátrix k-adik sorának és a B mátrix j-edik oszlopának skaláris szorzata.

Látjuk tehát, hogy nem minden mátrixpár esetében értelmes a fenti szorzás: ehhez az kell, hogy az első mátrix oszlopainak száma megegyezzék a második mátrix sorainak számával.

Ha nem okoz félreértést, akkor a mátrixszorzást jelölő szorzópontot a későbbiekben elhagyjuk.

4.1. Példa: Legyenek

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \text{ és } B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Akkor:

$$AB := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ és } BA := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

A példa azt is mutatja, hogy a mátrixszorzás általában – a valós számok közti szorzástól eltérően – *nem kommutatív* (még akkor sem, ha mindkét sorrendben vett szorzat értelmes). Ha két négyzetes mátrixra mégis teljesül, hogy $AB = BA$, akkor azt mondjuk, hogy A és B *felcserélhetők*. A példa további tanulsága, hogy a szorzatmátrix akkor is lehet zérus, ha egyik tényezője sem az. Ettől eltekintve, a mátrixszorzás teljesíti a szokásos műveleti azonosságokat:

4.3. Állítás: A mátrixszorzás asszociatív, azaz, ha $A \in \mathbf{M}_{n\times m}$, $B \in \mathbf{M}_{m\times p}$, $C \in \mathbf{M}_{p\times q}$ tetszőleges mátrixok, akkor $(AB)C = A(BC)$. Továbbá a mátrixszorzás az összeadás felett disztributív: ha A,$B \in \mathbf{M}_{n\times m}$, $C \in \mathbf{M}_{m\times p}$, akkor $(A + B)C = AC + BC$, és ha A,$B \in \mathbf{M}_{n\times m}$, $C \in \mathbf{M}_{p\times n}$, akkor $C(A + B) = CA + CB$.

Bizonyítás: Az $(AB)C$ mátrix kj-edik eleme: $((AB)C)_{kj} = \sum_{i=1}^{p}(AB)_{ki}c_{ij} = \sum_{i=1}^{p}\sum_{r=1}^{m} a_{kr}b_{ri}c_{ij}$, míg az $A(BC)$ mátrix kj-edik eleme: $(A(BC))_{kj} = \sum_{u=1}^{m} a_{ku}(BC)_{uj}c_{ij} = \sum_{u=1}^{m}\sum_{v=1}^{p} a_{ku}b_{uv}c_{vj}$. A disztributivitás igazolását – gyakorlatképpen – az Olvasóra hagyjuk.

A szorzás nem-kommutatív voltának további következménye, hogy egy sor egyszerű azonosság, mely számokra igaz, érvényét veszti mátrixok esetében. Így pl. ha A,$B \in \mathbf{M}_{n\times n}$, akkor általában $(A + B)^2 \neq A^2 + 2AB + B^2$: egyenlőség csak akkor áll fenn, ha A és B felcserélhetők. Egyébként pedig csak annyi igaz, hogy $(A + B)^2 = A^2 + AB + BA + B^2$.

A következő állítás, mely a mátrixszorzás definíciójából nyomban adódik, a zérus- és az egységmátrix különleges szerepére mutat rá: a négyzetes mátrixok között ezek úgy viselkednek, mint a számok közt a 0 és az 1. Pontosabban:

4.4. Állítás: Az $n \times n$-es zérusmátrix ($\mathbf{0}$), és az $n \times n$-es egységmátrix (I) tetszőleges $A \in \mathbf{M}_{n\times n}$ mátrixszal felcserélhető, éspedig $A\mathbf{0} = \mathbf{0}A = \mathbf{0}$, és $AI = IA = A$.

Jegyezzük még meg az alábbi speciális mátrixszorzásokat:

- (a) Azonos méretű négyzetes mátrixok szorzata ugyanolyan méretű négyzetes mátrix.
- (b) $n \times n$-es mátrix és $n \times 1$-es oszlopvektor szorzata $n \times 1$-es oszlopvektor.
- (c) $1 \times n$-es sorvektor szorzata $n \times 1$-es oszlopvektorral egy 1×1-es mátrixot, azaz egyetlen számot eredményez (skaláris szorzat).
- (d) $n \times 1$-es oszlopvektor szorzata $1 \times n$-es sorvektorral egy $n \times n$-es mátrixot ad (*diadikus szorzat).*

4.3. Mátrixszorzás és lineáris leképezések

Nyilvánvaló, hogy a rendezett szám-n-esek és az $n \times 1$-es mátrixok (oszlopvektorok) között kölcsönösen egyértelmű megfeleltetés létesíthető. Ennélfogva a továbbiakban $\mathbf{R}^n$ elemeit azonosítjuk $\mathbf{M_{n\times 1}}$ elemeivel: egy-egy rendezett szám-n-est akár $\mathbf{R}^n$, akár $\mathbf{M}_{n\times 1}$ elemének is tekinthetünk. Ilyen értelemben, ha adott egy $A \in \mathbf{M}_{n\times m}$ mátrix, és $x \in \mathbf{R}^m$ tetszőleges, akkor az $x \to Ax$ hozzárendelés egy $\mathbf{R}^m \to \mathbf{R}^n$ leképezést definiál. A mátrixszorzás előzőekben részletezett tulajdonságai miatt ez a leképezés *lineáris*. Azaz minden $n \times m$-es mátrix felfogható egy $\mathbf{R}^m \to \mathbf{R}^n$ lineáris leképezésnek. Ha az A mátrix valósítja meg a leképezést, akkor azt mondjuk, hogy *a leképezés mátrixa* A.

Megfordítva, minden $\mathbf{R}^m \to \mathbf{R}^n$ lineáris leképezéshez található egy $n \times m$-es mátrix, mely a fenti értelemben előállítja ezt a lineáris leképezést. Valóban, legyen $A : \mathbf{R}^m \to \mathbf{R}^n$ egy lineáris leképezés. Jelölje $e_1, e_2, ..., e_m$ a standard bázist $\mathbf{R}^m$-ben, és tekintsük azt az $n \times m$-es mátrixot, melynek oszlopai az $Ae_1, Ae_2, ..., Ae_m$ oszlopvektorok: ekkor ez a mátrix épp az A lineáris leképezés mátrixa.

Ily módon kölcsönösen egyértelmű megfeleltetést létesítettünk $\mathbf{M}_{n\times m}$ elemei és az $\mathbf{R}^m \to \mathbf{R}^n$ lineáris leképezések között. Ugyanazt az objektumot mátrixnak avagy lineáris leképezésnek fogjuk tekinteni a továbbiakban, aszerint, hogy melyik kényelmesebb.

Speciális eset: Tekintsük az $I \in \mathbf{M}_{n\times n}$ egységmátrixot. Tetszőleges $x \in \mathbf{R}^n$ esetén könnyen ellenőrizhetően $Ix = x$, azaz az egységmátrix épp az $\mathbf{R}^n$ tér identikus leképezésének mátrixa.

Legyenek $A \in \mathbf{M}_{n\times m}$, $B \in \mathbf{M}_{m\times r}$ tetszőleges mátrixok, $x \in \mathbf{R}^r$ tetszőleges (oszlop)vektor. A mátrixszorzás asszociativitása miatt

$$A(Bx) = (AB)x$$

Ugyanakkor, A-t és B-t leképezésként fogva fel, $A(Bx)$ az $A \circ B$ összetett függvény az x vektorra alkalmazva. Következésképp az $A \circ B$ összetett függvény mátrixa épp az AB szorzatmátrix.

Most néhány példát mutatunk lineáris leképezések mátrixának meghatározására.

4.2. Példa: Az $\mathbf{R}^2$ sík minden pontjához rendeljük hozzá az origó körüli, adott t szögű elforgatás útján kapott pontot. Már láttuk, hogy ezzel egy $\mathbf{R}^2 \to \mathbf{R}^2$ lineáris leképezést definiálunk. Határozzuk meg ennek mátrixát.

Megoldás: Legyen $(x,y) \in \mathbf{R}^2$ tetszőleges, $x := r\cos\theta$, $y := r\sin\theta$, ahol r az (x,y) pont helyvektorának hossza, θ pedig az irányszöge. Akkor az elforgatott pont koordinátái nyilván:

$$x' = r\cos(\theta+t) = r(\cos\theta\cos t - \sin\theta\sin t) = x\cos t - y\sin t$$

$$y' = r\sin(\theta+t) = r(\sin\theta\cos t + \cos\theta\sin t) = x\sin t + y\cos t$$

ahonnan tehát

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Következésképp az origó körüli t szögű elforgatás mátrixa az alábbi 2×2-es mátrix (ún. *forgatómátrix*):

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

4.3. Példa: Az $\mathbf{R}^3$ tér minden pontjához rendeljük hozzá a pontnak egy adott, az origóra illeszkedő síkra vett ortogonális vetületét. Már tudjuk, hogy ezzel egy $\mathbf{R}^3 \to \mathbf{R}^3$ lineáris leképezést definiáltunk. Határozzuk meg ennek mátrixát.

Megoldás: Jelölje $\mathbf{n} := (n_1,n_2,n_3) \in \mathbf{R}^3$ a szóban forgó sík normálvektorát, egyszerűség kedvéért tegyük fel, hogy ez egységnyi hosszú, azaz $||\mathbf{n}|| = 1$. Akkor egy tetszőleges $\mathbf{x} := (x_1,x_2,x_3) \in \mathbf{R}^3$ pont $\mathbf{n}$ irányú ortogonális vetülete: $P\mathbf{x} := \langle \mathbf{x},\mathbf{n}\rangle \cdot \mathbf{n}$, azaz:

$$P\mathbf{x} = (x_1n_1 + x_2n_2 + x_3n_3) \cdot \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} =$$

$$= \begin{pmatrix} n_1^2 & n_1n_2 & n_1n_3 \\ n_1n_2 & n_1^2 & n_2n_3 \\ n_1n_3 & n_2n_3 & n_3^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Mivel pedig a síkra vett ortogonális vetület helyvektora $\mathbf{x} - P\mathbf{x}$, azért ennek mátrixa $I - P$, ahol I a 3×3-as egységmátrix, P pedig a fenti egyenlőségből adódó

$$\begin{pmatrix} n_1^2 & n_1n_2 & n_1n_3 \\ n_1n_2 & n_2^2 & n_2n_3 \\ n_1n_3 & n_2n_3 & n_3^2 \end{pmatrix}$$

mátrix (ami épp a normálvektornak önmagával vett diadikus szorzata).

4.4. Példa: Legyen $\mathbf{a} \in \mathbf{R}^3$ adott vektor és minden $\mathbf{x} \in \mathbf{R}^3$-hez rendeljük hozzá az $\mathbf{x} \times \mathbf{a}$ vektort. Ezzel lineáris leképezést definiáltunk. Határozzuk meg ennek mátrixát.

Megoldás: A vektoriális szorzat definíciójából adódóan:

$$\mathbf{x} \times \mathbf{a} = \begin{pmatrix} x_2a_3 - x_3a_2 \\ x_3a_1 - x_1a_3 \\ x_1a_2 - x_2a_1 \end{pmatrix} = \begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

tehát a vektoriális szorzás mátrixa:

$$\begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix}$$

4.4. Mátrixok inverze és determinánsa

Mátrix inverze:

Az $A \in \mathbf{M}_{n\times n}$ négyzetes mátrix *invezének* azt az $A^{-1} \in \mathbf{M}_{n\times n}$ mátrixot nevezzük, melyre $A^{-1}A = I$ teljesül. Ha ilyen A^{-1} mátrix létezik, akkor A-t *regulárisnak*, ellenkező esetben *szingulárisnak* nevezzük.

Nem minden négyzetes mátrixnak van inverze. Így pl. a zérusmátrixnak nincs, ui. bármilyen mátrixszal szorozva balról a zérusmátrixot, a szorzat mindig a zérusmátrix, és sohasem az egységmátrix. Ugyanakkor pl. az egységmátrix reguláris, inverze önmagával egyezik.

Az inverz a reciprok fogalmának erős általánosítása. Ugyanakkor, ha az $A \in \mathbf{M}_{n\times n}$ mátrixot $\mathbf{R}^n \to \mathbf{R}^n$ lineáris leképezésnek tekintjük, akkor A^{-1} *mint leképezés, megegyezik* A *inverzével* (ez indokolja az elnevezést is). Valóban, $A^{-1}A = I$ következtében minden $x \in \mathbf{R}^n$ esetén $A^{-1}(Ax) = Ix = x$. Következésképp egy négyzetes A mátrix pontosan akkor reguláris, ha mint leképezés egy-egyértelmű, és ekkor az inverz egyértelműen meghatározott (bár ez a definícióból nem látszik azonnal). Továbbá, ha A reguláris, akkor A^{-1} is az, és $(A^{-1})^{-1} = A$. Ez utóbbiból pedig még az is nyomban adódik, hogy ha A reguláris mátrix, akkor az A^{-1} inverz mátrixszal A-t nemcsak balról szorozva kapunk egységmátrixot, hanem ugyanez áll a jobbról való szorzásra is: $AA^{-1} = I$. Röviden, egy reguláris mátrix mindig felcserélhető az inverzével.

Szorzat inverze kiszámítható a tényezők inverzeinek segítségével (a reciprok számításához hasonlóan), de mivel a mátrixszorzás általában nem kommutatív, azért a tényezők sorrendje nem közömbös:

4.5. Állítás: Ha $A,B \in \mathbf{M}_{n\times n}$ mindketten regulárisak, akkor AB is az, és $(AB)^{-1} = B^{-1}A^{-1}$

Bizonyítás: Valóban, az asszociativitást felhasználva: $B^{-1}A^{-1}AB = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$.

Megjegyezzük, hogy ugyanakkor $(A+B)^{-1}$ *sohasem* egyenlő $A^{-1}+B^{-1}$-gyel!

A regularitás és az inverz fogalmának jelentőségére az alábbi fontos példa mutat rá. Legyen $A = [a_{kj}] \in \mathbf{M}_{n\times n}$ adott mátrix, $b \in \mathbf{R}^n$ adott oszlopvektor. Akkor az $Ax = b$ mátrixegyenlet ($x \in \mathbf{R}^n$) a mátrixszorzás definíciója miatt ekvivalens az alábbi n-ismeretlenes lineáris egyenletrendszerrel:

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n$$

Ha a rendszer A mátrixa reguláris, és ismerjük az A^{-1} inverz mátrixot, akkor A^{-1}-gyel balról szorozva a fenti egyenlet mindkét oldalát, kapjuk, hogy $A^{-1}Ax = A^{-1}b$, azaz $x = A^{-1}b$; az egyenletrendszer megoldását tehát explicit formula adja. Sajnos, nagyméretű egyenletrendszerek esetében az inverz mátrix meghatározása semmivel sem könnyebb probléma, mint az eredeti egyenletrendszer megoldása. (Az inverz mátrix meghatározására általános algoritmust a 3.6.szakaszban adunk.) Mindazonáltal az $x = A^{-1}b$ explicit megoldóformula hasznos lehet, ha ugyanazt az egyenletrendszert egymás után több különböző jobb oldal mellett kell megoldani.

Néha, bizonyos speciális mátrixosztályok esetén, az inverz egyszerűen meghatározható. Erre mutatunk két példát:

1. *Diagonálmátrix inverze*: Ha a diagonálmátrix diagonálelemei mind 0-tól különböznek, akkor a mátrix reguláris, és

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}^{-1} = \begin{pmatrix} a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}^{-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn}^{-1} \end{pmatrix}$$

Ez a mátrixszorzás definíciójának felhasználásával könnyen ellenőrizhető.

2. *Forgatómátrix inverze*: A 2×2-es forgatómátrixok mindig reguláris, éspedig:

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}^{-1} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$$

azaz egy t szögű forgatás inverze megegyezik egy $(-t)$ szögű forgatással, a szemlélettel teljes összhangban.

Végül néhány állítást mondunk ki a reguláris mátrixok jellemzésére:

4.6. Állítás: Egy $A \in \mathbf{M}_{n\times n}$ mátrix pontosan akkor reguláris, ha (mint leképezés) csak a zérusvektort viszi a zérusvektorba, azaz $Ax = \mathbf{0}$ csak úgy lehetséges, ha $x = \mathbf{0}$.

Bizonyítás: Legyen A reguláris. Nyilván mindig $A\mathbf{0} = \mathbf{0}$, ha pedig $x \neq \mathbf{0}$, akkor az egy-egyértelműség miatt $Ax \neq \mathbf{0}$. Megfordítva, tegyük fel, hogy A csak a zérusvektort viszi a zérusvektorba Megmutatjuk, hogy ekkor A egy-egyértelmű, azaz reguláris. Legyen $x_1 \neq x_2$, akkor $x_1 - x_2 \neq \mathbf{0}$, így $A(x_1 - x_2) = Ax_1 - Ax_2 \neq \mathbf{0}$, azaz $Ax_1 \neq Ax_2$, tehát A valóban egy-egyértelmű.

Az állítás lényege, hogy a linearitás miatt elég az egy-egyértelműséget egyetlen vektor, nevezetesen a zérusvektor esetében ellenőrizni.

Magtér:

Egy $A \in \mathbf{M}_{n\times n}$ mátrix ill. lineáris leképezés *magterének* a

$$\ker A := \{x \in \mathbf{R}^n : Ax = \mathbf{0}\}$$

halmazt nevezzük, mely A linearitásából adódóan mindig altér $\mathbf{R}^n$-ben.

Ezzel a jelöléssel az előző állítás röviden úgy fogalmazható meg, hogy az $A \in \mathbf{M}_{n\times n}$ mátrix pontosan akkor reguláris, ha $\ker A = \{\mathbf{0}\}$.

4.7. Állítás: Egy $A \in \mathbf{M}_{n\times n}$ mátrix pontosan akkor reguláris, ha (mint leképezés), lineárisan független vektorokat lineárisan független vektorokba visz.

Bizonyítás: Legyen A reguláris, és legyenek $x_1,x_2,...,x_m \in \mathbf{R}^n$ lineárisan függetlenek ($m \le n$). Tekintsünk egy tetszőleges $\lambda_1Ax_1 + \lambda_2Ax_2 + ... + \lambda_mAx_m$ lineáris kombinációt. Ha ez a zérusvektor, akkor $A(\lambda_1x_1 + \lambda_2x_2 + ... + \lambda_mx_m) = \mathbf{0}$, így az előző állítás értelmében szükségképp $\lambda_1x_1 + \lambda_2x_2 + ... + \lambda_mx_m = \mathbf{0}$. Ám az x_j vektorok lineáris függetlensége miatt ezért $\lambda_1 = \lambda_2 = ... = \lambda_m = 0$, azaz $Ax_1,Ax_2,...,Ax_m$ valóban lineárisan függetlenek. Megfordítva, ha A lineárisan független vektorokat lineárisan független vektorokba visz, akkor semmilyen $x \ne \mathbf{0}$ vektort nem vihet a zérusvektorba: ha ui. x előáll $x = \lambda_1e_1 + \lambda_2e_2 + ... + \lambda_ne_n$ (nemtriviális!) lineáris kombinációként, ahol $e_1,e_2,...,e_n$ bázist alkotnak $\mathbf{R}^n$-ben, akkor Ax előáll az $Ae_1,Ae_2,...,Ae_n$ vektorok nemtriviális lineáris kombinációjaként: $Ax = \lambda_1Ae_1 + \lambda_2Ae_2 + ... + \lambda_nAe_n$, így nem lehet a zérusvektorral egyenlő.

Az állítás egyszerű következménye, hogy egy $A \in \mathbf{M}_{n\times n}$ mátrix pontosan akkor reguláris, ha képterének dimenziója n-nel egyenlő, azaz, ha a képtér a teljes $\mathbf{R}^n$-nel egyezik. Valóban, a képtér dimenziója n-nél nagyobb nyilván nem lehet, így ha $e_1,e_2,...,e_n$ bázist alkot $\mathbf{R}^n$-ben, és A reguláris, akkor $Ae_1,Ae_2,...,Ae_n$ lineárisan független rendszer lévén maga is bázist alkot $\mathbf{R}^n$-ben, tehát a képtér n-dimenziós. Az Olvasóra bízzuk annak végiggondolását, hogy ha $A \in \mathbf{M}_{n\times n}$ szinguláris, akkor képterének dimenziója n-nél határozottan kisebb.

A regularitás eldönthető egy, a fentieknél sokkal gépiesebb (bár meglehetősen számításigényes) módon is. Ehhez bevezetjük a mátrix determinánsának fogalmát. A definíciót rekurzív módon fogalmazzuk meg:

Determináns:

Az egyetlen $a \in \mathbf{R}$ szám alkotta 1×1-es mátrix determinánsa legyen maga az a szám. Ha pedig az $(n-1)$-edrendű mátrixok determinánsát már definiáltuk, akkor tetszőleges $A := [a_{kj}] \in \mathbf{M}_{n\times n}$ mátrix esetén definiáljuk az A mátrix determinánsát a

$$\det A := a_{11}D_{11} - a_{12}D_{12} + a_{13}D_{13} - \ldots \pm a_{1n}D_{1n},$$

formulával, ahol D_{1j} jelentse annak az $(n-1)$-edrendű mátrixnak a determinánsát, melyet A-ból az első sor és a j-edik oszlop elhagyásával kaptunk.

A determináns tehát mindig egyetlen *szám*: jegyezzük meg, hogy ha $A := [a_{kj}] \in \mathbf{M}_{2\times 2}$ egy másodrendű mátrix, akkor $\det A = a_{11}a_{22} - a_{12}a_{21}$. Bizonyítás nélkül megemlítjük, hogy $\det(AB) = \det(A) \cdot \det(B)$ mindig teljesül. Ugyanakkor általában $\det(A+B) \neq \det(A) + \det(B)$!

4.5. Példa:

$$\det \begin{pmatrix} 1 & -2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} =$$

$$= 1 \cdot \begin{pmatrix} -2 & 1 \\ 0 & 3 \end{pmatrix} - (-2) \cdot \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = 6.$$

A definíció már viszonylag kis n értékek mellett is gyakorlati szempontból használhatatlan. Megmutatható, hogy a determinánsnak a definíció szerinti kiszámítása $n!$-nál is több műveletet igényel. Így pl. egy mindössze 100-adrendű determináns kiszámítása – bármilyen ma és a belátható jövőben használatos számítógépek mellett is! – több időt venne igénybe, mint a Világegyetem jelenleg ismeretes életkora. Szerencsére vannak algoritmusok, melyek ennél lényegesen kevesebb művelet árán számítják ki a determinánst. Ezek részleteivel azonban e jegyzet keretein belül nem foglalkozhatunk.

Megmutatható, hogy a determináns kiszámításának rekurzív definíciója nem kell, hogy az első sorhoz kötődjék. Bármelyik, pl. a k-adik sor esetén

$$\det A = (-1)^{k+1} a_{k1} D_{k1} + (-1)^{k+2} a_{k2} D_{k2} + ... + (-1)^{k+n} a_{kn} D_{kn},$$

ahol D_{kj} jelenti annak az $(n-1)$-edrendű mátrixnak a determinánsát, melyet A-ból a k-adik sor és a j-edik oszlop elhagyásával kaptunk (*sor szerinti kifejtés*). Hasonló formula érvényes az oszlopok esetére is:

$$\det A = (-1)^{k+1} a_{1k} D_{1k} + (-1)^{k+2} a_{2k} D_{2k} + ... + (-1)^{k+n} a_{nk} D_{nk}$$

(*oszlop szerinti kifejtés*).

A determináns számunkra legfontosabb tulajdonságát az alábbi tétel mutatja, melyet bizonyítás nélkül adunk meg (a másodrendű eset a feladatok között szerepel):

4.1. Tétel: Egy $A \in \mathbf{M}_{n\times n}$ mátrix pontosan akkor reguláris, ha $\det A \neq 0$, másszóval, pontosan akkor szinguláris, ha $\det A = 0$.

A determináns fogalma a *vektoriális szorzatok* kiszámítását könnyen áttekinthetővé teszi. Egyszerű számolással igazolható, hogy ha $\mathbf{x} = (x_1,x_2,x_3)$ és $\mathbf{y} = (y_1,y_2,y_3)$ tetszőleges $\mathbf{R}^3$-beli vektorok, akkor a formális

$$\det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$$

determináns épp az $\mathbf{x} \times \mathbf{y}$ vektoriális szorzatvektorral egyenlő.

4.5. Lineáris egyenletrendszerek megoldhatósága

Ebben a szakaszban a mátrixok regularitásának és az

$$Ax = b$$

egyenlet megoldhatóságának kapcsolatát vizsgáljuk, ahol $A = [a_{kj}] \in \mathbf{M}_{n\times n}$ adott mátrix, $b \in \mathbf{R}^n$ adott jobb oldal.

Már láttuk, hogy a fenti tömör mátrixegyenlet ekvivalens az alábbi n-ismeretlenes lineáris egyenletrendszerrel:

$$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n = b_n$$

Az egyenletrendszert *homogénnak* nevezzük, ha $b = \mathbf{0}$, azaz $b_1 = b_2 = ... = b_n = 0$. Nyilvánvaló, hogy ekkor $x_1 = x_2 = ... = x_n = 0$ mindig megoldás: ezt *triviális megoldásnak* nevezzük, míg a homogén egyenlet minden olyan megoldását, ahol legalább egy x_j zérustól különbözik, *nemtriviális megoldásnak* nevezzük.

A homogén egyenletek esetében a jellemző probléma az, hogy létezik-e nemtriviális megoldás, míg az inhomogén egyenlet esetén az a kérdés, hogy van-e egyáltalán megoldása, és ha igen, akkor hány.

Nyilvánvaló, hogy a megoldás létezése (tetszőleges jobb oldal mellett) azzal ekvivalens, hogy A képtere a teljes $\mathbf{R}^n$ tér. A homogén egyenlet nemtriviális megoldásának létezése pedig, más megfogalmazásban, azt jelenti, hogy az A leképezés egy zérusvektortól különböző vektort is a zérusvektorba visz. Ily módon az előző szakasz eredményei minden további nélkül alkalmazhatók, és az alábbi fontos tételekhez jutunk:

4.2. Tétel: Az $A \in \mathbf{M}_{n\times n}$ mátrix pontosan akkor reguláris, ha az

$$Ax = b$$

egyenletnek minden jobb oldal mellett létezik megoldása. Ekkor a megoldás egyértelmű is (éspedig $A^{-1}b$-vel egyenlő).

4.3. Tétel: Az $A \in \mathbf{M}_{n\times n}$ mátrix pontosan akkor reguláris, ha az

$$Ax = \mathbf{0}$$

homogén egyenletnek csak a triviális megoldása létezik. Másszóval, A pontosan akkor szinguláris, ha a homogén egyenletnek létezik nemtriviális megoldása. Ekkor pedig végtelen sok nemtriviális megoldás is létezik (bármely x megoldás esetén annak tetszőleges konstansszorosa is megoldás).

A tételek alkalmazásaként egy egyszerű elegendő feltételt mutatunk egy mátrix szingularitására:

4.1. Következmény: Ha az $A \in \mathbf{M}_{n\times n}$ mátrixnak valamelyik sora vagy oszlopa csupa 0-ból áll, akkor a mátrix szinguláris.

Bizonyítás: Ha valamelyik, pl. a k-adik sor csupa 0, akkor az e_k standard báziselem mellett az $Ax = e_k$ egyenletnek nincs megoldása, hiszen Ax k-adik eleme biztosan 0, ezért a mátrix szinguláris. Ha pedig valamelyik, pl. a k-adik oszlop csupa 0, akkor $Ae_k = \mathbf{0}$, azaz a homogén egyenletnek van nemtriviális megoldása, ezért a mátrix ekkor is szinguláris.

4.6. Megoldási algoritmus: a Gauss-elimináció

Tekintsük az alábbi lineáris egyenletrendszert:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n &= b_3 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nn}x_n &= b_n
\end{aligned}$$

ahol feltesszük, hogy az A mátrix reguláris. A Gauss-elimináció (vagy kiküszöböléses módszer) lépései a következők:

1. Osszuk le az 1. egyenletet az a_{11} együtthatóval:

$$\begin{aligned}
x_1 + a'_{12}x_2 + a'_{13}x_3 + \ldots + a'_{1n}x_n &= b'_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n &= b_3 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nn}x_n &= b_n
\end{aligned}$$

2. Az 1. sor a_{k1}-szeresét vonjuk ki a k-adik sorból ($k = 2,3,\ldots,n$), ezáltal kiküszöböljük x_1-et a 2.,3.,...,n. egyenletből. Eredményül az alábbi szerkezetű egyenletrendszerhez jutunk:

$$\begin{aligned}
x_1 + a'_{12}x_2 + a'_{13}x_3 + \ldots + a'_{1n}x_n &= b'_1 \\
a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n &= b'_2 \\
a'_{32}x_2 + a'_{33}x_3 + \ldots + a'_{3n}x_n &= b'_3 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a'_{n2}x_2 + a'_{n3}x_3 + \ldots + a'_{nn}x_n &= b'_n
\end{aligned}$$

3. A 2.,3.,...,n. egyenlet már csak $(n-1)$ ismeretlent tartalmaz, így az 1., 2. pont lépéseit megismételhetjük: x_2-t kiküszöböljük a 3.,4.,...,n. egyenletből, majd hasonlóan, x_3-at a 4.,5.,...,n. egyenletből, és így tovább. Végül a következő alakú egyenletrendszert kapjuk:

$$\begin{aligned} x_1 + \tilde{a}_{12}x_2 + \tilde{a}_{13}x_3 + \ldots + \tilde{a}_{1n}x_n &= \tilde{b}_1 \\ x_2 + \tilde{a}_{23}x_3 + \ldots + \tilde{a}_{2n}x_n &= \tilde{b}_2 \\ x_3 + \ldots + \tilde{a}_{3n}x_n &= \tilde{b}_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ x_n &= \tilde{b}_n \end{aligned}$$

4. Az utolsó egyenletből x_n máris ismert: visszahelyettesítve az $(n-1)$-edik egyenletbe, x_{n-1} számítható; x_n-et és x_{n-1}-et visszahelyettesítve az $(n-2)$-edik egyenletbe, x_{n-2} számítható, és így tovább, így az egyes ismeretleneket index szerint csökkenő sorrendben határozzuk meg. Ezt a műveletsort a Gauss-elimináció *visszahelyettesítési* részének, míg az előző lépéseket *eliminációs résznek* nevezzük.

Az algoritmust az alábbi egyszerű példán szemléltetjük:

4.6. Példa: Oldjuk meg az alábbi egyenletrendszert:

$$\begin{array}{rcrcrcr} 2x_1 & - & 6x_2 & + & 10x_3 & = & -12 \\ 2x_1 & - & 5x_2 & + & 3x_3 & = & -4 \\ 3x_1 & - & 2x_2 & + & x_3 & = & 3 \end{array}$$

Megoldás: Osszuk le az 1. egyenletet 2-vel:

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

Az 1. egyenlet 2-szeresét vonjuk ki a 2. egyenletből, majd az 1. egyenlet 3-szorosát vonjuk ki a 3. egyenletből, ezzel a 2. és 3. egyenletből kiküszöböljük x_1-et:

$$\begin{array}{rcrcrcr} x_1 & - & 3x_2 & + & 5x_3 & = & -6 \\ & & x_2 & - & 7x_3 & = & 8 \\ & & 7x_2 & - & 14x_3 & = & 21 \end{array}$$

A 2. egyenletet már nem kell x_2 együtthatójával leosztani, lévén az 1-gyel egyenlő. Vonjuk ki a 2. egyenlet 7-szeresét a 3. egyenletből, ezzel a 3. egyenletből x_2-t is kiküszöböltük:

$$\begin{array}{rcrcrcr} x_1 & - & 3x_2 & + & 5x_3 & = & -6 \\ & & x_2 & - & 7x_3 & = & 8 \\ & & & & 35x_3 & = & -35 \end{array}$$

Elosztva a 3. egyenletet 35-tel, az eliminációs részt befejeztük:

$$\begin{array}{ccccccc} x_1 & - & 3x_2 & + & 5x_3 & = & -6 \\ & & x_2 & - & 7x_3 & = & 8 \\ & & & & x_3 & = & -1 \end{array}$$

Az utolsó egyenletből x_3 már ki van számítva. Visszahelyettesítve a 2. egyenletbe, innen x_2 is számítható (ugyanide jutunk, ha a 3. egyenlet 7-szeresét hozzáadjuk a 2. egyenlethez):

$$\begin{array}{ccccccc} x_1 & - & 3x_2 & + & 5x_3 & = & -6 \\ & & x_2 & & & = & 1 \\ & & & & x_3 & = & -1 \end{array}$$

Végül, x_2-t és x_3-t az 1. egyenletbe helyettesítve vissza, x_1 is számítható:

$$\begin{array}{ccccc} x_1 & & & = & 2 \\ & x_2 & & = & 1 \\ & & x_3 & = & -1 \end{array}$$

Ezzel az egyenletrendszer megoldását előállítottuk. Visszahelyettesítéssel meggyőződhetünk róla, hogy az így nyert megoldás valóban kielégíti az eredeti egyenletrendszert.

Vegyük észre, hogy a számítás végrehajtásához az x_1, x_2, x_3 szimbólumokat és az egyenlőségjeleket újra meg újra leírni felesleges: a számításokat voltaképpen csak az együtthatókon, azok mátrixán hajtjuk végre. Így a fenti számítási lépések az alábbi tömör formába írhatók (a mátrix utolsó oszlopa előtti függőleges vonal csak a jobb áttekinthetőséget szolgálja):

$$\left(\begin{array}{rrr|r} 2 & -6 & 10 & -12 \\ 2 & -5 & 3 & -4 \\ 3 & -2 & 1 & 3 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & -3 & 5 & -6 \\ 2 & -5 & 3 & -4 \\ 3 & -2 & 1 & 3 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & -3 & 5 & -6 \\ 0 & 1 & -7 & 8 \\ 0 & 7 & -14 & 21 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & -3 & 5 & -6 \\ 0 & 1 & -7 & 8 \\ 0 & 0 & 35 & -35 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & -3 & 5 & -6 \\ 0 & 1 & -7 & 8 \\ 0 & 0 & 1 & -1 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & -3 & 5 & -6 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right)$$

Megmutatható, hogy a Gauss-elimináció végrehajtása kb. $\frac{2}{3}n^3$ műveletet igényel (ez a becslés annál pontosabb, minél nagyobb a mátrix n rendje), ami azt mutatja, hogy numerikus szempontból a Gauss-elimináció nem "olcsó": ha az ismeretlenek száma duplájára nő, akkor a szükséges műveletszám kb. *nyolcszorosára* emelkedik.

Az algoritmust ebben a formában nem mindig lehet végrehajtani, ui. lehetséges, hogy valamelyik együttható, mellyel osztanunk kéne, 0-val egyenlő. Legyen pl. $a_{11} = 0$. Ekkor az első és valamelyik későbbi egyenlet cseréjével elérhető, hogy az új egyenletrendszer első egyenletében x_1 együtthatója ne legyen 0, ellenkező esetben a mátrix első oszlopa csupa 0-ból állna, de ekkor a mátrix szinguláris volna, kiinduló feltételünkkel ellentétben. Ugyanez áll az elimináció további lépéseiben fellépő (egyre kisebb méretű) egyenletrendszerekre. A numerikus számítás pontosságának szempontjából az a célszerű, hogy a együtthatók, melyekkel leosztjuk az egyenleteket (az ún *főegyütthatók*), abszolút értékben minél nagyobbak legyenek (hogy az osztás számítási hibája minél kisebb legyen). Ezért az egyenletek cseréjekor célszerű az aktuális, mondjuk k-adik egyenletet azzal a későbbi pl. r-edik egyenlettel felcserélni, melyre $|a_{rk}|$ a lehető legnagyobb ($r = k,k + 1,...,n$) még akkor is, ha $a_{kk} \neq 0$. Ezt a megoldási stratégiát *részleges főelemkiválasztásnak* nevezzük, és ez már minden reguláris A mátrix esetén működik. Valamivel több számítási munkával jár, de még nagyobb pontosságot biztosít a *teljes főelemkiválasztás*, amikor a k-adik egyenlettel való eliminációs lépéskor az összes hátralévő $|a_{pq}|$ érték maximumát keressük ($p,q = k,k+1,...,n$), és ekkor nemcsak az egyenleteket cseréljük meg, hanem az ismeretlenek sorrendjét is megváltoztatjuk, hogy a főegyüttható az imént meghatározott maximális abszolút értékű elem legyen.

A Gauss-elimináció egy válfaja a *Gauss–Jordan-elimináció*, amikor az aktuális pl. k-adik egyenlet segítségével nemcsak a későbbi egyenletekből küszöböljük ki a k-adik ismeretlent, hanem a *megelőzőekből* is. Így a visszahelyettesítési lépések elmaradnak, és az elimináció befejeztével azonnal nyerjük az ismeretlenek értékeit. (Egyszerűsége ellenére a Gauss–Jordan-elimináció műveletigénye nagyobb a Gauss-elimináció műveletigényénél).

4.7. Példa: Tekintsük az előző példa egyenletrendszerét. Az algoritmus első két lépése egyezik a Gauss-elimináció első két lépésével, eltérés csak a 3. lépéstől van:

$$\left(\begin{array}{rrr|r} 2 & -6 & 10 & -12 \\ 2 & -5 & 3 & -4 \\ 3 & -2 & 1 & 3 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & -3 & 5 & -6 \\ 2 & -5 & 3 & -4 \\ 3 & -2 & 1 & 3 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & -3 & 5 & -6 \\ 0 & 1 & -7 & 8 \\ 0 & 7 & -14 & 21 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 0 & -16 & 18 \\ 0 & 1 & -7 & 8 \\ 0 & 0 & 35 & -35 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & 0 & -16 & 18 \\ 0 & 1 & -7 & 8 \\ 0 & 0 & 1 & -1 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right)$$

A Gauss-elimináció szinguláris mátrixú egyenletrendszerek megoldására is alkalmas. Ekkor az a jellemző, hogy az elimináció valamelyik lépésében az egyik egyenlet *összes* együtthatója zérussá válik. Amennyiben az illető egyenlet jobb oldala nem zérus, akkor az egyenletrendszernek nincs megoldása; ha a jobb oldal is zérus, akkor van megoldás, sőt, ekkor mindig *végtelen sok* megoldás van. Ekkor ui. valamelyik ismeretlen (esetleg több is) szabadon megválasztható, a többi pedig ezek függvényében fejezhető ki.

Az elmondottakat egy példán szemléltetjük:

4.8. Példa: Oldjuk meg a következő egyenletrendszert:

$$\begin{array}{rcrcrcl} x_1 & - & 2x_2 & + & x_3 & = & 1 \\ -2x_1 & + & x_2 & + & x_3 & = & 4 \\ x_1 & + & x_2 & - & 2x_3 & = & 1 \end{array}$$

Megoldás: A Gauss-elimináció lépéseit az előző példákban megismert tömör jelölésmóddal írjuk le:

$$\left(\begin{array}{rrr|r} 1 & -2 & 1 & 1 \\ -2 & 1 & 1 & 4 \\ 1 & 1 & -2 & 1 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & -2 & 1 & 1 \\ 0 & -3 & 3 & 6 \\ 0 & 3 & -3 & 0 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & -2 & 1 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 3 & -3 & 0 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & -2 & 1 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 6 \end{array}\right)$$

Az utolsó egyenlet együtthatói mind 0-val lettek egyenlők, de a jobb oldal nem zérus. Ez ellentmondás, így az egyenletrendszernek nincs megoldása.

Ha viszont a megfelelő homogén egyenletet tekintjük:

$$\begin{array}{rcrcrcl} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ -2x_1 & + & x_2 & + & x_3 & = & 0 \\ x_1 & + & x_2 & - & 2x_3 & = & 0 \end{array}$$

akkor már tudjuk, hogy van nemtriviális megoldás, hiszen a mátrix szinguláris (ezt akár a determináns zérus voltából láthatjuk, akár onnan, hogy ha a mátrix reguláris volna, az előző egyenletrendszernek is lenne, éspedig egyetlen megoldása). Lássuk, hogyan működik a Gauss-elimináció ebben az esetben:

$$\left(\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ -2 & 1 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & 3 & -3 & 0 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 3 & -3 & 0 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Az utolsó egyenlet a semmitmondó $0 = 0$ egyenlőséggé egyszerűsödött. Valamelyik, célszerűen az utolsó ismeretlent így tetszőlegesen megválaszthatjuk: $x_3 := t$, ahol $t \in \mathbf{R}$ tetszőleges szám. Ezt beírva a 3. egyenlet helyére, a visszahelyettesítések már nehézség nélkül elvégezhetők:

$$\left(\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & t \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & -2 & 0 & -t \\ 0 & 1 & 0 & t \\ 0 & 0 & 1 & t \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & 0 & 0 & t \\ 0 & 1 & 0 & t \\ 0 & 0 & 1 & t \end{array}\right)$$

Tehát végtelen sok nemtriviális megoldás van, és ezek általános alakja: $x_1 = t$, $x_2 = t$, $x_3 = t$.

Végül megmutatjuk, hogyan használható a Gauss-elimináció *mátrixinvertálásra*. Legyen $A \in \mathbf{M}_{n\times n}$ egy reguláris mátrix. Ekkor érvényes az

$$AA^{-1} = I$$

mátrixegyenlőség. Jelölje az egyelőre ismeretlen A^{-1} inverz mátrix oszlopait a_1, a_2, ... , a_n, az I egységmátrix oszlopait pedig e_1, e_2, ... ,e_n (ezek épp a standard bázis elemei $\mathbf{R}^n$-ben):

$$A \cdot \left(\begin{array}{c|c|c|c} a_1 & a_2 & \dots & a_n \end{array}\right) = \left(\begin{array}{c|c|c|c} e_1 & e_2 & \dots & e_n \end{array}\right)$$

A mátrixszorzás definíciója értelmében ez a mátrixegyenlőség ekvivalens n db vektoregyenlőséggel, éspedig:

$$Aa_k = e_k \qquad (k = 1,2,\ldots,n)$$

Azt kaptuk tehát, hogy *a fenti n db egyenletrendszert megoldva, a megoldásvektorokból mint oszlopokból összeállított mátrix épp az eredeti A mátrix inverzével egyezik.*

Tehát egy mátrixinverzióhoz n db speciális jobb oldalú, de azonos mátrixú egyenletrendszert kell megoldani. Ez történhet egyidejűleg is, mivel az eliminációs lépésekkel egyidőben a jobb oldalakon végrehajtandó műveleteket egyszerre *mindegyik* jobb oldallal megtehetjük. Az eljárás a fentiekben használt tömör jelölésmóddal igen szemléletes: az elimináció elején a bal oldali részmátrix az eredeti mátrix, a jobb oldali pedig az egységmátrix, míg az algoritmus befejeztével a bal oldali részmátrixból egységmátrix lesz, ekkor a jobb oldali részmátrix az inverz mátrixszal lesz egyenlő.

4.9. Példa: Számítsuk ki az alábbi mátrix inverzét:

$$A := \begin{pmatrix} -3 & -2 & 0 \\ 0 & 3 & 2 \\ -2 & 0 & 1 \end{pmatrix}$$

Megoldás: Az algoritmus lépései az eddigiekben alkalmazott tömör jelölésmóddal:

$$\left(\begin{array}{rrr|rrr} -3 & -2 & 0 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 1 & 0 \\ -2 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 2/3 & 0 & -1/3 & 0 & 0 \\ 0 & 3 & 2 & 0 & 1 & 0 \\ -2 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 2/3 & 0 & -1/3 & 0 & 0 \\ 0 & 3 & 2 & 0 & 1 & 0 \\ 0 & 4/3 & 1 & -2/3 & 0 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 2/3 & 0 & -1/3 & 0 & 0 \\ 0 & 1 & 2/3 & 0 & 1/3 & 0 \\ 0 & 4/3 & 1 & -2/3 & 0 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 2/3 & 0 & -1/3 & 0 & 0 \\ 0 & 1 & 2/3 & 0 & 1/3 & 0 \\ 0 & 0 & 1/9 & -2/3 & -4/9 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 2/3 & 0 & -1/3 & 0 & 0 \\ 0 & 1 & 2/3 & 0 & 1/3 & 0 \\ 0 & 0 & 1 & -6 & -4 & 9 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 2/3 & 0 & -1/3 & 0 & 0 \\ 0 & 1 & 0 & 4 & 3 & -6 \\ 0 & 0 & 1 & -6 & -4 & 9 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -3 & -2 & 4 \\ 0 & 1 & 0 & 4 & 3 & -6 \\ 0 & 0 & 1 & -6 & -4 & 9 \end{array}\right)$$

Az inverz mátrix tehát:

$$A^{-1} = \begin{pmatrix} -3 & -2 & 4 \\ 4 & 3 & -6 \\ -6 & -4 & 9 \end{pmatrix}$$

amit az $A^{-1}A$ mátrixszorzás elvégzésével könnyen ellenőrizhetünk.

4.7. Sajátérték, sajátvektor

Sajátérték, sajátvektor:

Azt mondjuk, hogy az $A \in \mathbf{M}_{n\times n}$ négyzetes mátrixnak a λ szám *sajátértéke*, az $s \in \mathbf{R}^n$, $s \neq \mathbf{0}$ vektor pedig a λ-hoz tartozó *sajátvektora*, ha $As = \lambda s$ (ezt az egyenletet *sajátértékegyenletnek* nevezzük).

A definícióban az $s \neq \mathbf{0}$ kitétel érthető: ha ui, s gyanánt a zérusvektort is megengedjük, akkor a definíció értelmét veszti, hiszen ekkor minden szám sajátérték lenne az $s = \mathbf{0}$ sajátvektorral.

A definíció nehéz, mert egyszerre két új fogalmat is bevezet. Később látni fogunk olyan tételeket, melyek segítségével a sajátértékek a sajátvektoroktól elkülönítve határozhatók meg.

A fenti fogalmak szemléletes jelentése a következő. Általában, ha $x \in \mathbf{R}^n$ egy tetszőleges vektor, az x és az Ax vektorok egymáshoz képesti irányáról nem lehet semmit állítani. Abban a kivételes helyzetben, amikor valamely nemzérus s vektorra s és As párhuzamosak (azaz egyik valamilyen számszorosa a másiknak), akkor az ilyen s vektort sajátvektornak, az arányossági tényezőt pedig sajátértéknek nevezzük.

Világos, hogy a sajátvektorok legfeljebb egy konstans szorzó erejéig lehetnek egyértelműek: ha ui. s egy sajátvektor λ sajátértékkel, akkor $As = \lambda s$, de ekkor tetszőleges α nemzérus számra $A(\alpha s) = \lambda\alpha s$, azaz αs is sajátvektor, ugyanazzal a λ sajátértékkel.

4.10. Példa: Az $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ mátrixnak a 3 sajátértéke, egy hozzátartozó sajátvektor pedig $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, mert

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

További példák:

4.11. Példa: A 0 zérusmátrixnak a 0 szám sajátértéke (és csak az): minden nemzérus vektor sajátvektor.

4.12. Példa: Az I egységmátrixnak az 1 szám sajátértéke (és csak az): minden nemzérus vektor sajátvektor.

4.13. Példa: Ha A diagonálmátrix, akkor A sajátértékei a főátlóban szereplő számok (és csak azok): a sajátvektorok a standard bázis elemei.

Egyáltalán nem nyilvánvaló, hogy minden mátrixnak létezik-e sajátértéke. Ha csak valós elemű mátrixokra és vektorokra szorítkozunk, mint ahogy eddig is tettük, ez nem is igaz, de *komplex elemű* mátrixok és vektorok esetén már igaz (ekkor a sajátértékek is általában komplex számok).

Látni fogjuk ugyanis, hogy a sajátértékek előállnak egy speciális n-edfokú egyenlet megoldásaként. E jegyzet keretein belül azonban továbbra is csak valós elemű mátrixokkal és vektorokkal foglalkozunk, és általában csak valós sajátértékek lesznek számunkra érdekesek.

A sajátértékekkel a regularitás egyszerűen megfogalmazható:

4.8. Állítás: Egy $A \in \mathbf{M}_{n\times n}$ négyzetes mátrix pontosan akkor szinguláris, ha a 0 sajátértéke.

Bizonyítás: A 0 ui. pontosan akkor sajátérték, ha valamely $s \neq \mathbf{0}$ vektorra $As = \mathbf{0}$ teljesül, azaz, ha a homogén egyenletnek van nemtriviális megoldása.

Most pedig bebizonyítunk egy alapvető jelentőségű tételt, melynek alapján a sajátértékek – elvben – meghatározhatók:

4.4. Tétel: Egy $A \in \mathbf{M}_{n\times n}$ négyzetes mátrix sajátértékei (és csak azok) megegyeznek az $\det(A - \lambda I)$ n-edfokú polinom (az ún. *karakterisztikus polinom*) gyökeivel.

Bizonyítás: Az $As = \lambda s$ sajátértékegyenlet ui. azzal ekvivalens, hogy az $(A - \lambda I)s = \mathbf{0}$ homogén egyenletnek létezik nemtriviális megoldása (ti. az s sajátvektor), ami pedig pontosan akkor teljesül, ha a rendszer $(A - \lambda I)$ mátrixa szinguláris, azaz determinánsa 0.

A $\det(A - \lambda I) = 0$ egyenletet *karakterisztikus egyenletnek* is nevezzük.

A determináns definíciójából nyomban adódik, hogy $\det(A - \lambda I)$ a λ-nak valóban egy n-edfokú polinomja: az algebra alaptétele miatt ennek mindig létezik (általában komplex) gyöke. Következésképp bármely négyzetes mátrixnak van (általában komplex) sajátértéke: a sajátértékek általában akkor is komplex számok, ha a mátrix elemei mind valósak. Jegyezzük meg azonban, hogy ebben a speciális esetben (tehát valós elemű mátrixok esetén), *minden sajátérték esetén annak komplex konjugáltja is sajátérték*. Valóban, ha a karakterisztikus egyenlet a következő alakú:

$$a_0 + a_1\lambda + a_2\lambda^2 + ... + a_n\lambda^n = 0,$$

és ennek λ egy gyöke, akkor az egyenlőség mindkét oldalának komplex konjugáltját véve adódik, hogy λ-val együtt $\bar{\lambda}$ is gyök, mert

$$a_0 + a_1\bar{\lambda} + a_2\bar{\lambda}^2 + ... + a_n\bar{\lambda}^n = 0,$$

ahol kihasználtuk, hogy az a_j együtthatók valósak, így komplex konjugáltjuk önmagukkal egyezik. Ennek az észrevételnek egy érdekes következménye, hogy *minden páratlan rendű valós elemű mátrixnak van (legalább egy) valós sajátértéke*. Ellenkező esetben ui. minden sajátértékkel együtt annak konjugáltja is sajátérték lenne, azaz a sajátértékek száma páros lenne, ám a sajátértékek száma a mátrix rendjével egyezik, ami páratlan.

Ha a sajátértékeket már meghatároztuk, a sajátvektorok kiszámítása egyszerű: ekkor ui. már csak az $(A - \lambda I)s = \mathbf{0}$ homogén egyenletnek kell a nemtriviális megoldásait megkeresni.

4.14. Példa: Határozzuk meg az $A := \begin{pmatrix} -3 & 2 & 1 \\ 1 & -3 & 2 \\ 1 & 2 & -3 \end{pmatrix}$ mátrix sajátértékeit és sajátvektorait.

Megoldás: A karakterisztikus polinom:

$$\det(A - \lambda I) = \det \begin{pmatrix} -3-\lambda & 2 & 1 \\ 1 & -3-\lambda & 2 \\ 1 & 2 & -3-\lambda \end{pmatrix} =$$

$$= (-3-\lambda)\left[(3+\lambda)^2 - 4\right] - 2 \cdot (-3-\lambda-2) + (2+3+\lambda) =$$

$$= -\lambda^3 - 9\lambda^2 - 20\lambda.$$

Ennek gyökei a sajátértékek: 0, -4 és -5.

A $\lambda = 0$ sajátértékhez tartozó sajátvektor az alábbi egyenlet nemtriviális megoldása:

$$(A - 0 \cdot I)s = \begin{pmatrix} -3 & 2 & 1 \\ 1 & -3 & 2 \\ 1 & 2 & -3 \end{pmatrix} s = \mathbf{0},$$

ahonnan a sajátvektor (konstans szorzó erejéig egyértelműen):

$$s = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Hasonlóan, a $\lambda = -4$ sajátértékhez tartozó sajátvektor az alábbi egyenlet nemtriviális megoldása:

$$(A + 4 \cdot I)s = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} s = \mathbf{0},$$

ahonnan a sajátvektor (konstans szorzó erejéig egyértelműen):

$$s = \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix}$$

Végül, a $\lambda = -5$ sajátértékhez tartozó sajátvektor az alábbi egyenlet nemtriviális megoldása:

$$(A + 5 \cdot I)s = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix} s = \mathbf{0},$$

ahonnan a sajátvektor (konstans szorzó erejéig egyértelműen):

$$s = \begin{pmatrix} 2 \\ -3 \\ 2 \end{pmatrix}$$

Speciális eset: Másodrendű mátrixok esetén a karakterisztikus polinom másodfokú, így a gyökök meghatározása áttekinthető. Legyen $A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, akkor a karakterisztikus egyenlet:

$$\det(A - \lambda I) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} =$$

$$= \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0,$$

azaz

$$\lambda^2 - \mathrm{sp}(A)\lambda + \det(A) = 0.$$

ahol $\mathrm{sp}(A)$ jelöli a mátrix *nyomát*, azaz a főátlóbeli elemek összegét. Innen a gyökök és az együtthatók közti ismert összefüggés alapján a λ_1 és λ_2 sajátértékekre fennállnak az alábbi egyenlőségek:

$$\lambda_1 + \lambda_2 = \mathrm{sp}(A)$$

$$\lambda_1\lambda_2 = \det(A).$$

A sajátértékek és sajátvektorok rendszere a lineáris algebra egyik legfontosabb fogalomköre. Számos alkalmazása közül röviden vázolunk egyet, mely a numerikus matematikával kapcsolatos. A gyakorlatban az egyenletrendszerek számítógépes megoldása során mindig elkövetünk numerikus hibákat, mivel a számok ábrázolása értelemszerűen csak véges sok értékes jeggyel történik. A hibák egy másik forrása, hogy egy-egy egyenletrendszer együtthatói és/vagy jobb oldalának elemei maguk is csak pontatlanul ismertek (pl. azért, mert valamilyen mérés eredményei). Természetesen az a kívánatos, hogy ha csak "kicsit" hibázunk az adatokban és/vagy az egyenletmegoldás során, akkor ez csak "kicsi" hibát okozzon a megoldásban. Sajnos ez nem mindig van így. A következőkben példát mutatunk olyan (igen egyszerű és kisméretű) ún. *gyengén meghatározott* vagy *rosszul kondícionált* egyenletrendszerre, ahol az adatok kis hibája óriási hibát okoz a megoldásban. Egészen természetes igény, hogy ezt a jelenséget még az egyenletmegoldás előtt fel tudjuk ismerni és valamilyen "mérőszám" alapján jellemezni tudjuk. Ez a mérőszám lesz a *kondíciószám*, mely bizonyos mátrixok esetén a sajátértékekből származtatható.

Gyengén meghatározott egyenletrendszerek: Tekintsük a következő modellfeladatot:

$$\begin{aligned} 1000x + 999y &= 1 \\ 999x + 998y &= 1 \end{aligned}$$

Megoldása könnyen ellenőrizhetően: $x = 1$, $y = -1$, és ez az egyetlen megoldás, mivel a rendszer determinánsa nem 0 (ellenőrizzük!).

Tekintsük most ugyanezt az egyenletrendszert, de egy kicsit megváltoztatott jobb oldallal:

$$\begin{aligned} 1000x + 999y &= 1 \\ 999x + 998y &= 0.999 \end{aligned}$$

melynek egyetlen megoldása: $x = 0.001$, $y = 0$, ami egészen távol áll az előző megoldástól. A jelenség a *gyengén meghatározott* egyenletek tipikus esete. Kiderült (a részletekkel itt nem foglalkozhatunk), hogy az A mátrix kondícionáltságának mérőszámaként jól használható a

$$\mathrm{cond}(A) := \frac{|\lambda|_{\max}}{|\lambda|_{\min}}$$

ún. *kondíciószám*, legalábbis akkor ha az A mátrix önadjungált (azaz szimmetrikus a főátlójára, ld. a következő szakaszt). Itt $|\lambda|_{\max}$ ill. $|\lambda|_{\min}$ jelenti a mátrix legnagyobb ill. legkisebb abszolút értékű sajátértékének abszolút értékét. Definíció szerint $\mathrm{cond}(A) \geq 1$, és a kondíciószám nem változik, ha a mátrixot bármely nemzérus skalárral megszorozzuk, azaz érzéketlen a mátrixelemek "mértékegységére". Általános szabályként elmondható, hogy minél nagyobb a kondíciószám, annál érzékenyebb az egyenlet az adatok ill. a számítás hibájára. 10^5 körüli vagy ennél még nagyobb kondíciószám már nagyon gyengén meghatározott egyenletre utal. Könnyen ellenőrizhető, hogy a fenti modellfeladatban a mátrix sajátértékei $\lambda_1 \approx 2000$ és $\lambda_2 \approx -\frac{1}{2000}$, igy a kondíciószám $\mathrm{cond}(A) \approx 4 \cdot 10^6$.

Egy másik példa gyengén meghatározott (rosszul kondícionált) mátrixokra az alábbi mátrixsorozat:

$$H_n := [\frac{1}{k+j-1}] = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \dots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \dots & \frac{1}{n+2} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \dots & \frac{1}{2n-1} \end{pmatrix} \in \mathbf{M}_{n\times n}$$

H_n-t *n-edrendű Hilbert-mátrixnak* nevezzük. A Hilbert-mátrixok természetes módon felbukkannak bizonyos függvényközelítési problémákban (ld. a 4.7.szakaszt). H_n kondíciószáma a mátrix rendjének növelésével nagyon gyorsan nő: $\mathrm{cond}(H_6) \approx 1.5 \cdot 10^7$, de már $\mathrm{cond}(H_{10}) \approx 1.6 \cdot 10^{13}$. Nagyon rossz kondícionáltságuk miatt a Hilbert-mátrixok jól alkalmazhatók pl. különféle egyenletmegoldó algoritmusok tesztelésére.

4.8. Önadjungált mátrixok

Mátrix adjungáltja:

Az $A \in \mathbf{M}_{n\times m}$, $A = [a_{kj}]$ mátrix *transzponáltjának* (vagy *adjungáltjának*) a sorok és oszlopok felcserélésével nyert $A^* := [a_{jk}] \in \mathbf{M}_{m\times n}$ mátrixot nevezzük.

Önadjungált mátrix:

Az $A \in \mathbf{M}_{n\times n}$ négyzetes mátrix *szimmetrikus* (vagy *önadjungált*), ha $A^* = A$.

Másszóval, az A mátrix önadjungált, ha elemeit a főátlóra tükrözve, a mátrix nem változik, azaz $a_{kj} = a_{jk}$ teljesül minden $k,j = 1,2,\dots,n$-re. Így pl. az $n \times n$-es zérusmátrix és az egységmátrix önadjungáltak.

A szóhasználat sajnos nem egységes. Néha egy mátrix adjungáltjának azt a mátrixot nevezik, melynek kj-edik eleme a k-adik sor és j-edik oszlop elhagyásával kapott $(n-1)$-edrendű mátrix determinánsa. Ezt a fajta adjungáltat azonban nem A^*-gal, hanem az $\mathrm{adj}(A)$ szimbólummal szokás jelölni. E jegyzet keretein belül az adjungált mátrix mindig a fenti definícióval meghatározott mátrixot, tehát a sorok és az oszlopok felcserélésével kapott mátrixot fogja jelenteni.

Komplex elemű mátrixok esetén az adjungált mátrix nem a sorok és az oszlopok felcserélésével kapott mátrixot (tehát a transzponáltat), hanem a transzponált (elemenként képzett) komplex konjugáltját jelenti. Mivel e jegyzetben csak valós elemű mátrixokat vizsgálunk, a konjugálásnak nincs "hatása", így ezt fel sem tüntettük a definícióban.

Az adjungált mátrix és a skaláris szorzat közt érdekes kapcsolat áll fenn:

4.9. Állítás: Ha $A \in \mathbf{M}_{n\times n}$ tetszőleges mátrix, akkor minden $x,y \in \mathbf{R}^n$ esetén $\langle Ax,y\rangle = \langle x,A^*y\rangle$ teljesül. Következésképp, ha A önadjungált, akkor $\langle Ax,y\rangle = \langle x,Ay\rangle$ teljesül minden $x,y \in \mathbf{R}^n$ -re.

Bizonyítás: Valóban, a bal oldal: $\langle Ax,y\rangle = \sum_{k=1}^n (Ax)_k y_k = \sum_{k=1}^n \sum_{j=1}^n a_{kj}x_jy_k$, míg a jobb oldal: $\langle x,A^*y\rangle = \sum_{p=1}^n x_p(A^*y)_p = \sum_{p=1}^n \sum_{q=1}^n x_p(A^*)_{pq}y_q = \sum_{p=1}^n \sum_{q=1}^n a_{qp}x_py_q$.

Az alábbiakban összefoglaljuk az adjungálásra vonatkozó azonosságokat:

4.10. Állítás: Ha $A,B \in \mathbf{M}_{n\times n}$ tetszőleges mátrixok, akkor
(a) $(A+B)^* = A^* + B^*$
(b) $(c\cdot A)^* = c\cdot A^*$ minden $c \in \mathbf{R}$-re
(c) $(AB)^* = B^*A^*$
(d) $(A^*)^* = A$
(e) ha A reguláris, akkor A^* is az, és $(A^*)^{-1} = (A^{-1})^*$

Bizonyítás: Az állítások az utolsó kivételével közvetlenül adódik a definícióból: ezek ellenőrzését az Olvasóra bízzuk. Az (e) állítást így igazolhatjuk: a jobb oldali mátrixra teljesül, hogy $(A^{-1})^*A^* = (AA^{-1})^* = I^* = I$ (itt felhasználtuk a (c) állítást), tehát A^* inverze létezik, éspedig épp $(A^{-1})^*$-gal egyenlő.

Az önadjungált mátrixoknak számos fontos tulajdonságuk van, így pl.:

4.11. Állítás: Ha $A \in \mathbf{M}_{n\times n}$ önadjungált, akkor a különböző sajátértékekhez tartozó sajátvektorok ortogonálisak.

Bizonyítás: Legyenek $\lambda_1 \neq \lambda_2$, és $As_1 = \lambda_1 s_1$, $As_2 = \lambda_2 s_2$. Az első sajátértékegyenletet s_2-vel, a másodikat s_1-gyel szorozva skalárisan, kapjuk, hogy $\langle As_1,s_2\rangle = \lambda_1\langle s_1,s_2\rangle$, és $\langle As_2,s_1\rangle = \lambda_2\langle s_2,s_1\rangle$. A bal oldalak a **??**. Állítás miatt megegyeznek, innen: $(\lambda_1-\lambda_2)\cdot\langle s_1,s_2\rangle = 0$. Mivel pedig $\lambda_1 \neq \lambda_2$, azért innen szükségképp $\langle s_1,s_2\rangle = 0$, azaz s_1 és s_2 ortogonálisak.

Kvadratikus alak:

A $Q : \mathbf{R}^n \to \mathbf{R}$, $Q(x) := \langle Ax{,}x\rangle$ függvényt az $A \in \mathbf{M}_{n\times n}$ önadjungált mátrixhoz tartozó *kvadratikus alaknak* nevezzük.

A mátrixszorzás definíciója alapján könnyen ellenőrizhető, hogy a kvadratikus alak az A mátrix és az x vektor elemeivel a következő formában állítható elő:

$$\langle Ax{,}x\rangle = \sum_{k=1}^{n}\sum_{j=1}^{n} a_{kj}x_kx_j$$

Mátrix definitsége:

Az $A \in \mathbf{M}_{n\times n}$ önadjungált mátrix *pozitív szemidefinit*, ha a hozzátartozó kvadratikus alak csak nemnegatív értékeket vesz fel; *pozitív definit*, ha a kvadratikus alak pozitív minden $x \neq \mathbf{0}$ esetén. Hasonlóan, az A mátrix *negatív szemidefinit*, ha a hozzátartozó kvadratikus alakra $Q(x) \leq 0$ teljesül; *negatív definit*, ha $Q(x) < 0$ minden $x \neq \mathbf{0}$ esetén. Végül az A mátrixot *indefinitnek* nevezzük, ha a kvadratikus alak pozitív és negatív értékeket egyaránt felvesz.

Nyilván minden pozitív (negatív) definit mátrix egyúttal pozitív (negatív) szemidefinit is.

4.15. Példa: Az $A := \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ mátrix pozitív definit.

Bizonyítás: Legyen $A := \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbf{R}^2$ tetszőleges vektor, akkor

$$\langle \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\rangle = \langle \begin{pmatrix} 2x+y \\ y+2x \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\rangle =$$

$$= 2x^2 + yx + xy + 2y^2 = 2\cdot(x^2+xy+y^2) =$$

$$= 2\cdot((x+\frac{1}{2}y)^2 + \frac{3}{4}y^2),$$

ahonnan az állítás már következik.

4.16. Példa: A zérusmátrix pozitív szemidefinit, ugyanakkor negatív szemidefinit is. Az egységmátrix pozitív definit. Egy diagonálmátrix pozitív (negatív) definit, ha a diagonálelemei mind pozitívak (negatívak).

A definitség eldöntése számos alkalmazás (így pl. a többváltozós szélsőértékfeladatok) esetében alapvető fontosságú. Látni kell ugyanakkor, hogy a definíció alapján ezt általában nagyon nehéz megtenni. Ezen segít a következő tétel, mely szerint a definitség eldöntéséhez elég ismerni a sajátértékek előjelét:

4.5. Tétel: $A \in \mathbf{M}_{n\times n}$ önadjungált mátrix pontosan akkor

- pozitív (ill. negatív) szemidefinit, ha minden sajátértéke nemnegatív (ill. nempozitív);
- pozitív (ill. negatív) definit, ha minden sajátértéke pozitív (ill. negatív);
- indefinit, ha van pozitív és negatív sajátértéke is.

Csak a pozitív definitséget vizsgáljuk, a többi eset hasonlóan kezelhető. Legyen A pozitív definit. Tekintsük a sajátértékegyenletet: $As = \lambda s$. Mindkét oldalt szorozzuk meg skalárisan az s sajátvektorral: $0 < Q(s) = \langle As,s\rangle = \lambda \cdot \langle s,s\rangle = \lambda ||s||^2$, innen szükségképp $\lambda > 0$. A megfordítást (azt, hogy ha A minden sajátértéke pozitív, akkor A pozitív definit is) nem bizonyítjuk.

A definitség eldöntése sokszor még e tétel segítségével is nehéz, bár a tétel alkalmazásához nem kell ismerni magukat a sajátértékeket, csak azok előjelét. A következő tétel, melyet bizonyítás nélkül közlünk, lehetővé teszi a pozitív és negatív definitség megállapítását egészen gépies módon, egy sor determináns kiszámításával:

4.6. Tétel: Legyen $A = [a_{kj}] \in \mathbf{M}_{n\times n}$ önadjungált mátrix és jelölje A_k a mátrix bal felső sarkának elemeiből álló $k \times k$-as mátrixot ($k = 1,2,...,n$; az A_k mátrixokat az eredeti A mátrix *minormátrixainak* nevezzük):

$$A_k := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{pmatrix}$$

Akkor az A mátrix pontosan akkor

- pozitív definit, ha minden minormátrixának determinánsa pozitív, azaz $\det(A_k) > 0 \quad (k = 1,2,...,n)$;
- negatív definit, ha minormátrixainak determinánsai váltakozó előjelűek, pontosabban: $\operatorname{sign}(\det(A_k)) = (-1)^k \quad (k = 1,2,...,n)$, ahol "sign" az előjelfüggvényt jelöli.

A 2×2*-es mátrixok speciális esete*: Másodrendű önadjungált mátrixok esetén a definitség az előző tétel használata nélkül is könnyen eldönthető. Legyen $A = [a_{kj}] \in \mathbf{M}_{2\times 2}$ egy önadjungált mátrix: már láttuk, hogy ekkor fennállnak a

$$\lambda_1 + \lambda_2 = \mathrm{sp}(A), \quad \lambda_1\lambda_2 = \det(A).$$

összefüggések, ahol $\mathrm{sp}(A)$ jelöli az A mátrix *nyomát* (*spurját*), azaz a fődiagonálisban álló mátrixelemek összegét. Ezt az eredményt kombinálva a **??**. Tétellel, azonnal adódik, hogy:

4.2. Következmény: Az $A \in \mathbf{M}_{2\times 2}$ önadjungált mátrix pontosan akkor

- pozitív vagy negatív definit, ha $\det(A) > 0$;
- pozitív vagy negatív szemidefinit, ha $\det(A) \geq 0$;
- indefinit, ha $\det(A) < 0$;
- pozitív definit, ha $\det(A) > 0$, és $\mathrm{sp}(A) > 0$;
- negatív definit, ha $\det(A) > 0$, és $\mathrm{sp}(A) < 0$.

4.9. Néhány speciális mátrixosztály

Projektormátrix:

A $P \in \mathbf{M}_{n\times n}$ mátrixot *projektormátrixnak* vagy röviden projektornak nevezzük, ha $P^2 = P$.

Nyilvánvaló, hogy a zérusmátrix és az egységmátrix projektormátrixok (ezeket triviális projektoroknak nevezzük). Geometriai jelentéséből az is nyilvánvaló, hogy a síkra való vetítés mátrixa projektormátrix (ezt a definíció alapján is könnyű ellenőrizni). A projektorok egyszerűségük miatt játszanak fontos szerepet, ugyanakkor megmutatható, hogy egy sor mátrix (így pl. minden önadjungált mátrix) előállítható projektormátrixok lineáris kombinációjaként.

4.12. Állítás: Ha $P \in \mathbf{M}_{n\times n}$ projektor, akkor $I - P$ is az.

Bizonyítás: $(I - P)^2 = I^2 - IP - PI + P^2 = I - 2P + P = I - P$.

4.13. Állítás: Ha $P \in \mathbf{M}_{n\times n}$ projektor, akkor P sajátértékei csak a 0 és az 1 lehetnek.

Bizonyítás: Ha $Ps = \lambda s$, akkor $P^2 s = \lambda Ps = \lambda^2 s$. Ámde $Ps = P^2 s$, innen $\lambda(1-\lambda)s = \mathbf{0}$, ezért $s \neq \mathbf{0}$ miatt szükségképp $\lambda(1-\lambda) = 0$, ahonnan az állítás már következik.

Nemtriviális projektorra a legegyszerűbb példa a *diád*, azaz egyetlen $v = (v_1,v_2,\dots,v_n) \in \mathbf{R}^n$ egységnyi hosszúságú vektor önmagával vett diadikus szorzata:

$$P := \begin{pmatrix} v_1^2 & a_1v_2 & \dots & v_1v_n \\ v_2v_1 & v_2^2 & \dots & v_2v_n \\ \dots & \dots & \dots & \dots \\ v_nv_1 & v_nv_2 & \dots & v_n^2 \end{pmatrix}$$

A mátrixszorzás definíciója alapján könnyen látható, hogy minden $x \in \mathbf{R}^n$ esetén $Px = \langle x,v\rangle \cdot v$. Innen $P^2x = \langle x,v\rangle \cdot Pv = \langle x,v\rangle \cdot \langle v,v\rangle \cdot v = Px$ (mert $||v|| = 1$), tehát a fent P mátrix valóban projektor.

Legyen most $A \in \mathbf{M}_{n\times n}$ egy önadjungált mátrix, sajátértékei legyenek $\lambda_1,\lambda_2,\dots,\lambda_n$, a hozzátartozó sajátvektorok pedig $s_1,s_2,\dots,s_n$, melyekről feltehető, hogy egységnyi normájúak (ellenkező esetben osszuk el a sajátvektort a hosszával, így már egységnyi normájú sajátvektorhoz jutunk). Ha a sajátértékek mind különbözők, akkor már tudjuk, hogy a sajátvektorok egymásra páronként ortogonálisak, ezért bázist alkotnak $\mathbf{R}^n$-ben. Bizonyítás nélkül megjegyezzük, hogy ez *minden* önadjungált mátrix esetében igaz marad: $\mathbf{R}^n$-ben létezik A sajátvektoraiból álló ortonormált bázis, akkor is, ha a sajátértékek között vannak egyenlők (ez esetben e sajátértékekhez több, egymásra páronként ortogonális sajátvektor is létezik). Alapvető jelentőségű tétel, hogy ekkor A előáll a fenti típusú projektorok lineáris kombinációjaként:

4.7. Tétel: Ha $A \in \mathbf{M}_{n\times n}$ önadjungált, $\lambda_1,\lambda_2, \dots , \lambda_n$ sajátértékekkel és $s_1,s_2,\dots,s_n$ ortonormált sajátvektorrendszerrel, akkor tetszőleges $x \in \mathbf{R}^n$ esetén érvényes az alábbi előállítás:

$$Ax = \sum_{k=1}^{n} \lambda_k \langle x,s_k\rangle s_k$$

Bizonyítás: Valóban, már láttuk, hogy az $s_1,s_2,\dots,s_n$ ortonormált bázisban érvényes az $x = \sum_{k=1}^{n} \langle x,s_k\rangle s_k$ egyenlőség minden $x \in \mathbf{R}^n$ vektor mellett (**??**. Tétel). Az egyenlőség mindkét oldalát balról az A mátrixszal szorozva, a kívánt előállításhoz jutunk.

A fenti előállítást az A önadjungált mátrix *spektrálelőállításának* nevezzük. A jobb oldali összeg minden egyes tagjában $\langle x,s_k\rangle s_k$ nem más, mint egy projektor (az s_k sajátvektornak önmagával vett diadikus szorzata) alkalmazva az x vektorra. Tehát A előáll diádok lineáris kombinációjaként. Egyúttal azt is látjuk, hogy az önadjungált mátrixok egyértelműen meg vannak határozva a sajátértékeik és sajátvektoraik rendszerével: ezek ismeretében a mátrix rekonstruálható.

A spektrálelőállítás jelentőségét az adja, hogy ezzel az A mátrix bármely hatványa igen egyszerűen előállítható. Az egyenlőség mindkét oldalát A-val ismételten szorozva nyomban adódik, hogy tetszőleges m természetes számra:

$$A^m x = \sum_{k=1}^{n} \lambda_k^m \langle x,s_k\rangle s_k,$$

tehát $A^m x$ lényegében ugyanannyi művelet árán számítható ki mint Ax. Megjegyezzük, hogy ugyanakkor a mátrixszorzások tényleges (definíció szerinti) elvégzése igen műveletigényes: könnyen látható, hogy már két n-edrendű mátrix összeszorzása is kb. n^3 algebrai műveletet igényel.

Ezek alapján lehetséges egy (önadjungált) mátrixnak más, nem-hatvány alakú, de hatványsorral előállítható függvényét definiálni. Így pl. bevezethető a mátrixok exponenciális függvénye az

$$e^A x := \sum_{k=1}^{n} e^{\lambda_k} \langle x,s_k\rangle s_k,$$

és az így definiált mátrix-exponenciálisra a megszokott $(e^A)^k = e^{kA}$ teljesül. Ennek a konstrukciónak pl. a differenciálegyenletrendszerek vizsgálatában van nagy szerepe.

Még nagyobb jelentősége van annak, hogy a spektrálelőállítás ismeretében a mátrix *inverze* is ugyanilyen könnyedén határozható meg:

$$A^{-1} x = \sum_{k=1}^{n} \frac{1}{\lambda_k} \langle x,s_k\rangle s_k,$$

Szorozzuk meg ui. balról a jobb oldalt A-val: eredményül épp x-et kapunk (**??**. Tétel alapján), következésképp a jobb oldal valóban $A^{-1}x$-szel egyenlő.

Következésképp a sajátértékek és sajátvektorok ismeretében az önadjungált mátrixú

$$Ax = b$$

egyenlet megoldása explicit módon kifejezhető:

$$x = A^{-1}b = \sum_{k=1}^{n} \frac{1}{\lambda_k} \langle b, s_k \rangle s_k,$$

éspedig a Gauss-eliminációnál sokkal kevesebb (kb. $2n^2$) művelet árán. Mindehhez azonban, hangsúlyozzuk, az A mátrix sajátrendszerének ismerete szükséges, ennek meghatározása pedig – általában – még az inverz meghatározásánál is nehezebb probléma. Ezért e módszer csak speciális mátrixok esetén alkalmazható, amikor a sajátrendszer valamilyen egyéb meggondolások alapján már rendelkezésre áll.

Nilpotens mátrix:

Definíció: Az $A \in \mathbf{M}_{n\times n}$ mátrixot *nilpotensnek* nevezzük, ha valamely pozitív hatványa a zérusmátrixszalegyenlő (ekkor persze minden, ennél nagyobb kitevőjű hatványa is zérus).

4.17. Példa: Az

$$A := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbf{M}_{3\times 3}$$

mátrix nilpotens, $A^3 = \mathbf{0}$. Általában, ha $A \in \mathbf{M}_{n\times n}$ olyan, hogy a főátlója feletti átló elemei 1-gyel, az összes többi elem 0-val egyenlő, akkor A nilpotens mátrix: hatványozáskor az 1-esek átlója mintegy "felfelé csúszik", és az n-edik hatványnál már teljesen "kicsúszik" a mátrixból: $A^n = \mathbf{0}$.

A nilpotens mátrixok jelentőségét az adja, hogy lehetséges ilyen mátrixok *végtelen* hatványsorával dolgozni és ismert, sorokra vonatkozó eredményeket alkalmazni, miközben a mátrix nilpotens volta miatt a végtelen mátrix-hatványsor valójában egy *véges* összeg. Ilyen típusú eredmények közül az egyik legegyszerűbb az alábbi; két másikat pedig feladatként fogalmazunk meg a fejezet végén.

4.14. Állítás: Ha az $A \in \mathbf{M}_{n\times n}$ mátrix nilpotens, akkor az $I - A$ mátrix invertálható, és:

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k,$$

(ahol a jobb oldali formálisan végtelen sor csak véges sok tag összege, mivel alkalmas kitevőtől kezdve az összes mátrixhatvány a zérusmátrixszal egyenlő).

Bizonyítás: Legyen m egy olyan kitevő, melyre $A^m = \mathbf{0}$. Akkor a jobb oldal az $I + A + A^2 + ... + A^{m-1}$ véges összeggel egyenlő. Ezt jobbról $(I - A)$-val szorozva: $(I+A+A^2+...+A^{m-1})(I-A) = I+A+A^2+...+A^{m-1}-A-A^2-...-A^{m-1}-A^m = I - A^m = I$, tehát definíció szerint $(I - A)$ invertálható, és inverze $I + A + A^2 + ... + A^{m-1}$.

Az állítás a valós analízisből ismert, az $|a| < 1$ számokra érvényes $\frac{1}{1-a} = \sum_{k=0}^{\infty} a^k$ egyenlőség pontos megfelelője. Az $|a| < 1$ feltételt az előző állításban az A mátrix nilpotens voltának feltétele helyettesíti: megjegyezzük azonban, hogy az állítás ennél tágabb mátrixosztályokra is igaz marad (ennek részleteivel nem foglalkozunk).

Ortogonális mátrix:

Definíció: Az $A \in \mathbf{M}_{n\times n}$ mátrixot *ortogonális mátrixnak* nevezzük, ha oszlopai mint (oszlop)vektorok, ortonormált rendszert alkotnak $\mathbf{R}^n$-ben.

4.18. Példa: Az Olvasó könnyen ellenőrizheti, hogy $\mathbf{M}_{2\times 2}$ forgatómátrixai, azaz az

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

alakú mátrixok minden $t \in \mathbf{R}$ esetén ortogonális mátrixok.

Az ortogonális mátrixok különleges szerepe – amint azt a forgatómátrixok példáján már láttuk – abban áll, hogy inverzük rendkívül egyszerűen határozható meg:

4.15. Állítás: Ha az $A \in \mathbf{M}_{n\times n}$ mátrix ortogonális, akkor invertálható, és $A^{-1} = A^*$.

Bizonyítás: Az A^*A szorzatmátrix kj-edik eleme az A^* mátrix k-adik sorárak (tehát az A mátrix k-adik oszlopának) és az A mátrix j-edik oszlopának skaláris szorzata. Használva az oszlopvektorok ortonormáltságát, innen az adódik, hogy A^*A épp az I egységmátrixszal egyenlő, tehát A valóban invertálható, és $A^{-1} = A^*$.

4.10. Feladatok

1. Keressünk olyan (valós elemű) $J \in \mathbf{M}_{2\times 2}$ mátrixot, melyre $J^2 = -I$.

2. Igazoljuk, hogy ha $A,B \in \mathbf{M}_{n\times n}$ felcserélhető reguláris mátrixok, akkor

(a) A^{-1} és B is felcserélhetők;

(b) A és B^{-1} is felcserélhetők;

(c) A^{-1} és B^{-1} is felcserélhetők.

3. Igazoljuk, hogy ha $A,B \in \mathbf{M}_{n\times n}$ tetszőleges négyzetes mátrixok, akkor $\mathrm{sp}(AB) = \mathrm{sp}(BA)$, ahol $\mathrm{sp}(M)$ jelöli az M mátrix fődiagonálisában álló elemek összegét, azaz a mátrix *nyomát*.

4. Az előző feladat eredményének felhasználásával igazoljuk, hogy az

$$AB - BA = I$$

egyenlőség semmilyen négyzetes A,B mátrixokra nem teljesül.

5. Igazoljuk 2×2-es mátrixokra, hogy egy mátrix pontosan akkor szinguláris, ha a determinánsa zérus.

6. Regulárisak-e az alábbi mátrixok?

(a) $A := \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 2 \end{pmatrix}$

(b) $A := \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 3 \\ 1 & -3 & -1 \end{pmatrix}$

7. Oldjuk meg Gauss-eliminációval az alábbi egyenletrendszert:

$$\begin{array}{rcrcrcr} x & + & 4y & + & 2z & = & 5 \\ -3x & + & 2y & + & z & = & -1 \\ 4x & - & y & - & z & = & 2 \end{array}$$

8. Számítsuk ki Gauss-eliminációval az alábbi mátrix inverzét:

$$A := \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ -1 & 3 & 2 \end{pmatrix}$$

9. Hogyan alakul a

$$\begin{array}{ccccccc} tx & + & y & + & z & = & 1 \\ x & + & ty & + & z & = & t \\ x & + & y & + & tz & = & t^2 \end{array}$$

egyenletrendszer megoldhatósága a t valós paraméter különböző értékei mellett?

10. Számítsuk ki az alábbi mátrix sajátértékeit és sajátvektorait:

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

11. Mutassuk meg, hogy ha $t \neq k\pi$ $(k \in \mathbf{Z})$, akkor az $A_t := \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$ forgatómátrixnak nincs valós sajátértéke.

12. Igazoljuk, hogy ha λ sajátértéke az $A \in \mathbf{M}_{n\times n}$ mátrixnak, akkor λ^k sajátértéke A^k-nak ($k \in \mathbf{N}$), továbbá, ha A még reguláris is, akkor λ^{-1} sajátértéke A^{-1}-nek.

13. Mutassuk meg, hogy a vektoriális szorzás mátrixának a 0 mindig sajátértéke. Mi a hozzátartozó sajátvektor?

14. Definitség szempontjából osztályozzuk az $A := \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ és a $B := \begin{pmatrix} -3 & 8 \\ 8 & 3 \end{pmatrix}$ önadjungált mátrixokat.

15. Igazoljuk, hogy ha A,$B \in \mathbf{M}_{n\times n}$ önadjungáltak és felcserélhetők, akkor AB is önadjungált.

16. Mutassuk meg, hogy ha $A \in \mathbf{M}_{n\times n}$ nilpotens mátrix, akkor minden sajátértéke szükségképp 0.

17. Mutassuk meg, hogy ha $A \in \mathbf{M}_{n\times n}$ nilpotens mátrix, és $|c| < 1$, akkor $(cI + A)^m \to 0$ elemenként, ha $m \to +\infty$.

18. Egy négyzetes A mátrixot antiszimmetrikusnak nevezünk, ha $A^* = -A$. Igazoljuk, hogy ha A reguláris és antiszimmetrikus, akkor nem lehet valós sajátértéke!

19. Legyen $A \in \mathbf{M}_{n\times n}$ olyan mátrix, hogy minden sajátértékének abszolút értéke 1-nél kisebb. Igazoljuk, hogy ekkor $I - A$ reguláris.

20. Van-e olyan 3×3-as valós elemű mátrix, melynek sajátértékei az i, $2i$ és a $3i$ imaginárius számok?

Megoldások

1. Ilyen mátrix a következő: $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

A $J^2 = -I$ egyenlőség a komplex $i^2 = -1$ egyenlőség pontos megfelelője. Ily módon a komplex számok egyértelműen reprezentálhatók 2×2-es mátrixokkal. Feleltessük meg a $z = a + ib$ komplex számnak a $Z := aI + bJ = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ mátrixot, akkor ez a megfeleltetés összeg- és szorzattartó.

2.

(a) Jelölje $C := A^{-1}B$, akkor $CA = A^{-1}BA = A^{-1}AB = B$, innen $C = BA^{-1}$.

(b) Jelölje $C := AB^{-1}$, akkor $BC = BAB^{-1} = ABB^{-1} = A$, innen $C = B^{-1}A$.

(c) $A^{-1}B^{-1} = (BA)^{-1} = (AB)^{-1} = B^{-1}A^{-1}$

3. $\mathrm{sp}(AB) = \sum_{k=1}^n (AB)_{kk} = \sum_{k=1}^n \sum_{j=1}^n A_{kj}B_{jk}$, és $\mathrm{sp}(BA) = \sum_{p=1}^n (BA)_{pp} = \sum_{p=1}^n \sum_{q=1}^n B_{pq}A_{qp} = \mathrm{sp}(AB)$.

4. Ugyanis $\mathrm{sp}(AB - BA) = \mathrm{sp}(AB) - \mathrm{sp}(BA) = 0$, míg $\mathrm{sp}(I) = n \neq 0$.

5. Legyen $A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, és vizsgáljuk az $Ax = 0$ homogén egyenletet:

$$\begin{array}{rcrcl} a_{11}x_1 & + & a_{12}x_2 & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & = & 0 \end{array}$$

Az első egyenletet a_{21}-gyel, a másodikat a_{11}-gyel szorozva, és az így nyert egyenleteket kivonva, kapjuk, hogy:

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = 0$$

Haonlóan, az első egyenletet a_{22}-vel, a másodikat a_{12}-vel szorozva, és az így nyert egyenleteket kivonva, kapjuk, hogy:

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = 0$$

Ha tehát a determináns nem 0, akkor csak a triviális $x_1 = x_2 = 0$ megoldás létezik. Ha viszont a determináns 0, akkor van nemtriviális megoldás is.

6. (a) reguláris (mert $\det(A) = 1 \cdot 1 \cdot 2 \neq 0$)

(b) szinguláris (mert $\det(A) = 1\cdot(-1+9)-(-1)\cdot(-1-3)+1\cdot(-3-1) = 0$).

A (b) esetben a mátrix szingularitása úgy is belátható, hogy az $Ax = \mathbf{0}$ homogén egyenletnek van nemtriviális megoldása, pl. $(2,1,-1)$.

7. A számítás sémája a következő:

$$\left(\begin{array}{rrr|r} 1 & 4 & 2 & 5 \\ -3 & 2 & 1 & -1 \\ 4 & -1 & -1 & 2 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 4 & 2 & 5 \\ 0 & 14 & 7 & 14 \\ 0 & -17 & -9 & -18 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & 4 & 2 & 5 \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & -17 & -9 & -18 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 4 & 2 & 5 \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & -\frac{1}{2} & -1 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & 4 & 2 & 5 \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & 2 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 4 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

A megoldás tehát: $x = 1$, $y = 0$, $z = 2$.

8. A számítás sémája a következő:

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ -1 & 3 & 2 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ 0 & 3 & 3 & 1 & 0 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & \frac{1}{3} & 0 & \frac{1}{3} \end{array}\right)$$

Cseréljük meg a 2. és 3. egyenletet (a hozzájuk tartozó jobboldalakkal együtt, hogy az eliminációt folytatni tudjuk:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 2 & 0 & 1 & 0 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 & \frac{1}{2} & 0 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{1}{3} & -\frac{1}{2} & \frac{1}{3} \\ 0 & 0 & 1 & 0 & \frac{1}{2} & 0 \end{array}\right)$$

Az inverz mátrix tehát:

$$A^{-1} = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ \frac{1}{3} & -\frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

9. Ha a rendszer determinánsa nem 0, akkor a rendszer egyértelműen megoldható. Vizsgáljuk meg tehát, hogy milyen t értékek mellett zérus a determináns.

A determináns kifejtése:

$$t(t^2-1)-(t-1)+(1-t)=(t-1)(t^2+t-2)$$

Ezért a determináns pontosan akkor zérus, ha $t=1$ vagy
$t^2+t-2=0$, azaz $t=1$ vagy $t=-2$.

A $t=1$ esetben az egyenlet:

$$\begin{array}{ccccccc} x & + & y & + & z & = & 1 \\ x & + & y & + & z & = & 1 \\ x & + & y & + & z & = & 1 \end{array}$$

ennek pedig végtelen sok megoldása van: két ismeretlen, pl. y és z szabadon megválasztható, ezekután $x=1-y-z$.

A $t=-2$ esetben az egyenlet:

$$\begin{array}{rcrcrcr} -2x & + & y & + & z & = & 1 \\ x & - & 2y & + & z & = & -2 \\ x & + & y & - & 2z & = & 4 \end{array}$$

Összeadva az egyenleteket, a bal oldalak összege 0, a jobboldalaké 1, ami nem lehetséges. Ekkor tehát az egyenletrendszernek nincs megoldása.

10. A karakterisztikus polinom:

$$\det(A-\lambda I) = \det\begin{pmatrix} -2-\lambda & 1 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & 1 & -2-\lambda \end{pmatrix} =$$

$$= (-2-\lambda)[(2+\lambda)^2-1] - (-2-\lambda-1) + (1+2+\lambda) =$$

$$= -(\lambda+2)^3 + 2 + \lambda + 2 + \lambda + 1 + 1 + 2 + \lambda =$$

$$= -\lambda^3 - 6\lambda^2 - 9\lambda = -\lambda(\lambda+3)^2$$

A sajátértékek a karakterisztikus polinom gyökei, tehát a 0 és a -3 számok.

A $\lambda = 0$ sajátértékhez tartozó sajátvektor a

$$(A - 0\cdot I)s = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} s = \mathbf{0}$$

homogén egyenlet nemtriviális megoldása. Ilyen megoldás (konstans szorzótól eltekintve) egyetlenegy van, az $s = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ vektor.

A $\lambda = -3$ sajátértékhez tartozó sajátvektor a

$$(A + 3\cdot I)s = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} s = \mathbf{0}$$

homogén egyenlet nemtriviális megoldása. Itt most két ismeretlen is megválasztató szabadon, így két, lineárisan független megoldás is létezik, pl. az $s = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, és az $s = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ vektor.

11. A karakterisztikus egyenlet:

$$\lambda^2 - \mathrm{sp}(A_t)\lambda + \det(A_t) = \lambda^2 - 2\lambda\cos t + 1 = 0,$$

innen $\lambda = \cos t \pm \sqrt{\cos^2 t - 1}$. Valós megoldás tehát csak akkor van, ha $|\cos t| = 1$, azaz, ha $t = k\pi$ alakú, ahol k egész szám.

12. Ha $As = \lambda s$, akkor $A^2s = \lambda As = \lambda^2 s$, $A^3s = \lambda A^2s = \lambda^3 s$, és így tovább, $A^k s = \lambda A^{k-1}s = \lambda^k s$. Ha pedig A még reguláris is, akkor $As = \lambda s$-ból $s = \lambda A^{-1}s$, azaz $A^{-1}s = \lambda^{-1}s$ következik.

13. Jelölje $A \in \mathbf{M}_{3\times 3}$ az $\mathbf{x} \to \mathbf{x}\times\mathbf{a}$ leképezés mátrixát, akkor $A\mathbf{a} = \mathbf{a}\times\mathbf{a} = \mathbf{0}$. Tehát 0 sajátérték, a hozzűtartozó sajátvektor pedig épp az $\mathbf{a}$ vektor.

14. Az A mátrix pozitív definit, mert determinánsa és nyoma egyaránt pozitív. A B mátrix indefinit, mert determinánsa negatív.

15. Ekkor $(AB)^* = B^*A^* = BA = AB$, azaz AB valóban önadjungált.

16. Legyen $N \in \mathbf{N}$ akkora, hogy $A^N = 0$. Ha most λ sajátértéke A-nak, a hozzátartozó sajátvektor pedig s, akkor $As = \lambda s$: innen $A^2s = \lambda As = \lambda^2 s$, $A^3s = \lambda A^2 s = \lambda^3 s$, és így tovább, $A^N s = \lambda A^{N-1}s = \lambda^N s$. Mivel pedig $A^N = 0$, azért $\lambda^N s = 0$, innen $s \neq \mathbf{0}$ miatt szükségképp $\lambda = 0$.

17. Legyen $N \in \mathbf{N}$ akkora, hogy $A^N = 0$. Mivel A és I nyilván felcserélhető, a binomiális tétel alkalmazható az $(cI + A)^m$ mátrixhatványra, és:

$$(cI + A)^m = \sum_{k=0}^{m} \binom{m}{k} A^k c^{m-k} I^{m-k} = \sum_{k=0}^{N-1} \binom{m}{k} c^{m-k} A^k$$

A jobb oldalon most már a tagok száma *m-től független*, és minden tag elemenként 0-hoz tart, ha $m \to +\infty$ (ui. $\binom{m}{k}$ végtelenbe tart ugyan, ha $m \to +\infty$, de csak polinomiális sebességgel (lévén m-nek egy N-nél alacsonyabb fokú polinomja), míg $|c| < 1$ miatt c^{m-k} exponenciális sebességgel tart 0-hoz, így kettőjük szorzata még mindig 0-hoz tart.

18. Tegyük fel, hogy λ egy valós sajátértéke A-nak. Legyen $As = \lambda s$ ($s \neq \mathbf{0}$). Szorozzuk meg az egyenlőség mindkét oldalát skalárisan s-sel: $\langle As, s\rangle = \lambda ||s||^2$. Másrészt $\langle As, s\rangle = \langle s, A^*s\rangle = -\langle s, As\rangle = -\lambda||s||^2$. Innen $s \neq \mathbf{0}$ miatt szükségképp $\lambda = 0$, azaz a mátrix nem lehet reguláris.

19. Ekkor ui. $I - A$ sajátértékei $1 - \lambda$ alakúak, ahol λ sajátértéke A-nak (valóban, ha $As = \lambda s$, akkor $(I - A)s = s - \lambda s = (1 - \lambda)s$, és

megfordítva, ha $(I - A)s = (1 - \lambda)s$, akkor ebből $As = \lambda s$ következik). Mivel pedig $|\lambda| < 1$, azért $1 - \lambda \neq 0$, tehát $I - A$ valóban reguláris.

20. Nincs. A karakterisztikus polinom ui. egy pontosan 3-fokú polinom, melynek határértéke a $-\infty$-ben $-\infty$, a $+\infty$-ben pedig $+\infty$ (ha a harmadfokú tag előjele pozitív) ill. a $-\infty$-ben $+\infty$, a $+\infty$-ben pedig $-\infty$ (ha a harmadfokú tag előjele negatív). Mindkét esetben legalább egy valós gyöke van (Bolzano tétele miatt), tehát nem lehet mindhárom sajátérték tiszta képzetes.

Megjegyezzük, hogy ez az eredmény onnan is adódik, hogy a karakterisztikus polinom *valós együtthatós*, így ismeretes, hogy gyökei valósak vagy konjugált komplex gyökpárokat alkotnak. A feladatban pedig nem ilyen gyökök vannak adva.

5. Többváltozós függvények

Ebben a fejezetben $\mathbf{R}^n \to \mathbf{R}$ típusú függvények analízisével foglalkozunk. Látni fogjuk, hogy már a folytonosság és különösen a differenciálhatóság fogalma lényegesen bonyolultabb a megfelelő egyváltozós fogalmaknál. A szélsőérték problémák esetében az analízisbeli és a lineáris algebrai eszközök egy természetes összekapcsolódását figyelhetjük meg. Végül a Riemann-integrál fogalmát általánosítjuk többváltozós függvények esetére.

5.1. Többváltozós függvények bevezetése

Legyen $f : \mathbf{R}^n \to \mathbf{R}$ egy leképezés, jelölje $\Omega \subset \mathbf{R}^n$ az értelmezési tartományát. f-et egyaránt tekinthetjük olyan leképezésnek, mely *több változóhoz* rendel egy számot, vagy olyannak, amely *vektorhoz* rendel számot, ahogy az épp kényelmesebb. Egy-egy konkrét leképezés megadása általában explicit formulákkal történik (pl. $f : \mathbf{R}^2 \to \mathbf{R}$, $f(x,y) := e^{x+y}$), ritkábban implicit módon. Ez utóbbi általános alakja:

$$g(x_1,x_2,...,x_n,y) = 0,$$

ahol g egy $(n+1)$-változós kifejezés. Amennyiben ebből az egyenlőségből az utolsó változó (y) egyértelműen kifejezhető az első n változó függvényében, akkor a fenti egyenletet ezen n-változós függvény implicit alakjának nevezzük.

Hasonlóan az egyváltozós függvényekhez, az $\{(x_1,x_2,...,x_n,f(x_1,...,x_n) : (x_1,x_2,...,x_n) \in \Omega\} \subset \mathbf{R}^{n+1}$ halmazt az f függvény *grafikonjának* nevezzük. Sajnos, ennek csak $n = 1$ és $n = 2$ esetben van szemléletes jelentése: előbbi esetben a grafikon általában egy síkgörbe, míg a másodikban egy térbeli felület.

Kétváltozós függvények esetében egy másik szemléltetési forma a *szintvonalas ábrázolás*, ahol az értelmezési tartomány azon pontjait tüntetjük fel, melyekben a függvényértékek egy-egy előre meghatározott értékkel egyenlők (ezt a technikát a térképeken mindmáig előszeretettel használják). Ez olykor háromváltozós függvények esetében is alkalmazható (itt a függvényt a *szintfelületek* rendszere szemlélteti). Számítógépes grafika segítségével oly módon is lehet szemléltetni a függvényeket, hogy az értelmezési tartomány és/vagy a grafikon pontjait az ott érvényes függvényértéktől függően más-más színűre színezzük. Ilyen ábrázolási módokat pl. a MAPLE vagy a MATLAB matematikai szoftverek messzemenően támogatnak.

5.1. Példa: A következő formula implicit módon egy olyan kétváltozós leképezést definiál, melynek grafikonja egy origó középpontú, egységnyi sugarú félgömb-felület:

$$x^2 + y^2 + z^2 - 1 = 0 \quad (|x|,|y| \le 1,\, z \ge 0)$$

A leképezés explicit alakja: $z = f(x,y) = \sqrt{1 - x^2 - y^2}$, szintvonalai pedig az xy-síkon origó közepű koncentrikus körök.

5.2. Folytonosság

A folytonosságot az egyváltozós függvények folytonosságának mintájára definiáljuk. Ehhez előbb azonban be kell vezetni a *vektorsorozatok* konvergenciájának fogalmát:

Vektorsorozat konvergenciája:

Azt mondjuk, hogy az $(x_m) \subset \mathbf{R}^n$, $x_m := (x_m^{(1)},x_m^{(2)},\ldots,x_m^{(n)})$ vektorsorozat *konvergens*, éspedig az $x := (x^{(1)},x^{(2)},\ldots,x^{(n)}) \in \mathbf{R}^n$ vektorhoz tart, ha mindegyik $(x_m^{(j)})$ komponens-sorozat tart az x limeszvektor megfelelő $x^{(j)}$ komponenséhez, azaz $x_m^{(j)} \to x^{(j)}$ ($j = 1,2,\ldots,n$, $m \to +\infty$).

Nyilvánvaló, hogy $x_m \to x$ pontosan akkor teljesül, ha az $||x_m - x||$ *számsorozat* 0-hoz tart.

Ezekután a folytonosság már nehézség nélkül definiálható. A definíció lényege itt is – akárcsak egyváltozós esetben – az, hogy a függvény "konvergenciatartó" legyen:

Folytonosság:

Azt mondjuk, hogy az $f : \Omega \to \mathbf{R}$ függvény *folytonos az $x \in \Omega$ helyen*, ha minden $(x_m) \subset \Omega$, $x_m \to x$ vektorsorozat esetén $f(x_m) \to f(x)$.

Sajnos, a jelöléstechnika e téren sem egységes. Alsó indexszel pl. hol az $\mathbf{R}^n$-beli vektorok komponenseit jelöljük, hol egy vektorsorozat indexét (ekkor a komponenseket jobb híján zárójelben tett felső indexekkel jelöljük). Elterjedt ezért – legalábbis 2- és 3-változós függvények esetében –, hogy a vektorkomponenseket különböző *betűkkel* jelöljük (pl. x, y, z), ekkor az index a félreértés veszélye nélkül sorozat-elemet jelölhet.

A Lipschitz-féle feltétel teljesülése most is elegendő a folytonossághoz:

5.1. Állítás: Ha van oly $C \geq 0$ szám, hogy minden $x_1, x_2 \in \Omega$ esetén

$$|f(x_1) - f(x_2)| \leq C \cdot ||x_1 - x_2||$$

teljesül, akkor f folytonos Ω minden pontjában.

Bizonyítás: Ekkor ui. minden $x \in \Omega$ és $x_m \to x$ vektorsorozat esetén $|f(x) - f(x_m)| \leq C \cdot ||x - x_m|| \to 0$, azaz $f(x_m) \to f(x)$.

Könnyen látható, hogy a szokásos függvényműveletek a folytonosságot megtartják, azaz folytonos függvényekből a szokásos függvényműveletekkel nyert függvények maguk is folytonosak (ahol egyáltalán értelmezettek). Sőt, folytonos egyváltozós függvényekből "összerakott" függvények szintén folytonosak:

5.2. Állítás: Legyenek $f_1, f_2, ..., f_n : \mathbf{R} \to \mathbf{R}$ folytonos (egyváltozós) függvények. Akkor az

$$f(x_1, x_2, ..., x_n) := f_1(x_1) + f_2(x_2) + ... + f_n(x_n)$$

és az

$$f(x_1, x_2, ..., x_n) := f_1(x_1) \cdot f_2(x_2) \cdot ... \cdot f_n(x_n)$$

előírással értelmezett $f : \mathbf{R}^n \to \mathbf{R}$ (n-változós) függvények szintén folytonosak. Ha pedig $f : \mathbf{R}^n \to \mathbf{R}$ folytonos n-változós függvény, $g : \mathbf{R} \to \mathbf{R}$ pedig folytonos egyváltozós függvény, akkor az összetett $g \circ f$ függvény is folytonos (n-változós) függvény.

Következésképp folytonos függvényekből "képlettel" nem is lehet mást, mint folytonos függvényt értelmezni.

A definíció egyszerűsége ellenére a folytonosság ill. szakadás megállapítása már kétváltozós esetben sem mindig magától értetődő. Jellemzően az okozhat nehézséget, ha egy függvényt egyes helyeken eltérő módon definiálunk, tipikusan ott, ahol egy formula értelmét veszti (pl. 0-val való osztás vagy negatív számból való gyökvonás stb. miatt). A jelenséget két példán keresztül érzékeltetjük:

5.2. Példa: Tekintsük az alábbi előírással értelmezett kétváltozós függvényt:

$$f(x,y) := \begin{cases} \frac{x^2y^2}{x^2+y^2} & \text{ha } x^2+y^2 \neq 0 \\ 0 & \text{ha } x=y=0 \end{cases}$$

Az origón kívül a függvény nyilván mindenütt folytonos, a kérdés az, hogy az origóban folytonos-e. Megmutatjuk, hogy f folytonos az origóban. Legyenek ui. $x_n,y_n \to 0$ tetszőleges zérussorozatok, akkor:

$$|f(x_n,y_n)| = \frac{x_n^2y_n^2}{x_n^2+y_n^2} \leq \frac{x_n^2(x_n^2+y_n^2)}{x_n^2+y_n^2} = x_n^2 \to 0,$$

tehát $f(x_n,y_n) \to 0 = f(0,0)$, azaz f valóban folytonos az origóban.

5.3. Példa: Tekintsük most az alábbi függvényt, melyet majdnem ugyanúgy definiálunk, mint az előbb:

$$f(x,y) := \begin{cases} \frac{xy}{x^2+y^2} & \text{ha } x^2+y^2 \neq 0 \\ 0 & \text{ha } x=y=0 \end{cases}$$

Az origón kívül a függvény most is mindenütt folytonos, azonban az origóban nem folytonos. Hogy ezt megmutassuk, elég egyetlen olyan $x_n,y_n \to 0$ zérussorozat-párt találni, melyre $f(x_n,y_n)$ *nem* tart 0-hoz. Legyen pl. $x_n := y_n := \frac{1}{n}$, akkor nyilván $x_n,y_n \to 0$, de

$$f(x_n,y_n) = \frac{\frac{1}{n}\cdot\frac{1}{n}}{\frac{1}{n^2}+\frac{1}{n^2}} \equiv \frac{1}{2}.$$

Ezért f valóban nem folytonos az origóban. Látható tehát, hogy a folytonosság nem mindig állapítható meg "ránézésre". Szemléltetésként javasoljuk az Olvasónak, hogy a fenti két függvény grafikonját állítsa elő pl. a MAPLE segítségével, és figyelje meg az origó körüli viselkedésüket.

5.3. Többváltozós függvények differenciálhatósága

Az egyváltozós $f : \mathbf{R} \to \mathbf{R}$ függvény egy $x \in \mathbf{R}$ helyen vett $a := f'(x)$ deriváltjának szokásos definíciója:

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = a$$

általánosításra alkalmatlan, hiszen többváltozós f függvény esetén h vektor, és a vele való osztás értelmetlen. Ám ez a definíció könnyen láthatóan ekvivalens az alábbival:

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - a\cdot h}{|h|} = 0,$$

és ezt már könnyű általánosítani. Ennek a definíciónak a szemléletes jelentése a következő: a függvényérték $f(x+h) - f(x)$ változása kis h értékek mellett jól közelíthető az $a \cdot h$ szorzattal, abban az értelemben, hogy kettőjük különbsége nemcsak, hogy 0-hoz tart, de még $|h|$-val osztva is a 0-hoz tart, mikor $h \to 0$. Röviden: $f(x+h) - f(x)$ közelítően arányos h-val, és a derivált épp az arányossági tényező.

Ez a megfogalmazás már értelemszerűen általánosítható többváltozós függvények esetére:

Differenciálhatóság, gradiens:

Legyen $f : \Omega \to \mathbf{R}$ egy n-változós függvény, és legyen $x \in \Omega$ az értelmezési tartomány *belső pontja*, azaz tegyük fel, hogy nemcsak maga x, de még egy x középpontú, valamely $r > 0$ sugarú $\{y \in \mathbf{R}^n : ||x-y|| < r\}$ gömb pontjai is mind Ω-ba esnek. Azt mondjuk, hogy az f függvény *differenciálható az x helyen, és deriváltja az $a \in \mathbf{R}^n$ vektor*, ha

$$\lim_{h\to \mathbf{0}} \frac{f(x+h) - f(x) - \langle a,h\rangle}{||h||} = 0.$$

A fenti $a \in \mathbf{R}^n$ vektort az f függvény x helyen vett *gradiensvektorának* vagy röviden *gradiensének* nevezzük, és a grad $f(x)$ szimbólummal (néha $\bigtriangledown f(x)$-szel) jelöljük.

A deriváltra – mint az várható – az egyváltozós deriválthoz hasonló összefüggések igazak. Megjegyezzük azonban, hogy a definíció gyakorlati számításokra teljesen alkalmatlan: ezt a problémát majd a következő szakaszban, a parciális deriváltak fogalmának bevezetésével küszöböljük ki.

A következőkben összefoglaljuk a többváltozós derivált legfontosabb tulajdonságait. Az állítások a definícióból az egyváltozós esettel analóg módon következnek, így a bizonyításokat elhagyjuk.

5.3. Állítás: Ha $f,g:\Omega\to\mathbf{R}$ differenciálhatók az $x\in\Omega$ helyen, akkor $f+g$, fg és $\frac{f}{g}$ is differenciálhatók x-ben (utóbbinál feltéve, hogy $g(x)\neq 0$), és:

$$\mathrm{grad}(f+g)(x)=\mathrm{grad}\ f(x)+\mathrm{grad}\ g(x)$$

$$\mathrm{grad}(fg)(x)=(\mathrm{grad}\ f(x))\cdot g(x)+f(x)\cdot(\mathrm{grad}\ g(x))$$

$$\mathrm{grad}\frac{f}{g}(x)=\frac{(\mathrm{grad}\ f(x))\cdot g(x)-f(x)\cdot(\mathrm{grad}\ g(x))}{g^2(x)}$$

5.4. Parciális és iránymenti derivált, a második derivált mátrixa

Legyen $f:\Omega\to\mathbf{R}$ egy n-változós függvény, és legyen $x\in\Omega$ az értelmezési tartomány belső pontja. Legyen $e\in\mathbf{R}^n$ egy tetszőleges, 1 normájú (irány)vektor. A későbbiekben kiemelt szerepet fog játszani a következő egyváltozós függvény:

$$g_e(t):=f(x+t\cdot e)$$

Ezeknek a g_e függvényeknek $n=2$ esetén igen szemléletes jelentésük van. Képzeljünk el egy hegyes-völgyes terepet. Ez felfogható egy $f:\mathbf{R}^2\to\mathbf{R}$ függvény grafikonjaként: minden síkbeli (x,y) pont esetén $f(x,y)$ jelentse a terepmagasságot. Helyezzünk egy megfigyelőt a terep $(x,y,f(x,y))$ pontjára. Ekkor a g_e függvény grafikonja az ezen a ponton áthaladó, vízszintesen e irányú *útvonal*, azaz a g_e függvény értékei a terepmagasságok egy kitüntetett e irányban.

Ily módon *egyetlen többváltozós* függvény vizsgálatát visszavezetjük *sok* (méghozzá végtelen sok) *egyváltozós* függvény egyidejű vizsgálatára. Első eredményünk ezen g_e függvények deriváltjaira vonatkozik. A szóhasználat egyszerűsítése érdekében bevezetjük egy $x\in\mathbf{R}^n$ pont *környezetének* fogalmát: ezalatt egy tetszőleges olyan $\mathbf{R}^n$-beli halmazt értünk, mely tartalmaz egy x középpontú (nemnulla sugarú) gömböt. Ily módon $\mathbf{R}$-ben egy x szám környezete olyan számhalmaz, mely tartalmaz egy x középpontú, pozitív hosszúságú intervallumot.

5.4. Állítás: Ha az $f : \Omega \to \mathbf{R}$ differenciálható az $x \in \Omega$ belső pont egy környezetében, akkor az összes fenti g_e függvény differenciálható a 0 egy környezetében, éspedig:

$$g_e'(t) = \langle \text{grad } f(x + t \cdot e), e \rangle,$$

ahol $e \in \mathbf{R}^n$ tetszőleges, 1 normájú irányvektor.

Bizonyítás: Legyen $\tau \in \mathbf{R}$ tetszőleges, írjuk fel a g_e-re vonatkozó különbségi hányadost:

$$\frac{g(t+\tau) - g(t)}{\tau} = \frac{f(x + te + \tau e) - f(x + te)}{\tau} =$$

$$= \frac{f(x + te + \tau e) - f(x + te) - \langle \text{grad } f(x + t \cdot e), \tau e \rangle}{||\tau e||} \cdot \frac{|\tau|}{\tau} +$$

$$+ \langle \text{grad } f(x + t \cdot e), e \rangle$$

Jelölje $h := \tau e$, akkor innen:

$$\frac{g(t+\tau) - g(t)}{\tau} =$$

$$= \frac{f(x + te + h) - f(x + te) - \langle \text{grad } f(x + t \cdot e), h \rangle}{||h||} \cdot \frac{|\tau|}{\tau} +$$

$$+ \langle \text{grad } f(x + t \cdot e), e \rangle$$

Ha most $\tau \to 0$, akkor $h \to \mathbf{0}$; ekkor a jobb oldal első tagja az f függvény deriválhatósága miatt 0-hoz tart, míg a bal oldal $g_e'(t)$-hez. Innen az állítás adódik.

Az állítás az összetett függvény deriválási szabályát általánosítja. A fenti g_e függvény ui. f-nek és a $t \to x + te$ vektorértékű függvénynek a kompozíciója. Ez utóbbi függvény deriváltját értelemszerűen e-nek definiálva, az állítás pontos megfelelője az összetett függvény deriváltjára vonatkozó formulának (a szorzás szerepét a skaláris szorzat veszi át). A dolog tovább általánosítható: legyen $h := (h_1, \ldots, h_n) : \mathbf{R} \to \mathbf{R}^n$ egy *vektorfüggvény*, ahol a h_j komponensfüggvények differenciálható egyváltozós függvények. A h vektorfüggvény deriváltját $h' := (h_1', \ldots, h_n')$-nek definiálva, könnyen megmutatható, hogy az $f \circ h$ összetett függvény deriváltja egy t helyen a $\langle \text{grad } f(h(t)), h'(t) \rangle$ szám. Erre az általánosabb eredményre azonban e jegyzet keretein belül nem lesz szükségünk.

Elnevezés: A fenti g_e függvény 0-ban vett deriváltját az f függvény x-ben vett e irány menti deriváltjának nevezzük, és a $\frac{\partial f}{\partial e}(x)$ szimbólummal (ritkábban $\partial_e f(x)$-szel) jelöljük. A **??**. Állítás értelmében tehát:

$$\frac{\partial f}{\partial e}(x) = \langle \text{grad } f(x), e \rangle$$

Az állítás egyúttal a gradiensvektor komponenseinek kényelmes kiszámítására is alkalmas. Legyen $e := e_k$, (ahol $e_1,\dots,e_n$ jelöli $\mathbf{R}^n$ standard bázisát). Akkor g_e deriváltja egy, csak a k-adik x_k változóra vonatkozó különbségi hányados határértéke. Ugyanakkor $\langle \text{grad } f(x), e \rangle$ épp a gradiensvektor k-adik komponense. Kaptuk tehát, hogy *a gradiensvektor k-adik komponensét úgy számíthatjuk ki, hogy f-et mint egyváltozós függvényt deriváljuk, csak a k-adik változót tekintve "valódi" változónak, míg az összes többit konstansnak tekintjük.*

Elnevezés: Az e_k irány menti deriváltat az f függvény k-adik változója szerinti *parciális deriváltjának* nevezzük, és a $\frac{\partial f}{\partial x_k}(x)$ szimbólummal jelöljük. További használatos jelölések: $D_k f(x)$, $\partial_k f(x)$, $f'_{x_k}(x)$. Ha a változókat nem indexszel, hanem külön betűkkel jelöljük (tipikusan x, y, z), akkor a parciális deriváltakat értelemszerűen így jelöljük: $\frac{\partial f}{\partial x}$, $\partial_x f$, f'_x, és így tovább.

Ezeket a fogalmakat használva, azt kaptuk, hogy a gradiensvektor komponensei a parciális deriváltak (az áttekinthetőség kedvéért az argumentumokat elhagyva):

$$\text{grad } f = (D_1 f, D_2 f, \dots, D_n f)$$

5.4. Példa: Legyen $f(x,y,z) := x^2 \sin(5y + z^3)$, akkor

$$\frac{\partial f}{\partial x} = 2x \sin(5y + z^3)$$

$$\frac{\partial f}{\partial y} = 5x^2 \cos(5y + z^3)$$

$$\frac{\partial f}{\partial z} = 3z^2 x^2 \cos(5y + z^3)$$

5.5. Példa: Legyen $f : \mathbf{R}^n \to \mathbf{R}$, $f(x) := ||x||^2 = \sum_{j=1}^n x_j^2$. Akkor $k = 1,2,\ldots,n$-re

$$\frac{\partial f}{\partial x_k} = 2x_k,$$

innen pedig:

$$\text{grad } f(x) = 2x$$

5.6. Példa: Legyen $f : \mathbf{R}^n \to \mathbf{R}$, $f(x) := ||x|| = \sqrt{\sum_{j=1}^n x_j^2}$. Akkor $k = 1,2,\ldots,n$-re

$$\frac{\partial f}{\partial x_k} = \frac{2x_k}{2\sqrt{\sum_{j=1}^n x_j^2}},$$

azaz:

$$\text{grad } f(x) = \frac{x}{||x||}$$

5.7. Példa: Legyen $f : \mathbf{R}^n \to \mathbf{R}$, $f(x) := \langle Ax,x\rangle$, ahol $A = [a_{kj}] \in \mathbf{M}_{n\times n}$ egy önadjungált mátrix. Alkalmazzuk most a többváltozós függvények deriváltjának definícióját: tetszőleges $h \in \mathbf{R}^n$ esetén

$$f(x+h) - f(x) = \langle Ax + Ah, x + h\rangle - \langle Ax,x\rangle =$$

$$= \langle Ax,x\rangle + \langle Ax,h\rangle\langle Ah,x\rangle + \langle Ah,h\rangle - \langle Ax,x\rangle =$$

$$= 2\langle Ax,h\rangle + \langle Ah,h\rangle$$

Mivel pedig a Cauchy-egyenlőtlenség miatt $|\frac{\langle Ah,h\rangle|}{||h||} \leq \frac{||Ah||\cdot||h||}{||h||} = ||Ah|| \to 0$, ha $h \to \mathbf{0}$, azért innen a gradiensvektor már meghatározható, éspedig:

$$\text{grad } f(x) = 2Ax.$$

A gradiensvektornak – legalábbis $n = 2$ és $n = 3$ esetben – szemléletes jelentése van, de ez eltér az egyváltozós függvények deriváltjának jelentésétől:

5.5. Állítás: A gradiensvektor a függvény *legmeredekebb változásának irányába mutat*, azaz tetszőleges $e \in \mathbf{R}^n$ egységvektor esetén

$$\left|\frac{\partial f}{\partial e}(x)\right| \leq \left|\frac{\partial f}{\partial e^*}(x)\right|,$$

ahol e^* a grad $f(x)$ irányba mutató egységvektor (feltéve, hogy grad $f(x) \neq \mathbf{0}$).

Bizonyítás: Nyilván $e^* = \frac{\text{grad } f(x)}{||\text{grad } f(x)||}$, innen a jobb oldal:

$$\left|\frac{\partial f}{\partial e^*}(x)\right| = \frac{\langle \text{grad } f(x), \text{grad } f(x)\rangle}{||\text{grad } f(x)||} = ||\text{grad } f(x)||$$

A bal oldali kifejezés pedig:

$$\left|\frac{\partial f}{\partial e}(x)\right| = |\langle \text{grad } f(x), e\rangle|$$

Az állítás most már a Cauchy-egyenlőtlenség egyenes következménye.

Végül a másodrendű deriváltat általánosítjuk többváltozós függvényekre. Tegyük fel, hogy az $f : \Omega \to \mathbf{R}$ függvény az $x \in \Omega$ belső pontban *kétszer differenciálható*, azaz mindegyik parciális deriváltfüggvénye differenciálható. Képezzük a x_j szerinti parciális deriváltfüggvény x_k szerinti parciális deriváltját az x helyen ($j,k = 1,2,...,n$). Az így kapott számokat az f függvény x helyen vett *másodrendű parciális deriváltjainak* nevezzük, és a $\frac{\partial^2 f}{\partial x_k \partial x_j}(x)$ szimbólummal jelöljük. További használatos jelölések: $D_k D_j f(x)$, $D_{kj} f(x)$, $\partial_{kj} f(x)$, $f''_{x_k x_j}(x)$. Ha $j = k$, akkor a megfelelő, ún. *tiszta másodrendű parciális deriváltakat* még így is jelölik: $\frac{\partial^2 f}{\partial x_k^2}(x)$ vagy $D_k^2 f(x)$. (Megkülönböztetésül, $j \neq k$ esetén a $D_k D_j f(x)$ deriváltakat *vegyes másodrendű parciális deriváltaknak* is nevezik.)

Elnevezés: A másodrendű parciális deriváltakból összeállított

$$D^2 f(x) := \begin{pmatrix} D_{11} f(x) & D_{12} f(x) & \ldots & D_{1n} f(x) \\ D_{21} f(x) & D_{22} f(x) & \ldots & D_{2n} f(x) \\ \ldots & \ldots & \ldots & \ldots \\ D_{n1} f(x) & D_{n2} f(x) & \ldots & D_{nn} f(x) \end{pmatrix} \in \mathbf{M}_{n \times n}$$

mátrixot az f függvény x helyen vett *második derivált mátrixának* vagy *Hesse-féle mátrixának* nevezzük.

A Hesse-mátrix a legtöbb gyakorlati esetben *önadjungált* (szimmetrikus). Érvényes ui. az alábbi tétel, melyet bizonyítás nélkül közlünk:

5.1. Tétel: Ha az f függvény másodrendű parciális deriváltjai mind léteznek és folytonosak valamely $x \in \Omega$ pontban, akkor ott

$$D_{kj}f(x) = D_{jk}f(x),$$

azaz a parciális deriválások sorrendje felcserélhető.

5.8. Példa: Verifikáljuk a fenti tételt a következő konkrét esetben. Legyen $f(x,y) := x^6 + x^2 \sin y^4 + e^{2y}$. Akkor:

$$\frac{\partial f}{\partial x} = 6x^5 + 2x \sin y^4, \qquad \frac{\partial f}{\partial y} = 4x^2y^3 \cos y^4 + 2e^{2y}$$

innen pedig:

$$\frac{\partial^2 f}{\partial y \partial x} = 8xy^3 \cos y^4, \ \text{ és } \ \frac{\partial^2 f}{\partial x \partial y} = 8xy^3 \cos y^4.$$

5.5. Többváltozós függvények lokális szélsőértékei

Lokális szélsőérték:

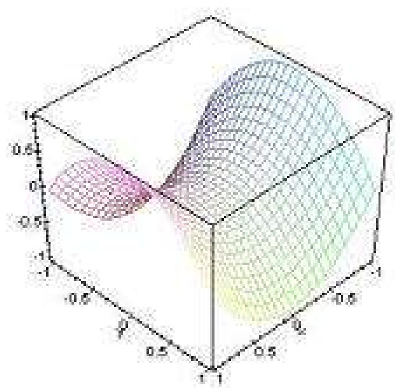
Azt mondjuk, hogy az $f : \Omega \to \mathbf{R}$ függvénynek *lokális maximuma (ill. minimuma) van az $x_0 \in \Omega$ pontban*, ha x_0-nak van olyan Ω_0 környezete, hogy $f(x) \leq f(x_0)$ (ill. $f(x) \geq f(x_0)$) teljesül minden $x \in \Omega_0$ esetén.

A következő tétel pontos analogonja az egyváltozós függvényekre vonatkozó megfelelő tételnek:

5.2. Tétel: Ha az f függvény differenciálható az Ω értelmezési tartomány egy x_0 belső pontjában, és ott lokális szélsőértéke (maximuma vagy minimuma) van, akkor x_0-ban mindegyik parciális derivált eltűnik, azaz grad $f(x_0) = \mathbf{0}$.

Bizonyítás: Tekintsük $\mathbf{R}^n$-ben az e_k standard bázisvektort, és jelölje $g_k(t) := f(x_0 + t \cdot e_k)$. Könnyen látható, hogy g_k-nak a 0-ban szintén szélsőértéke van (éspedig ugyanolyan típusú, mind f-nek x_0-ban). A **??**. Állítás értelmében ezért $g_k'(0) = \langle \text{grad } f(x_0), e_k \rangle = D_k f(x_0) = 0$. Ez igaz minden $k = 1,2,\ldots,n$-re, és ezzel az állítást igazoltuk.

A tétel megfordítása nem igaz! Abból, hogy egy függvénynek valahol mindegyik parciális deriváltja eltűnik, általában nem következik, hogy ott a függvénynek lokális szélsőértéke lenne. Ellenpéldaként tekintsük az $f(x,y) := x^2 - y^2$ előírással értelmezett függvényt. Az origóban a függvényérték és mindkét parciális derivált zérussal egyenlő, a függvénynek itt mégsincs szélsőértéke: minden $x \neq 0$, $y = 0$ koordinátájú helyen f pozitív, míg minden $x = 0$, $y \neq 0$ helyen negatív értéket vesz fel. Szemléletesen: a függvénynek az origóban az $(1,0)$ irány szerint lokális minimuma, míg a $(0,1)$ irány szerint lokális maximuma van. Az ilyen tulajdonságú helyet *nyeregpontnak* nevezzük. Az elnevezést a fenti függvény grafikonjának nyeregfelületre emlékeztető alakja indokolja (ld. az ábrát). Most megmutatjuk,



5.1. ábra. *Példa nyeregfelületre és nyeregpontra*

hogy – az egyváltozós esethez hasonlóan – ha az elsőrendű parciális deriváltak eltűnésén kívül a második deriváltra további feltételek teljesülnek, akkor ez már biztosítja a szélsőértékhely létezését, sőt, a szélsőérték jellege is eldönthető. Figyeljük meg, hogy míg egyváltozós esetben ehhez elég volt a második derivált *előjelét* vizsgálni, addig többváltozós esetben a második derivált mátrix *definitségének* vizsgálata szükséges.

5.3. Tétel: Ha az f függvény kétszer folytonosan differenciálható az Ω értelmezési tartomány egy x_0 belső pontjában, grad $f(x_0) = \mathbf{0}$, továbbá a $D^2 f(x_0)$ Hesse-mátrix definit, akkor f-nek az x_0 helyen biztosan lokális szélsőértéke van, éspedig lokális minimuma, ha $D^2 f(x_0)$ pozitív definit, ill. lokális maximuma, ha $D^2 f(x_0)$ negatív definit. f-nek nincs szélsőértéke (mégpedig nyeregpontja van), ha $D^2 f(x_0)$ indefinit.

Bizonyítás: Legyen $e = (e_1, e_2, \ldots, e_n) \in \mathbf{R}^n$ teszőleges egységvektor. Vezessük be most is a $g_e(t) := f(x_0 + t \cdot e_k)$ előírással értelmezett függvényt. A **??**. Állítás következtében

$$g_e'(t) = \langle \text{grad } f(x_0 + t \cdot e), e \rangle = \sum_{j=1}^{n} D_j f(x_0 + t \cdot e) \cdot e_j,$$

ezért $g_e'(0) = 0$.

Még egyszer deriválva g_e-t, a jobb oldali szumma minden tagjában újra alkalmazhatjuk a **??**. Állítást (most már a parciális deriváltfüggvényekre), innen a $t = 0$ helyen:

$$g_e''(0) = \sum_{j=1}^{n} \langle \text{grad } D_j f(x_0), e \rangle = \sum_{k=1}^{n} \sum_{j=1}^{n} D_{kj} f(x_0) \cdot e_j \cdot e_k$$

Ámde a jobb oldalon, mint az könnyen ellenőrizhető, épp a Hesse-mátrix kvadratikus alakja áll (az e vektorra alkalmazva), azaz

$$g_e''(0) = \langle D^2 f(x_0) e, e \rangle$$

Ezekután a tétel állítása már egyszerűen látható. Már tudjuk, hogy $g_e'(0) = 0$. Ha most $D^2 f(x_0)$ pozitív definit, akkor $g_e''(0) > 0$, így az egyváltozós g_e függvénynek a 0-ban lokális minimuma van. *Ez igaz minden e irányra*, így f-nek x_0-ban valóban lokális minimuma van. A lokális minimum esete ugyanígy látható be. Végül, ha $D^2 f(x_0)$ indefinit, akkor a kvadratikus alak pozitív és negatív értéket egyaránt felvesz, tehát van olyan irány mely szerint x_0-ban f-nek lokális minimuma, és olyan irány is, mely szerint lokális maximuma van, azaz x_0 nyeregpontja f-nek: ekkor tehát nincs lokális szélsőérték.

Speciális eset: Kétváltozós függvények esetében a Hesse-mátrix 2×2-es mátrix, melynek definitségét az előző fejezet eredményei alapján nagyon könnyű eldönteni. Innen nyerjük a kétváltozós függvényekre vonatkozó alábbi tételt:

5.1. Következmény: Ha a kétváltozós f függvény kétszer folytonosan differenciálható az Ω értelmezési tartomány egy (x_0,y_0) belső pontjában, $\frac{\partial f}{\partial x}(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0) = 0$, továbbá a $D^2 f(x_0)$ Hesse-mátrix determinánsa pozitív, akkor f-nek az (x_0,y_0) helyen biztosan lokális szélsőértéke van, éspedig lokális minimuma, ha még $\mathrm{sp}(D^2 f(x_0,y_0)) > 0$, ill. lokális maximuma, ha még $\mathrm{sp}(D^2 f(x_0,y_0)) < 0$ is teljesül. f-nek nincs szélsőértéke (mégpedig nyeregpontja van), ha a $D^2 f(x_0)$ Hesse-mátrix determinánsa negatív.

A fenti tételeket néhány példával illusztráljuk:

5.9. Példa: Keressük meg az $f(x,y) := 2x^2 - xy - 3y^2 + 5x - 1$ előírással értelmezett kétváltozós függvény lokális szélsőértékeit.

Megoldás: Az elsőrendű parciális deriváltak:

$$\frac{\partial f}{\partial x} = 4x - y + 5, \qquad \frac{\partial f}{\partial y} = -x - 6y.$$

Innen kapjuk, hogy a Hesse-mátrix konstans mátrix:

$$D^2 f(x,y) = \begin{pmatrix} 4 & -1 \\ -1 & -6 \end{pmatrix}$$

Lokális szélsőérték ott lehet, ahol mindkét parciális derivált eltűnik: azonban e hely(ek)et fel sem érdemes térképezni, mert a Hesse-mátrix mindenütt indefinit (mert determinánsa negatív), így lokális szélsőérték sehol sincs, a függvénynek nyeregpontja van ott, ahol az elsőrendű parciális deriváltak eltűnnek.

5.10. Példa: Legyenek $(x_1,y_1),(x_2,y_2),...,(x_N,y_N) \in \mathbf{R}^2$ adott síkbeli pontok. Keressük azt az $(x,y) \in \mathbf{R}^2$ pontot, melyre a

$$\sum_{j=1}^{N} ((x - x_j)^2 + (y - y_j)^2)$$

távolság-négyzetösszeg a lehető legkisebb.

Megoldás: Jelölje $f(x,y) := \sum_{j=1}^{N} ((x - x_j)^2 + (y - y_j)^2)$, meg kell keresni az így definiált kétváltozós f függvény minimumhelyét. Minimumhely ott lehet, ahol mindkét parciális derivált eltűnik, azaz

$$\frac{\partial f}{\partial x} = 2\sum_{j=1}^{N} (x - x_j) = 0, \qquad \frac{\partial f}{\partial y} = 2\sum_{j=1}^{N} (y - y_j) = 0,$$

ahonnan $x = \frac{1}{N}\sum_{j=1}^{N} x_j$, $y = \frac{1}{N}\sum_{j=1}^{N} y_j$, tehát (x,y) a pontrendszer *súlypontja*. Itt pedig f-nek valóban lokális minimuma van, mert a Hesse-mátrix

$$D^2 f(x,y) = \begin{pmatrix} 2N & 0 \\ 0 & 2N \end{pmatrix} = 2N \cdot I$$

nyilván pozitív definit.

A fenti példa nemcsak az $\mathbf{R}^2$ síkon definiálható, hanem $\mathbf{R}^n$-ben is (tetszőleges $n \in \mathbf{N}$-re): a minimumhelyet mindig az adott vektorrendszer súlypontja adja.
A példa a *távolságnégyzetek* összegének minimalizálására vonatkozik. A *távolságösszegek* minimalizálása egészen más – és sokkal nehezebb – feladat!

5.11. Példa: Egy kerti (téglatest alakú) medencét szeretnénk építeni. A medence belső oldalait ki kell csempézni. Hogyan válasszuk meg a medence méreteit, hogy a térfogata 32 m^3, a csempézés költsége (ami a kicsempézendő felülettel egyenesen arányosnak vehető) pedig a lehető legkisebb legyen?

Megoldás: Jelölje x, y a medence alapjának méreteit, z pedig a mélységét. Akkor kicsempézendő felület nagysága:

$$F = xy + 2xz + 2yz$$

Azonban x, y és z nem függetlenek, a térfogat előírt: $V = xyz$, ahonnan az egyik hosszúságadat pl. z kifejezhető a másik kettő függvényében: $z = \frac{V}{xy}$. Ezt beírva a felületet megadó formulába, F már csak x-től és y-tól függ:

$$F(x,y) = xy + \frac{2V}{y} + \frac{2V}{x}$$

Ezzel a kifejezéssel értelmezett függvényt kell minimalizálni. Szélsőérték ott lehet, ahol mindkét parciális derivált eltűnik, azaz:

$$\frac{\partial F}{\partial x} = y - \frac{2V}{x^2} = 0$$
$$\frac{\partial F}{\partial y} = x - \frac{2V}{y^2} = 0$$

Ennek az x, y-ra nézve nemlineáris egyenletnek megoldására nyilván teljesül, hogy $x^2y = xy^2 = V$, ahonnan egyetlen megoldás adódik, éspedig $x = y = \sqrt[3]{2V}$. Számítsuk ki a Hesse-mátrixot:

$$D^2 F(x,y) = \begin{pmatrix} \frac{4V}{x^3} & 1 \\ 1 & \frac{4V}{y^3} \end{pmatrix}$$

Ennek determinánsa a fenti helyen: $\frac{16V^2}{x^3y^3} - 1 = \frac{16V^2}{4V^2} - 1 = 3$, azaz pozitív. Nyoma nyilván pozitív, így a fenti helyen valóban lokális minimum van. A konkrét adatokkal: $x = y = 4$ m, ahonnan a medence mélysége: $z = 2$ m.

5.12. Példa: Szimmetrikus trapéz keresztmetszetű egyenes csatornát építünk. Belső oldalait (mindenütt azonos vastagsággal) ki kell betonozni. Hogyan válasszuk meg a trapéz méreteit, hogy a csatorna keresztmetszete előírt T nagyságú, a betonozás folyóméterenkénti költsége pedig a lehető legkisebb legyen?

Megoldás: Jelölje a a csatorna szélességét a fenéken (azaz a trapéz rövidebb párhuzamos oldalát), b a csatorna oldalfalának hosszát (azaz a trapéz szárait). Jelölje továbbá m a csatorna mélységét (a trapéz magasságát), α pedig az oldalfal lejtési szögét (a trapéz szára és hosszabbik párhuzamos oldala által bezárt szöget). A folyóméterenkénti betonozási költség nyilván arányos az $(a + 2b)$ hossz-összeggel, ezt kell tehát minimalizálni. Ám a és b egymástól nem függetlenek, a trapéz T területe előírt. Célszerű áttérni az m és az α változókra, azaz minden adatot ezek függvényében kifejezni. Elemi trigonometriai összefüggések alapján:

$$m = b \sin\alpha, \qquad T = \frac{2a + 2b\cos\alpha}{2} \cdot m = (a + b\cos\alpha) \cdot m$$

Az első egyenlőségből b azonal kifejezhető: ezt beírva a második egyenlőségbe, ezekután onnan a is kifejezhető m és α segítségével:

$$b = \frac{m}{\sin\alpha}, \qquad a = \frac{T}{m} - m \cdot \operatorname{ctg}\alpha$$

Így a minimalizálandó $(a + 2b)$ mennyiség is kifejezhető m és α függvényében:

$$f(m,\alpha) := \frac{T}{m} - m \cdot \operatorname{ctg}\alpha + \frac{2m}{\sin\alpha}$$

Szélsőérték ott lehet, ahol az elsőrendű parciális deriváltak eltűnnek, azaz ahol

$$\frac{\partial f}{\partial\alpha} = \frac{m}{\sin^2\alpha} - \frac{2m\cos\alpha}{\sin^2\alpha} = m \cdot \frac{1 - 2\cos\alpha}{\sin^2\alpha} = 0$$

$$\frac{\partial f}{\partial m} = -\frac{T}{m^2} - \operatorname{ctg}\alpha + \frac{2}{\sin\alpha} = 0$$

Az első egyenletből azonnal kapjuk, hogy $\cos\alpha = \frac{1}{2}$, azaz $\alpha = 60°$. Ezt beírva a második egyenletbe, m is számítható: $m = \sqrt{\frac{T}{\sqrt{3}}}$.

A másodrendű deriváltakat kiszámítva a fenti α, m értékek mellett, kapjuk, hogy:

$$\frac{\partial^2 f}{\partial\alpha} = m \cdot \frac{2\sin^3\alpha - (1 - 2\cos\alpha) \cdot 2\sin\alpha\cos\alpha}{\sin^4\alpha} > 0$$

$$\frac{\partial^2 f}{\partial m} = \frac{2T}{m^3} > 0$$
$$\frac{\partial^2 f}{\partial\alpha\partial m} = \frac{1-2\cos\alpha}{\sin^2\alpha} = 0$$

Ezért a Hesse-mátrix pozitív definit (mert determinánsa és nyoma is pozitív), így f-nek a fenti helyen valóban lokális minimuma van.

A minimumhoz tartozó a, b értékek a, b kifejezéseiből most már minden további nélkül adódnak. A részletszámítások levezetését az Olvasóra bízzuk. Az eredmény: $a = b = \frac{2}{3}\sqrt{T\sqrt{3}}$.

5.6. Feltételes szélsőérték feladatok

Az előző szakasz két utolsó példáján is megfigyelhető, hogy a gyakorlati szélsőérték feladatokban eléggé tipikus, hogy a szóba jöhető "változók" nem függetlenek, közöttük valamilyen kényszerkapcsolat előírt: a **??**. Példa esetében a medence térfogata, a **??**. Példa esetében pedig a csatorna keresztmetszete. Ilyenkor az egyik lehetséges megoldási technika az, amit az előző szakaszban alkalmaztunk: a kényszerfeltételi egyenlőségből kifejezzük az egyik változót, és ezt beírjuk az illető változó összes előfordulási helyére. Egy másik – bizonyos értelemben sokkal természetesebb – megoldási módszer a problémát ún. *feltételes szélsőérték feladatként* kezelni, melyet az alábbiakban vázolunk.

Feltételes szélsőérték:

Legyenek $f,g : \mathbf{R}^n \to \mathbf{R}$ adott függvények. Azt mondjuk, hogy az f függvénynek az $x_0 \in \mathbf{R}^n$ helyen *lokális feltételes maximuma (ill. minimuma) van a $g(x) = 0$ feltétel mellett*, ha $g(x_0) = 0$, és x_0-nak van oly Ω_0 környezete, hogy minden olyan $x \in \Omega_0$ esetén, melyre $g(x) = 0$, teljesül, hogy

$$f(x) \le f(x_0) \qquad (\text{ill. } f(x) \ge f(x_0)).$$

Így pl. a **??**. Példában a feltétel az, hogy a medence térfogata 32 m^3 legyen, azaz most $g(x,y,z) = xyz - 32$.

A feltételes szélsőértékek létezésére ad szükséges feltételt az alábbi alapvető jelentőségű tétel, mely az ún. *Lagrange-féle multiplikátor módszert* alapozza meg:

5.4. Tétel: Ha $f,g : \mathbf{R}^n \to \mathbf{R}$ folytonosan differenciálható függvények az $x_0 \in \mathbf{R}^n$ pont egy környezetében, f-nek x_0-ban lokális feltételes szélsőértéke van a $g(x) = 0$ feltétel mellett, továbbá grad $g(x_0) \neq \mathbf{0}$, akkor a grad $f(x_0)$ és a grad $g(x_0)$ vektorok párhuzamosak, azaz van oly $\lambda \in \mathbf{R}$ szám (az ún. *Lagrange-féle multiplikátor* vagy *Lagrange-szorzó*), hogy

$$\text{grad } f(x_0) = \lambda \cdot \text{grad } g(x_0)$$

Bizonyítás: A szigorú bizonyítás meghaladja e jegyzet kereteit, mert itt nem tárgyalt tételekre épül. Röviden vázolunk viszont egy nem egészen korrekt levezetést, mely mindazonáltal jól szemlélteti a tétel elméleti hátterét.

Csak a feltételes maximum esetével foglalkozunk, a feltételes minimum esete hasonlóan kezelhető. A többváltozós derivált definíciójából adódóan, x_0 egy kis környezetében lévő $x := x_0 + h$ alakú vektorokra:

$$g(x) \approx g(x_0) + \langle \text{grad } g(x_0), h \rangle.$$

Így tehát a $g(x) = 0$ feltétel nem minden x-re teljesül x_0 egy környezetéből, hanem csak azokra, melyek $x = x_0 + h$ alakúak, ahol $\langle \text{grad } g(x_0), h \rangle = 0$, azaz h ortogonális a gradiensvektorra. Ezekre a h-kra teljesül, hogy

$$f(x_0 + h) \leq f(x_0)$$

Ámde $f(x_0 + h) \approx f(x_0) + \langle \text{grad } f(x_0), h \rangle$, innen kapjuk, hogy minden (elég kis normájú) grad $g(x_0)$-ra ortogonális h vektor esetén $f(x_0) + \langle \text{grad } f(x_0), h \rangle \leq f(x_0)$, azaz:

$$\langle \text{grad } f(x_0), h \rangle \leq 0$$

h helyébe $(-h)$-t írva, ellenkező irányú egyenlőtlenséget kapunk. Mindkét egyenlőtlenség csak úgy teljesülhet, ha minden (elég kis normájú) grad $g(x_0)$-ra ortogonális h vektor esetén

$$\langle \text{grad } f(x_0), h \rangle = 0,$$

azaz a grad $f(x_0)$ vektor ortogonális az összes ilyen h-ra. Ebből már következik, hogy a grad $f(x_0)$ és a grad $g(x_0)$ vektorok szükségképp egymás számszorosai.

Javasoljuk az Olvasónak, hogy a fenti "bizonyítást" gondolja át $n \leq 3$ esetén (amikor az ortogonalitás a geometriai merőlegességgel egyezik, így a dolog eléggé szemléletes) olyan feltétel mellett, amikor g elsőfokú polinom, azaz a $g(x) = 0$ feltételt egy egyenes vagy egy sík pontjai teljesítik. Ekkor a kívánt eredmény korrekt módon levezethető.

A tétel gyakorlati alkalmazása rendszerint a következő sémában foglalható össze. Bevezetve az

$$L(x, \lambda) := f(x) - \lambda \cdot g(x)$$

formulával értelmezett $(n+1)$-változós függvényt (a *Lagrange-függvényt*), felírjuk és megoldjuk az alábbi, általában nemlineáris egyenletrendszert:

$$\frac{\partial L}{\partial x_k}(x,\lambda) = \frac{\partial f}{\partial x_k}(x) - \lambda \cdot \frac{\partial g}{\partial x_k}(x) = 0 \qquad (k = 1,2,\ldots,n)$$

$$\frac{\partial L}{\partial \lambda}(x,\lambda) = -g(x) = 0$$

A feltételes szélsőértékhely(ek) a megoldások között van(nak): ez(eke)t aztán külön-külön kell vizsgálni és megállapítani, hogy ott valóban feltételes szélsőérték van-e, és ha igen, milyen típusú.

A módszert néhány példán keresztül szemléltetjük.

5.13. Példa: Keressük meg az

$$f(x,y) := 2xy$$

kifejezéssel értelmezett függvény feltételes maximumhelyét a

$$g(x,y) := x + 3y - 12 = 0$$

feltétel mellett.

1. *Megoldás*: A problémát visszavezethetjük feltétel nélküli szélsőérték feladatra. A feltételből pl. x kifejezhető: $x = 12 - 3y$, ezt visszahelyettesítve f kifejezésébe, egyváltozós szélsőérték feladatot nyerünk:

$$f(20 - 3y,y) = 2 \cdot (12 - 3y)y = 24y - 6y^2 =: F(y)$$

Szélsőérték ott lehet, ahol az első derivált 0, azaz $F'(y) = 24 - 12y = 0$, tehát az $y = 2$ helyen. Itt pedig valóban lokális maximum van, mert $F''(y) = -12 < 0$. Az így kapott y értéket a feltételbe visszahelyettesítve kapjuk, hogy $x = 6$, azaz a feltételes szélsőértékhely koordinátái: $x = 6$, $y = 2$.

2. *Megoldás*: A Lagrange-féle módszert alkalmazzuk. A Lagrange-függvény:

$$L(x,y,\lambda) = 2xy - \lambda \cdot (x + 3y - 12)$$

Feltételes szélsőérték ott lehet, ahol a Lagrange-függvény parciális deriváltjai eltűnnek, azaz:

$$\frac{\partial L}{\partial x} = 2y - \lambda = 0$$

$$\frac{\partial L}{\partial y} = 2x - 3\lambda = 0$$

$$x + 3y = 12$$

Az első két egyenletből: $3\lambda = 6y = 2x$, innen $x = 3y$. Ezt beírva a 3. egyenletbe: $x + 3y = 6y = 12$, innen $y = 2$, $x = 6$.

Megjegyezzük, hogy a **??**. Tétel alapján nem lehet megállapítani, hogy a lehetséges helye(ke)n valóban van-e lokális feltételes szélsőérték, és ha igen, az minimum-e vagy maximum. Ezt egyéb meggondolásokkal (pl. feltétel nélküli szélsőérték feladatra való átírással) lehet vizsgálni.

5.14. Példa: Legyen $A \in \mathbf{M}_{n\times n}$ egy önadjungált mátrix. Jelölje S az n-dimenziós egységgömb felületét, azaz az 1 normájú vektorok halmazát: $S := \{x \in \mathbf{R}^n : ||x|| = 1\}$. Határozzuk meg az $\langle Ax,x\rangle$ kvadratikus alak maximumhelyét az S halmazon.

Megoldás: A probléma egy feltételes szélsőértékfeladat.

Meghatározandó az $f : \mathbf{R}^n \to \mathbf{R}^n$, $f(x) := \langle Ax,x\rangle$ függvény feltételes maximumhelye a $g(x) := ||x||^2 - 1 = 0$ feltétel mellett. A Lagrange-függvény:

$$L(x,\lambda) = \langle Ax,x\rangle - \lambda \cdot (||x||^2 - 1)$$

Kiszámítva az x szerinti gradiensvektort, az alábbi egyenletrendszerre jutunk:

$$2Ax - \lambda \cdot 2x = 0, \qquad ||x||^2 = 1$$

azaz $Ax = \lambda x$, $||x|| = 1$, tehát a maximumhely szükségképp sajátvektora A-nak, és a Lagrange-féle multiplikátor épp a hozzátartozó sajátérték. Megjegyezzük, hogy a feltételfüggvényt definiálhattuk volna pl. $g(x) := ||x|| - 1 = 0$ alakban is, de akkor a Lagrange-multiplikátor jelentése nem annyira szemléletes.

5.15. Példa: Tekintsük ismét a **??**. Példát:
Egy kerti (téglatest alakú) medencét szeretnénk építeni. A medence belső oldalait ki kell csempézni. Hogyan válasszuk meg a medence méreteit, hogy a térfogata 32 m^3, a csempézés költsége (ami a kicsempézendő felülettel egyenesen arányosnak vehető) pedig a lehető legkisebb legyen?

Megoldás: A probléma feltételes szélsőérték feladatként azonnal megfogalmazható. Jelölje ismét x, y a medence alapjának méreteit, z pedig a mélységét, akkor kicsempézendő felület nagysága:

$$F = xy + 2xz + 2yz$$

A térfogat: $V = xyz = 32$, így minimalizálandó az

$$F(x,y,z) := xy + 2xz + 2yz$$

formulával értelmezett függvény a

$$g(x,y,z) := xyz - 32 = 0$$

feltétel mellett. A Lagrange-függvény:

$$L(x,y,z,\lambda) := xy + 2xz + 2yz - \lambda \cdot (xyz - 32)$$

A feltételes minimumhelyen tehát teljesülnek az alábbi egyenlőségek:

$$\begin{aligned}
\frac{\partial L}{\partial x} &= y + 2z - \lambda yz = 0 \\
\frac{\partial L}{\partial y} &= x + 2z - \lambda xz = 0 \\
\frac{\partial L}{\partial z} &= 2x + 2y - \lambda xy = 0 \\
xyz &= 32
\end{aligned}$$

Az 1. és 2. egyenletből λ-t kifejezve:

$$\lambda = \frac{1}{z} + \frac{2}{y} = \frac{1}{z} + \frac{2}{x},$$

innen $x = y$. Ezt a 3. egyenletbe beírva kapjuk, hogy $\lambda = \frac{4}{x}$, amit visszaírva a 2. egyenletbe: $x + 2z - 4z = 0$, azaz $z = \frac{x}{2}$. Végül innen és a 4. egyenletből most már adódik, hogy $x = y = 4$, $z = 2$.

5.16. Példa: Tekintsük ismét a **??**. Példát:
Szimmetrikus trapéz keresztmetszetű egyenes csatornát építünk. Belső oldalait (mindenütt azonos vastagsággal) ki kell betonozni. Hogyan válasszuk meg a trapéz méreteit, hogy a csatorna keresztmetszete előírt T nagyságú, a betonozás folyóméterenkénti költsége pedig a lehető legkisebb legyen?

Megoldás: Jelölje ismét a a csatorna szélességét a fenéken (azaz a trapéz rövidebb párhuzamos oldalát), b a csatorna oldalfalának hosszát (azaz a trapéz szárait). Jelölje α az oldalfal lejtési szögét (a trapéz szára és hosszabbik párhuzamos oldala által bezárt szöget). A folyóméterenkénti betonozási költség nyilván arányos az $(a+2b)$ hossz-összeggel. Mivel a csatorna m mélysége nyilván $m = b\sin\alpha$, a csatorna keresztmetszete (a trapéz területe): $T = (a + b\cos\alpha)\cdot b\sin\alpha$. A probléma tehát a következő feltételes szélsőértékfeladatként fogalmazható meg: minimalizáljuk az

$$f(a,b,\alpha) := a + 2b$$

előírással értelmezett függvényt a

$$g(a,b,\alpha) := (a + b\cos\alpha)\cdot b\sin\alpha - T = ab\sin\alpha + b^2\cos\alpha\sin\alpha - T =$$

$$= ab\sin\alpha + \frac{b^2}{2}\sin 2\alpha - T = 0$$

feltétel mellett. A Lagrange-függvény:

$$L(a,b,\alpha,\lambda) = a + 2b - \lambda(ab\sin\alpha + \frac{b^2}{2}\sin 2\alpha - T)$$

A feltételes minimumhelyen tehát teljesülnek az alábbi egyenlőségek:

$$\frac{\partial L}{\partial a} = 1 - \lambda b\sin\alpha = 0$$

$$\frac{\partial L}{\partial b} = 2 - \lambda(a\sin\alpha + 2b\sin\alpha\cos\alpha) = 0$$

$$\frac{\partial L}{\partial \alpha} = -\lambda(ab\cos\alpha + b^2\cos 2\alpha) = 0$$

$$ab\sin\alpha + \frac{b^2}{2}\sin 2\alpha = T$$

Az 1. és a 2. egyenletből λ-t kifejezve:

$$\lambda = \frac{1}{b\sin\alpha} = \frac{2}{a\sin\alpha + 2b\sin\alpha\cos\alpha},$$

ahonnan kapjuk, hogy $a = 2b(1-\cos\alpha)$. Ezt beírva a 3. egyenletbe: $2b\cos\alpha - 2b\cos^2\alpha + bcos^2\alpha - b^2\sin^2\alpha = 0$, ahonnan $\cos\alpha = \frac{1}{2}$, azaz $\alpha = 60°$. Ezt visszaírva az imént nyert $a = 2b(1-\cos\alpha)$ egyenlőségbe, kapjuk, hogy $a = b$. Végül innen, és a 4. egyenletből: $a = b = \frac{2}{3}\sqrt{T\sqrt{3}}$.

Látjuk tehát, hogy *ugyanazt* a problémát legtöbbször feltételes és feltétel nélküli szélsőérték feladatként is meg lehet fogalmazni. Általános észrevételként megállapítható, hogy a feltételes szélsőérték feladatként való megfogalmazás sokkal természetesebb és sokkal gépiesebb is. Azonban a Lagrange-multiplikátor módszer által szolgáltatott egyenletrendszer mindig nagyobb méretű, és gyakorlati megoldása sokszor nehezebb. A feltétel nélküli szélsőérték feladatként való megfogalmazás nem mindig kézenfekvő (ld. a **??**. Példát), de ha ez már megtörtént, a megoldandó egyenletrendszer kisebb méretű (nincs benne a Lagrange-multiplikátor és egy további változót kifejezhetünk a feltételi egyenletből, így az ismeretlenek száma 2-vel kevesebb), és általában könnyebb megoldani. Így gyakorlati szélsőérték feladatok esetén célszerű mindkét megközelítést készenlétben tartani.

5.7. Néhány alkalmazás

Lineáris regresszió
Legyenek $(x_1,y_1),(x_2,y_2),...,(x_N,y_N)$ adott számpárok (pl. mérési adatok).

Tegyük fel, hogy az x-ek és y-ok közt közel lineáris kapcsolat van, azaz $y_k \approx a \cdot x_k + b$, ahol a, b egyelőre ismeretlen paraméterek. Feladatunk tehát meghatározni a-t és b-t úgy, hogy az $y = ax + b$ egyenletű egyenes a "lehető legjobban" illeszkedjék az (x_k, y_k) $(k = 1,2,...,N)$ adatokra.

A fenti probléma jellemzően *statisztikai* probléma, és számos gyakorlati alkalmazásban felbukkan, ahol egymástól függő adatokat kell kiértékelni. Ilyen problémák gyakran előfordulnak a biológiában, orvostudományban, közgazdaságban stb.

A probléma igazi nehézsége az, hogy a "lehető legjobb" illeszkedés eléggé intuitív, és számos különböző módon lehet értelmezni. Első pillanatra talán a legtermészetesebb az illeszkedés hibáját a

$$\max\{|ax_1 + b - y_1|, |ax_2 + b - y_2|, ..., |ax_N + b - y_N|\}$$

számmal, azaz a maximális abszolút eltéréssel mérni. Ennek matematikai kezelése azonban eléggé nehézkes, különösen nagy N esetén. *Technikailag* legkönnyebb az alábbi hiba használata:

$$E(a,b) := \sum_{k=1}^{N} (ax_k + b - y_k)^2$$

Ez is jól méri az illeszkedés hibáját abban az értelemben, hogy ha az adatok történetesen illeszkednek az egyenesre, akkor nyilván $E(a,b) = 0$, és minél inkább távol esnek az egyenestől, $E(a,b)$ értéke annál nagyobb. Kézenfekvő tehát "legjobban illeszkedő" egyenesnek azt nevezni, melyre $E(a,b)$ minimális: ezt *regressziós egyenesnek* nevezzük.

A probléma egy lehetséges megoldása tehát a fenti $E(a,b)$ kétváltozós függvény minimalizálása. Minimum ott lehet, ahol mindkét parciális derivált eltűnik, azaz

$$\frac{\partial E}{\partial a} = 2\sum_{k=1}^{N} (ax_k + b - y_k)x_k = 0$$

$$\frac{\partial E}{\partial b} = 2\sum_{k=1}^{N} (ax_k + b - y_k) = 0$$

Innen a-ra és b-re a következő egyenletrendszert nyerjük:

$$a\sum_{k=1}^{N} x_k^2 + b\sum_{k=1}^{N} x_k = \sum_{k=1}^{N} x_k y_k$$

$$a\sum_{k=1}^{N} x_k + b\sum_{k=1}^{N} 1 = \sum_{k=1}^{N} y_k$$

Az egyenletrendszernek egyetlenegy megoldása van, mert a rendszer determinánsa pozitív, ui. a Cauchy-egyenlőtlenség szerint

$$(\sum_{k=1}^{N} 1 \cdot x_k)^2 \leq (\sum_{k=1}^{N} 1^2) \cdot (\sum_{k=1}^{N} x_k^2),$$

és egyenlőség csak akkor fordul elő, ha az $(1,1,...,1)$ és az $(x_1,x_2,...,x_N)$ vektorok egymás számszorosai, azaz mindegyik x_k egyenlő; másszóval, ha az $(x_1,y_1),...,(x_N,y_N)$ adatpontok egy függőleges egyenesen helyezkednek el (ekkor egyúttal ez a legjobban közelítő egyenes is, de ez nem írható fel $y = ax + b$ alakban). Következésképp ettől a speciális esettől eltekintve az egyenletrendszer megoldása valóban egyértelmű. Az így adódó a, b paraméter mellett pedig valóban minimuma van E-nek, mert a második derivált mátrixa:

$$D^2F(x,y) = 2 \cdot \begin{pmatrix} \sum_{k=1}^{N} x_k^2 & \sum_{k=1}^{N} x_k \\ \sum_{k=1}^{N} x_k & \sum_{k=1}^{N} 1 \end{pmatrix},$$

és ez pozitív definit, mivel már láttuk, hogy determinánsa pozitív, nyoma pedig szintén pozitív.

Függvények közelítése polinomokkal

Legyen $f : [0,1] \to \mathbf{R}$ adott folytonos függvény. Keressünk olyan $p(x) := a_0 + a_1x + a_2x^2 + ... + a_nx^n$ legfeljebb n-edfokú polinomot, mely a "lehető legjobban" közelíti f-et!

A kérdés most is az, hogy hogyan definiáljuk a "lehető legjobb" közelítést. Ennek egyik legegyszerűbb módja, ha a közelítés hibáját az

$$E(a_0,a_1,....,a_n) := \int_0^1 (\sum_{j=0}^{n} a_jx^j - f(x))^2 dx$$

négyzetintegrál értékével mérjük. Valóban, ha f maga egy legfeljebb n-edfokú polinom, akkor a $p := f$ választás mellett $E = 0$. Kézenfekvő tehát "legjobban illeszkedő" polinomnak azt a p polinomot nevezni, melyre a fenti E minimális.

A probléma egy lehetséges megoldása tehát a fenti $(n + 1)$-változós E függvény minimalizálása. Minimum ott lehet, ahol az összes elsőrendű parciális derivált eltűnik, azaz

$$\frac{\partial E}{\partial a_k} = 2\sum_{j=1}^{n}\int_0^1 (a_j x^j - f(x))x^k dx = 0 \qquad (k = 0,1,...,n)$$

ahonnan az ismeretlen $a_0,a_1,...,a_n$ együtthatókra a következő egyenletrendszert nyerjük:

$$\sum_{j=0}^{n} \frac{1}{k+j+1} \cdot a_j = b_k \qquad (k = 0,1,...,n),$$

ahol $b_k := \int_0^1 f(x)x^k dx$.

A rendszer mátrixa $(n+1)$-edrendű Hilbert-mátrix:

$$H_{n+1} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{n+1} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \dots & \frac{1}{n+2} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \dots & \frac{1}{n+3} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{1}{n+1} & \frac{1}{n+2} & \frac{1}{n+3} & \dots & \frac{1}{2n+1} \end{pmatrix},$$

Bizonyítás nélkül megemlítjük, hogy H_{n+1} pozitív definit, így az egyenletrendszer egyértelműen megoldható (bár, mint már láttuk, nagyon gyengén meghatározott: az adatok ill. a számítások kis hibája a számított együtthatókban nagy változásokat eredményez). Az így adódó $a_0,a_1,...,a_n$ paraméterek mellett pedig valóban minimuma van E-nek, mert a második derivált mátrixa könnyen láthatóan $D^2E = 2H_{n+1}$, tehát szintén pozitív definit.

5.8. Többszörös integrálok

Ebben a szakaszban az integrálfogalmat általánosítjuk többváltozós függvényekre. Már most megjegyezzük, hogy a kérdéskör szabatos matematikai tárgyalása nagyon nehéz, és eddigi matematikai eszközeinkkel nem is lehetséges. Ezzel a mérték- és az integrálelmélet foglalkozik. Ugyanakkor az alkalmazások szempontjából legtöbb esetben nincs szükség ezekre az eszközökre, csak néhány, az integrálokra vonatkozó tételre. Ezért ezen szakasz tárgyalásmódjában az alábbi kompromisszumot tesszük. Lesznek tételek, melyeket bizonyítás nélkül mondunk ki és fogadunk el, és lesznek olyanok is, melyek bizonyításában a szemléletet részesítjük előnyben, és a levezetés nem minden lépését igazoljuk. Viszont a bizonyítás nélkül kimondott ill. felhasznált állításokra ill. következtetésekre minden esetben fel fogjuk hívni a figyelmet.

Az integrálfogalmat először kétváltozós, éspedig *téglalapon értelmezett* függvényekre általánosítjuk, de látni fogjuk, hogy ugyanez a konstrukció tetszőleges n-változós függvények esetében is végigvihető, értelemszerű változtatásokkal.

A témakörben valamennyire jártas Olvasóinknak felhívjuk a figyelmét, hogy a kimondásra kerülő definíciók egy része nem törekszik teljes általánosságra (pl. egy tartomány felbontásának fogalma vagy az integrálközelítő összegek értelmezése esetében). Céljainknak ez is megfelel: igyekeztünk az anyagot úgy felépíteni, hogy a lehető legkevesebb előismeretet követeljen.

Legyenek tehát $[a,b],[c,d] \subset \mathbf{R}$ véges intervallumok, és jelölje T a

$$T := [a,b] \times [c,d] = \{(x,y) \in \mathbf{R}^2 : a \leq x \leq b, \;\; c \leq y \leq d\}$$

téglalapot. Legyen $a = x_0 < x_1 < ... < x_N = b$ és $c = y_0 < y_1 < ... < y_M = d$ az $[a,b]$ ill. a $[c,d]$ intervallumok egy-egy – nem feltétlen egyenközű – felbontása. Jelölje $\Delta x_k := x_k - x_{k-1}$ ill. $\Delta y_j := y_j - y_{j-1}$ a lépésközöket ($k = 1,2,...,N$, $j = 1,2,...,M$). Akkor a $T_{kj} := [x_{k-1},x_k] \times [y_{j-1},y_j]$ téglalapok az eredeti T téglalap egy felbontását szolgálatják: T területe nyilván megegyezik a $\sum_{k=1}^{N} \sum_{j=1}^{M} \Delta x_k \Delta y_j$ területösszeggel.

Alsó, felső integrálközelítő összegek:

Az

$$S_{-}^{(N,M)}(f) := \sum_{k=1}^{N} \sum_{j=1}^{M} f_{kj}^{(min)} \Delta x_k \Delta y_j$$

számot az f függvénynek a fenti felbontáshoz tartozó *alsó integrálközelítő összegének* nevezzük. Hasonlóan, az

$$S_{+}^{(N,M)}(f) := \sum_{k=1}^{N} \sum_{j=1}^{M} f_{kj}^{(max)} \Delta x_k \Delta y_j$$

számot az f függvénynek a fenti felbontáshoz tartozó *felső integrálközelítő összegének* nevezzük, ahol $f_{kj}^{(min)}$ ill. $f_{kj}^{(max)}$ jelöli az f függvény minimális ill. maximális értékét a T_{kj} téglalapon.

Nyilván minden f folytonos függvény és T minden felbontása esetén $S_{-}^{(N,M)}(f) \leq S_{+}^{(N,M)}(f)$.

Alsó, felső integrál:

Az alsó integrálközelítő összegek halmazának $S_-(f)$-fel jelölt felső határát (szuprémumát) az f függvény *alsó integráljának* nevezzük. Hasonlóan, a felső integrálközelítő összegek halmazának $S_+(f)$-fel jelölt alsó határát (infimumát) az f függvény *felső integráljának* nevezzük.

Nyilván minden f folytonos függvény esetén $S_-(f) \leq S_+(f)$.

Riemann-integrál:

Azt mondjuk, hogy az f függvény *Riemann-integrálható a T téglalapon*, ha alsó és felső integrálja egyenlő. Ezt a közös értéket f-nak T-n vett Riemann-integráljának nevezzük és az $\int_T f$, $\int_T f(x,y)dxdy$ szimbólumok valamelyikével jelöljük.

Világos, hogy a definíció nehézség nélkül általánosítható kettőnél több változós függvényekre is. Két- és háromváltozós függvények esetén (utóbbi esetben T egy téglatest) szokásos még a $\int\int_T f(x,y)dxdy$ ill. az $\int\int\int_T f(x,y,z)dxdydz$ jelölés is. Elterjedt szóhasználat kétváltozós függvény Riemann-integrálját *kettős*, háromváltozós függvény Riemann-integrálját *hármas integrálnak* nevezni.

Bizonyítás nélkül megemlítjük, hogy – az egyváltozós függvények esetéhez hasonlóan – *minden, a T téglalapon folytonos függvény Riemann-integrálható is*.

A most bevezetett integrálfogalom az analízis egyik legfontosabb fogalma. Számos fizikai jelentése közül megemlítünk kettőt:

(a) Ha V egy háromdimenziós téglatest, f pedig valamilyen, V-ben eloszlott anyag koncentrációfüggvénye, azaz $f(x,y,z)$ jelenti az anyag koncentrációját az $(x,y,z) \in V$ pontban, akkor az $\int\int\int_V f(x,y,z)dxdydz$ hármas integrál az illető anyag teljes tömege a V térfogatban. Hasonlóan, ha f az energiasűrűség eloszlását írja le, akkor ez a hármas integrál a teljes energiát jelenti.

(b) Legyen $\Omega \subset \mathbf{R}^2$ egy korlátos, összefüggő halmaz, $T \supset \Omega$ egy bővebb téglalap. Képzeljünk el egy Ω alakú, mindenütt egyenletesen vékony és homogén lemezből készült síkidomot. Ennek a lemezdarabnak a súlypontja (M_x,M_y), ahol

$$M_x = \frac{\int\int_T x \cdot f(x,y)dxdy}{\int\int_T f(x,y)dxdy}$$

$$M_y = \frac{\int\int_T y \cdot f(x,y)dxdy}{\int\int_T f(x,y)dxdy}$$

és f az Ω tartomány *karakterisztikus függvénye*, azaz $f(x,y) = 1$, ha $(x,y) \in \Omega$, ill. $f(x,y) = 0$, ha (x,y) nincs Ω-ban.

A Riemann-integrál fenti definíciója az egyváltozós függvények Riemann-integráljának pontos megfelelője. Látni kell azonban, hogy a definíció gyakorlati számításokra teljesen alkalmatlan. Szemléletes jelentését viszont jól mutatja a következő tétel, melyet bizonyítás nélkül mondunk ki:

5.5. Tétel: Tekintsük az $[a,b]$ és $[c,d]$ intervallumok felbontásainak egy *korlátlanul finomodó sorozatát*, azaz tegyük fel, hogy a $\max_{k=1,...,N} \Delta x_k$ és a $\max_{j=1,...,M} \Delta y_j$ számok is 0-hoz tartanak. Legyen f_{kj} egy tetszőleges függvényérték, melyet f a T_{kj} téglalapon felvesz. Akkor a

$$\sum_{k=1}^{N} \sum_{j=1}^{M} f_{kj} \Delta x_k \Delta y_j$$

Riemann-összegek sorozata az $\int \int_T f(x,y)dxdy$ Riemann-integrálhoz tart (a konkrét felbontástól és az f_{kj} függvényértékek megválasztásától függetlenül).

A tétel egyúttal a T téglalapon vett Riemann-integrál közelítő kiszámítására is eljárást ad: elég egy "kellően finom" felbontás mellett egy Riemann-összeget kiszámítani. Arra nézve viszont, hogy ez milyen pontosan közelíti a Riemann-integrált, a tétel nem mond semmit. A pontossággal nem fogunk foglalkozni, mert a kettős integrálok kiszámítását *egyváltozós integrálok egymás utáni kiszámítására* vezetjük vissza a következő tétel segítségével:

5.6. Tétel: Ha f folytonos függvény T-n, akkor

$$\int \int_T f(x,y)dxdy =$$

$$= \int_a^b \left(\int_c^d f(x,y)dy \right) dx = \int_c^d \left(\int_a^b f(x,y)dx \right) dy$$

Bizonyítás vázlat: Közelítsük az $\int \int_T f(x,y)dxdy$ integrált az alábbi Riemann-összeggel:

$$\sum_{k=1}^{N} \sum_{j=1}^{M} f(x_k,y_j) \Delta x_k \Delta y_j$$

Összegezve először j szerint, minden rögzített k indexre:

$$\sum_{j=1}^{M} f(x_k,y_j) \Delta y_j \approx \int_c^d f(x_k,y)dy$$

Δx_k-val szorozva, és k szerint összegezve:

$$\sum_{k=1}^{N}\sum_{j=1}^{M} f(x_k,y_j)\Delta y_j \Delta x_k \approx \sum_{k=1}^{N}\left(\int_c^d f(x_k,y)dy\right)\cdot \Delta x_k$$

A jobb oldalon szintén egy Riemann-összeg áll, éspedig az $F(x) := \int_c^d f(x,y)dy$ formulával értelmezett F függvény egy Riemann-összege. Ezért:

$$\sum_{k=1}^{N}\sum_{j=1}^{M} f(x_k,y_j)\Delta y_j \Delta x_k \approx \int_a^b F(x)dx = \int_a^b \left(\int_c^d f(x,y)dy\right)dx$$

Ha most a felbontás korlátlanul finomodik, akkor a közelítő egyenlőségek pontos egyenlőségekbe mennek át: ezt azonban nem bizonyítjuk.

A másik egyenlőséget ugyanilyen meggondolásokkal láthatjuk be.

A tétel azonnali következménye, hogy ha $f(x,y) = g(x)\cdot h(y)$ alakú, akkor a kettős integrál kiszámítása különösen egyszerű:

5.2. Következmény: Legyenek $g : [a,b] \to \mathbf{R}$, $h : [c,d] \to \mathbf{R}$ folytonos függvények, és legyen $f(x,y) = g(x)\cdot h(y)$. Akkor:

$$\int\int_T f(x,y)dxdy = \left(\int_a^b g(x)dx\right)\cdot\left(\int_c^d h(y)dy\right)$$

Az előző tétel és következmény értelemszerűen általánosítható kettőnél több változós függvények esetére.

A fenti tételek alkalmazását az alábbi példákon illusztráljuk:

5.17. Példa: Legyen $T := [0,2]\times[0,1]$. Számítsuk ki a $\int\int_T (x+xy^2)dxdy$ kettős integrált.

Megoldás: $\int\int_T (x+xy^2)dxdy = \int_0^2\left(\int_0^1 (x+xy^2)dy\right)dx$. A belső integrálban az integrálás y szerint történik, x-et konstansnak tekintve:

$$\int_0^1 (x+xy^2)dy = \left[xy + \frac{xy^3}{3}\right]_0^1 = \frac{4x}{3}$$

Most már a külső integrál is kiszámítható, és:

$$\int\int_T (x+xy^2)dxdy = \int_0^2 \frac{4x}{3}dx = \left[\frac{4x^2}{6}\right]_0^2 = \frac{8}{3}.$$

Eljárhattunk volna úgy is, hogy előbb x szerint integrálunk:

$$\int_0^2 (x + xy^2)dx = (1 + y^2) \cdot \left[\frac{x^2}{2}\right]_0^2 = 2 \cdot (1 + y^2),$$

majd pedig y szerint, így kapjuk, hogy

$$\int\int_T (x + xy^2)dxdy = \int_0^1 2 \cdot (1 + y^2)dy = 2 \cdot \left[y + \frac{y^3}{3}\right]_0^1 = \frac{8}{3}.$$

5.18. Példa: Számítsuk ki a $\int\int_T x \cdot \sin(x + y)dxdy$ kettős integrált a $0 \leq x \leq \pi$, $0 \leq y \leq \frac{\pi}{2}$ egyenlőtlenségek által meghatározott T téglalapon.

Megoldás: Célszerű (de nem kötelező) először y szerint integrálni. $t := x + y$ helyettesítéssel:

$$\int_0^{\pi/2} x \cdot \sin(x + y)dy = x \cdot \int_x^{x+\pi/2} \sin t dt =$$

$$= x \cdot (-\cos(\frac{\pi}{2} + x) + \cos x) = x \cdot (\sin x + \cos x).$$

Most x szerint integrálunk, parciális integrálást alkalmazva: legyenek $u := x$, $v' := \sin x + \cos x$, akkor $u' = 1$, és $v = -\cos x + \sin x$, innen:

$$\int\int_T x \cdot \sin(x + y)dxdy =$$

$$= [x \cdot (-\cos x + \sin x)]_0^\pi - \int_0^\pi (-\cos x + \sin x)dx =$$

$$= \pi - [-\sin x - \cos x]_0^\pi = \pi - 2.$$

A megoldáshoz a **??**. Következmény alkalmazásával is eljuthatunk:

$$\int\int_T x \cdot \sin(x + y)dxdy =$$

$$= \int_0^\pi \int_0^{\pi/2} (x \cdot \sin x \cos y + x \cdot \cos x \sin y)dydx =$$

$$= \left(\int_0^\pi x \sin x dx\right) \cdot \left(\int_0^{\pi/2} \cos y dy\right) +$$

$$+ \left(\int_0^\pi x \cos x dx\right) \cdot \left(\int_0^{\pi/2} \sin y dy\right)$$

A fellépő (egyváltozós) integrálokat kiszámítva, a jobb oldal értékére a következő adódik: $\pi \cdot 1 + (-2) \cdot 1 = \pi - 2$.

5.19. Példa: Legyen $T := [0,1] \times [0,1]$. Számítsuk ki a $\int\int_T \frac{1}{(x+y+1)^2} dxdy$ kettős integrált.

Megoldás: Először y szerint integrálva ($t := x + y + 1$ helyettesítéssel), majd x szerint integrálva:

$$\int\int_T \frac{1}{(x+y+1)^2} dxdy = \int_0^2 \left(\int_0^1 \frac{1}{(x+y+1)^2} dy \right) dx =$$

$$= \int_0^2 \left(\int_{x+1}^{x+2} \frac{1}{t^2} dt \right) dx = \int_0^2 \left(-\frac{1}{x+2} + \frac{1}{x+1} \right) dx =$$

$$= [-\log|x+2| + \log|x+1|]_0^1 = \log\frac{4}{3}$$

A Riemann-integrált most téglalapnál általánosabb halmazok esetére fogjuk definiálni. Legyen T ismét téglalap, $\Omega \subset T$ egyelőre tetszőleges (korlátos) halmaz, és $f : \Omega \to \mathbf{R}$ adott folytonos függvény. Értelmezzük az alsó és felső integrálközelítő összegeket ugyanúgy mint idáig, azzal a különbséggel, hogy az összegekbe csak azon T_{kj} téglalapoknak megfelelő tagok kerülnek, melyekre $T_{kj} \subset \Omega$:

$$S_-^{(N,M)}(f) := \sum_{T_{kj} \subset \Omega} f_{kj}^{(min)} \Delta x_k \Delta y_j$$

$$S_+^{(N,M)}(f) := \sum_{T_{kj} \subset \Omega} f_{kj}^{(max)} \Delta x_k \Delta y_j$$

Ezekután a Riemann-integrált és a Riemann-összegeket is az előbbiekhez hasonlóan lehet definiálni. A helyzet most bonyolultabb, mert a felbontás sűrítésével az integrálközelítő összegekben figyelembe vett T_{kj} téglalapok összessége általában nem marad állandó, hanem egy olyan Ω_{NM} halmazt alkot, mely bizonyos értelemben "konvergál" Ω-hoz, ha a felbontás korlátlanul finomodik. A T_{kj} téglalapok területeinek összegei ekkor – a gyakorlat számára fontos esetek túlnyomó többségében – egy számhoz tartanak, melynek szemléletes jelentése az Ω halmaz *területe*. A geometriai terület fogalma mindazonáltal nem általánosítható *tetszőleges* Ω halmazra: ezen kérdések részleteivel a *mértékelmélet* foglalkozik. Jelen jegyzet keretein belül csak egy nagyon speciális esetet tárgyalunk, mikor Ω bizonyos egyenesek és bizonyos folytonos egyváltozós függvények grafikonjai által határolt halmaz (ezeket *normáltartományoknak* nevezzük). Itt ki lesz használva Ω kétdimenziós volta: az eredmény magasabb dimenziós terekre csak nehézkesen általánosítható.

5.7. Tétel: Legyenek $g,h : [a,b] \to \mathbf{R}$ folytonos függvények, melyekre $g \leq h$ teljesül az $[a,b]$ intervallumon. Tekintsük az

$$\Omega := \{(x,y) \in \mathbf{R}^2 : \; a \leq x \leq b, \, g(x) \leq y \leq h(y)\}$$

normáltartományt. Legyen $f : \Omega \to \mathbf{R}$ folytonos függvény, akkor

$$\int\int_T f(x,y)dxdy = \int_a^b \left(\int_{g(x)}^{h(x)} f(x,y)dy \right) dx$$

Bizonyítás vázlat: A **??**. Tétel bizonyítás-vázlatához hasonló gondolatmenetet követünk. Az $\int\int_T f(x,y)dxdy$ integrál egy Riemann-összege:

$$\sum_{k=1}^{N} \sum_j f(x_k,y_j)\Delta y_j \Delta x_k,$$

ahol a j szerinti összegezés – rögzített k mellett – olyan j indexekre terjed ki, melyekre $g(x_k) \leq y_j \leq h(x_k)$. Ekkor a j szerinti összeg az $\int_{g(x_k)}^{h(x_k)} f(x_k,y)dy$ integrál egy Riemann-féle összege, ezért:

$$\sum_j f(x_k,y_j)\Delta y_j \approx \int_{g(x_k)}^{h(x_k)} f(x_k,y)dy$$

Δx_k-val szorozva, és k szerint összegezve:

$$\sum_{k=1}^{N} \sum_j f(x_k,y_j)\Delta y_j \Delta x_k \approx \sum_{k=1}^{N} \left(\int_{g(x_k)}^{h(x_k)} f(x_k,y)dy \right) \cdot \Delta x_k$$

A jobb oldalon szintén egy Riemann-összeg áll, éspedig az $F(x) := \int_{g(x_k)}^{h(x_k)} f(x,y)dy$ formulával értelmezett F függvény egy Riemann-összege. Ezért:

$$\sum_{k=1}^{N} \sum_j f(x_k,y_j)\Delta y_j \Delta x_k \approx \int_a^b F(x)dx =$$

$$= \int_a^b \left(\int_{g(x_k)}^{h(x_k)} f(x,y)dy \right) dx$$

Ha most a felbontás korlátlanul finomodik, akkor a közelítő egyenlőségek pontos egyenlőségekbe mennek át: ezt azonban nem bizonyítjuk.

A tételben az integrálások sorrendje – értelemszerűen – most nem cserélhető fel.

Hasonló tétel fogalmazható meg az

$$\Omega := \{(x,y) \in \mathbf{R}^2 : \ a \leq y \leq b, g(y) \leq x \leq h(y)\}$$

típusú (tehát vízszintes egyenesek közötti) normáltartományra is. Ez esetben tetszőleges $f : \Omega \to \mathbf{R}$ folytonos függvényre:

$$\int\int_T f(x,y)dxdy = \int_a^b \left(\int_{g(y)}^{h(y)} f(x,y)dx \right) dy.$$

5.20. Példa: Legyen Ω egy origó közepű R sugarú kör. Számítsuk ki az $\int\int_\Omega 1\ dxdy$ kettős integrált.

Megoldás: Ω normáltartomány, az $[-R,R] \to \mathbf{R}$, $x \to -\sqrt{R^2 - x^2}$ és a $[-R,R] \to \mathbf{R}$, $x \to \sqrt{R^2 - x^2}$ függvények grafikonjai határolják. Innen:

$$\int\int_\Omega 1\ dxdy = \int_{-R}^R \left(\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} 1 \right) dy = 2 \cdot \int_{-R}^R \sqrt{R^2 - x^2} dx =$$

$$= 4 \cdot \int_0^R \sqrt{R^2 - x^2} dx$$

A jobb oldalon $x = R \sin t$ helyettesítést alkalmazva:

$$\int\int_\Omega 1\ dxdy = 4 \cdot \int_0^{\pi/2} \sqrt{R^2 - R^2 \sin^2 t} \cdot R \cos t\ dt =$$

$$= 4R^2 \int_0^{\pi/2} \cos^2 t\ dt = 4R^2 \int_0^{\pi/2} \frac{1 + \cos 2t}{2}\ dt =$$

$$= 4R^2 \cdot \frac{1}{2} \cdot \frac{\pi}{2} == R^2 \pi.$$

Az azonosan 1 függvénynek az Ω tartományon vett integrálja tehát Ω területével egyezik, ami a Riemann-összegekkel való közelítés értelmében szemléletesen nyilvánvaló.

5.21. Példa: Számítsuk ki az $\int\int_\Omega (x^2 + y)\ dxdy$ integrált, ahol Ω az $y = x^2$ és az $y^2 = x$ egyenletű parabolák által határolt tartomány.

Megoldás:

$$\int\int_\Omega (x^2 + y)\ dxdy = \int_0^1 \left(\int_{x^2}^{\sqrt{x}} (x^2 + y) dy \right) dx =$$

$$= \int_0^1 \left[x^2y + \frac{y^2}{2}\right]_{x^2}^{\sqrt{x}} dx == \int_0^1 (x^{5/2} + \frac{x}{2} - x^4 - \frac{x^4}{2})dx =$$

$$= \left[\frac{2x^{7/2}}{7} + \frac{x^2}{4} - \frac{3}{2} \cdot \frac{x^5}{5}\right]_0^1 = \frac{33}{140}.$$

5.22. Példa: Határozzuk meg az egységnyi sugarú félkör súlypontját.

Megoldás: Helyezzük a félkört a koordinátarendszerbe úgy, hogy a félkörvonal a $[-1,1] \to \mathbf{R}$, $x \to \sqrt{1-x^2}$ függvény grafikonja legyen. Ekkor a súlypont koordinátái:

$$M_x = \frac{\int\int_\Omega x\, dxdy}{\int\int_\Omega 1\, dxdy},$$

$$M_y = \frac{\int\int_\Omega y\, dxdy}{\int\int_\Omega 1\, dxdy},$$

ahol Ω jelöli a szóban forgó félkört. A nevező épp e félkör területe, azaz $\frac{\pi}{2}$. Továbbá:

$$\int\int_\Omega x\, dxdy =$$

$$= \int_{-1}^1 \left(\int_0^{\sqrt{1-x^2}} x\, dy\right) dx = \int_{-1}^1 x\sqrt{1-x^2} dx = 0,$$

mert az integrandusz páratlan függvény. A másik integrál:

$$\int\int_\Omega y\, dxdy = \int_{-1}^1 \left(\int_0^{\sqrt{1-x^2}} y\, dy\right) dx =$$

$$= \int_{-1}^1 \left[\frac{1}{2}y^2\right]_0^{\sqrt{1-x^2}} dx =$$

$$= \frac{1}{2} \cdot \int_{-1}^1 (1-x^2)dx = \frac{1}{2} \cdot \left[x - \frac{x^3}{3}\right]_{-1}^1 = \frac{2}{3}$$

Innen a súlypont koordinátái: $M_x = 0$, $M_y = \frac{4}{3\pi}$.

Speciális eset: integrálás polárkoordináták szerint

Ha az integrálási tartomány kör vagy körszerű tartomány (körcikk, körgyűrű), akkor az integrálás sokszor egyszerűsíthető polárkoordináták alkalmazásával. Ekkor az Ω tartományt téglalapok helyett körcikkekkel fedjük le. Pontosabban, legyen Ω_0 egy origó középpontú R sugarú kör, ahol

R akkora, hogy $\Omega_0 \supset \Omega$ teljesüljön. Legyen $0 = r_0 < r_1 < ... < r_M = R$ a $[0,R]$, $0 = \phi_0 < \phi_1 < ... < \phi_N = 2\pi$ pedig a $[0,2\pi]$ intervallum egy felbontása. Ezek létrehoznak egy olyan felbontását Ω_0-nak, mely csupa körcikkből áll. Az $[r_{j-1},r_j]$ sugár- és a $[\phi_{k-1},\phi_k]$ szögintervallum által meghatározott B_{kj} körcikk területe: $\frac{r_{j-1}+r_j}{2}\Delta r_j \Delta\phi_k$, ahol $\Delta r_j := r_j - r_{j-1}$, $\Delta\phi_k := \phi_k - \phi_{k-1}$.

Legyen $f : \Omega \to \mathbf{R}$ tetszőleges folytonos függvény. Tetszőleges $(x,y) \in \mathbf{R}^2$ esetén jelölje r és ϕ az (x,y) pont *polárkoordinátáit*, azaz azokat a számokat, melyekre $x = r\cos\phi$ és $y = r\sin\phi$ (nyilván $r = \sqrt{x^2+y^2}$ és tg$\phi = \frac{y}{x}$). Jelölje F az f függvény *polárkoordinátás alakját*, azaz az $F(r,\phi) := f(r\cos\phi, r\sin\phi)$ előírással értelmezett függvényt. Akkor az $\int\int_\Omega f(x,y)dxdy$ integrál egy Riemann-összege (a körcikk-felbontásra nézve):

$$\sum_{k,j} f(r_j\cos\phi_k, r_j\sin\phi_k)\cdot\frac{r_{j-1}+r_j}{2}\Delta r_j\Delta\phi_k =$$

$$= \sum_{k,j} F(r_j,\phi_k)\cdot\frac{r_{j-1}+r_j}{2}\Delta r_j\Delta\phi_k,$$

ahol az összegzés csak azokra a k,j indexpárokra terjed ki, melyekre a B_{kj} körcikk Ω-ban van: $B_{kj} \subset \Omega$. Ha a felbontás elég finom, akkor $\frac{r_{j-1}+r_j}{2} \approx r_j$, így a jobb oldali összeg közelítően: $\sum_{k,j} F(r_j,\phi_k)\cdot r_j\Delta r_j\Delta\phi_k$ alakba írható, mely viszont az $(r,\phi) \to r\cdot F(r,\phi)$ leképezés egy Riemann-összege. Igazolható, hogy ha a felbontás korlátlanul finomodik, akkor a közelítő egyenlőségek pontos egyenlőségekkel válthatók fel, így nyertük, hogy:

5.8. Tétel: Jelölje $\tilde{\Omega}$ az Ω tartomány polárkoordinátás megfelelőjét, azaz $\tilde{\Omega} := \{(r,\phi) \in \mathbf{R}^2 : (r\cos\phi, r\sin\phi) \in \Omega$. Legyen $f : \Omega \to \mathbf{R}$ folytonos függvény, jelölje F ennek polárkoordinátás alakját, azaz $F(r,\phi) := f(r\cos\phi, r\sin\phi)$. Akkor

$$\int\int_\Omega f(x,y)dxdy = \int\int_{\tilde{\Omega}} F(r,\phi)rdrd\phi.$$

A tételt jellemzően akkor lehet jól alkalmazni, ha a transzformált $\tilde{\Omega}$ tartomány már normáltartomány az (r,ϕ) síkon, éspedig Ω-nál egyszerűbb. Így pl. egy origó középű R sugarú negyedkör megfelelője az (r,ϕ) síkon a $[0,1]\times[0,\frac{\pi}{2}]$ téglalap; az R_1, R_2 sugarú körök közötti körgyűrűtartomány megfelelője az (r,ϕ) síkon a $[R_1,R_2]\times[0,2\pi]$ téglalap, és így tovább.

5.23. Példa: Határozzuk meg az egységsugarú félkör súlypontját.

Megoldás: A feladatot most polárkoordináták használatával oldjuk meg. A félkörnek az az (r,ϕ) síkon a $0 \le r \le 1$, $0 \le \phi \le \pi$ egyenlőtlenségek által meghatározott téglalap felel meg. Ezért:

$$\int\int_\Omega 1\, dxdy = \int_0^\pi \int_0^1 r dr d\phi = \left(\int_0^\pi 1\, d\phi\right) \cdot \left(\int_0^1 r dr\right) = \frac{\pi}{2}$$

$$\int\int_\Omega x\, dxdy = \int_0^\pi \int_0^1 r\cos\phi \cdot r dr d\phi =$$

$$= \left(\int_0^\pi \cos\phi\, d\phi\right) \cdot \left(\int_0^1 r^2 dr\right) = 0$$

$$\int\int_\Omega y\, dxdy = \int_0^\pi \int_0^1 r\sin\phi \cdot r dr d\phi =$$

$$= \left(\int_0^\pi \sin\phi\, d\phi\right) \cdot \left(\int_0^1 r^2 dr\right) = \frac{2}{3}$$

innen ismét megkaptuk a súlypont koordinátáit: $(0, \frac{4}{3\pi})$.

5.24. Példa: Számítsuk ki az $\int\int_\Omega xy\, dxdy$ integrált, ahol Ω az $\frac{1}{2} \le r \le 1$, $1 \le \phi \le \frac{\pi}{2}$ egyenlőtlenségek által meghatározott körgyűrűcikk.

Megoldás: Áttérve polárkoordinátákra:

$$\int\int_\Omega xy\, dxdy = \int_0^{\pi/2} \int_{1/2}^1 r\cos\phi\, r\sin\phi\, r dr d\phi =$$

$$= \left(\int_0^{\pi/2} \cos\phi \sin\phi\, d\phi\right) \cdot \left(\int_{1/2}^1 r^3 dr\right) =$$

$$= \frac{1}{2}\left(\int_0^{\pi/2} \sin 2\phi\, d\phi\right) \cdot \left(\int_{1/2}^1 r^3 dr\right) = \frac{15}{128}.$$

5.9. Feladatok

1. Igazoljuk, hogy az

$$f(x,y) := \begin{cases} x \cdot \sin \frac{1}{y} + y \cdot \sin \frac{1}{x}, & \text{ha } xy \neq 0 \\ 0, & \text{ha } xy = 0 \end{cases}$$

formulával értelmezett függvény folytonos az origóban.

2. Igazoljuk, hogy az

$$f(x,y) := \begin{cases} \frac{x^2+y^2}{\sqrt{x^2+y^2+1}-1}, & \text{ha } x^2+y^2 \neq 0 \\ 0, & \text{ha } x^2+y^2 = 0 \end{cases}$$

formulával értelmezett függvény nem folytonos az origóban.

3. Legyen $f : \mathbf{R}^n \to \mathbf{R}$, $f(x) := \log ||x||$ $(x \neq \mathbf{0})$. Számítsuk ki a gradiensvektort.

4. Mutassuk meg, hogy

(a) az

$$u(x,y) := \log \frac{1}{\sqrt{x^2+y^2}}$$

formulával értelmezett (kétváltozós) függvényre $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \equiv 0$ teljesül $((x,y) \neq (0,0))$;

(b) az

$$u(x,y,z) := \frac{1}{\sqrt{x^2+y^2+z^2}}$$

formulával értelmezett (háromváltozós) függvényre $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \equiv 0$ teljesül $((x,y,z) \neq (0,0,0))$.

5. Hol és milyen típusú lokális szélsőértéke van az

$$f : \mathbf{R}^2 \to \mathbf{R}, \quad f(x,y) := x^2 + xy + y^2 - 5x - 4y + 1$$

függvénynek?

6. Egy téglatest egy csúcsba összefutó éleinek összhossza 24 cm. Legfeljebb mekkora a térfogata?

7. Van egy 2 méteres madzagunk, ezzel átkötünk egy téglatest alakú csomagot, méghozzá két irányban is. Legfeljebb mekkora lehet a csomag térfogata?

8. Tervezzünk egy 16 m^3 térfogatú téglatest alakú medencét. A fenék kicsempézése olyan csempével történik, melynek négyzetméterára feleakkora mint az oldalfalakra kerülő csempéé. Hogyan válasszuk meg a méreteket, hogy a csempézés költsége minimális legyen?

9. Legyen Ω az egységkör. Számítsuk ki az $\int\int_\Omega (x+y)^2 dxdy$ integrált.

Megoldások

1. Legyenek $x_n,y_n \to 0$ tetszőleges zérussorozatok, akkor:

$$|f(x_n,y_n) - f(0,0)| \leq |x_n| \cdot |\sin\frac{1}{y_n}| + |y_n| \cdot |\sin\frac{1}{x_n}| \leq |x_n| + |y_n|$$

$$\to 0,$$

tehát f valóban folytonos az origóban.

2. Legyenek $x_n,y_n \to 0$ tetszőleges zérussorozatok, akkor:

$$|f(x_n,y_n) - f(0,0)| \leq \frac{(x_n^2+y_n^2)\cdot(\sqrt{x_n^2+y_n^2+1}+1)}{x_n^2+y_n^2+1-1} \to 2 \neq 0$$

tehát f nem folytonos az origóban.

3. grad $f(x) = \frac{x}{||x||^2}$

4.

(*a*) Nyilván $u(x,y) = -\frac{1}{2}\log(x^2+y^2)$, innen

$$\frac{\partial u}{\partial x} = -\frac{x}{x^2+y^2},$$

$$\frac{\partial^2 u}{\partial x^2} = -\frac{(x^2+y^2) - x\cdot 2x^2}{(x^2+y^2)^2} = -\frac{-x^2+y^2}{(x^2+y^2)^2}$$

Hasonlóan:

$$\frac{\partial u}{\partial y} = -\frac{y}{x^2+y^2},$$

$$\frac{\partial^2 u}{\partial y^2} = -\frac{(x^2+y^2) - y\cdot 2y^2}{(x^2+y^2)^2} = -\frac{-y^2+x^2}{(x^2+y^2)^2},$$

így kettőjük összege valóban azonosan zérus (ahol értelmezett).

(*b*) Nyilván $u(x,y,z) = (x^2+y^2+z^2)^{-1/2}$, innen

$$\frac{\partial u}{\partial x} = -x\cdot(x^2+y^2+z^2)^{-3/2},$$

$$\frac{\partial^2 u}{\partial x^2} = -(x^2+y^2+z^2)^{-3/2} + x\cdot 3x\cdot(x^2+y^2+z^2)^{-5/2} =$$

$$= \frac{-x^2 - y^2 - z^2 + 3x^2}{(x^2 + y^2 + z^2)^{5/2}}$$

Hasonlóan:

$$\frac{\partial u}{\partial y} = -y \cdot (x^2 + y^2 + z^2)^{-3/2},$$

$$\frac{\partial^2 u}{\partial y^2} = -(x^2 + y^2 + z^2)^{-3/2} + y \cdot 3y \cdot (x^2 + y^2 + z^2)^{-5/2} =$$

$$= \frac{-x^2 - y^2 - z^2 + 3y^2}{(x^2 + y^2 + z^2)^{5/2}}$$

és

$$\frac{\partial u}{\partial z} = -z \cdot (x^2 + y^2 + z^2)^{-3/2},$$

$$\frac{\partial^2 u}{\partial z^2} = -(x^2 + y^2 + z^2)^{-3/2} + z \cdot 3z \cdot (x^2 + y^2 + z^2)^{-5/2} =$$

$$= \frac{-x^2 - y^2 - z^2 + 3z^2}{(x^2 + y^2 + z^2)^{5/2}}$$

A három kifejezés összege tehát valóban azonosan zérus (ahol értelmezett).

5. Szélsőérték ott lehet, ahol mindkét változó szerinti parciális derivált zérus, azaz

$$\frac{\partial f}{\partial x} = 2x + y - 5 = 0, \qquad \frac{\partial f}{\partial y} = x + 2y - 4 = 0$$

azaz az $(x,y) = (2,1)$ helyen. A második derivált mátrix konstans mátrix:

$$D^2 f(x,y) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

mégpedig pozitív definit (nyoma és determinánsa egyaránt pozitív). Az $(x,y) = (2,1)$ helyen a függvénynek tehát lokális minimuma van.

6. A feladat egy feltételes szélsőértékfeladat: maximalizálandó a $V(x,y,z) := xyz$ térfogat a $g(x,y,z) := x + y + z - 24 = 0$ feltétel mellett. A Lagrange-függvény: $L(x,y,z,\lambda) = xyz - \lambda(x + y + z - 24)$. Feltételes szélsőérték ott lehet, ahol L parciális deriváltjai eltűnnek, azaz:

$$\frac{\partial L}{\partial x} = yz - \lambda = 0$$

$$\frac{\partial L}{\partial y} = xz - \lambda = 0$$

$$\frac{\partial L}{\partial z} = xy - \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = -(x + y + z - 24) = 0$$

Innen $x = y = z = 8$ cm (tehát, ha a téglatest kocka), azaz a térfogat legfeljebb 512 cm^3. Az Olvasó bizonyítsa be, hogy itt valóban lokális feltételes maximum van.

7. A feladat egy feltételes szélsőértékfeladat: maximalizálandó a $V(x,y,z) := xyz$ térfogat a $g(x,y,z) := 2x + 4y + 2z - 2 = 0$ feltétel mellett. A Lagrange-függvény: $L(x,y,z,\lambda) = xyz - \lambda(2x + 4y + 2z - 2)$. Feltételes szélsőérték ott lehet, ahol L parciális deriváltjai eltűnnek, azaz:

$$\frac{\partial L}{\partial x} = yz - 2\lambda = 0$$

$$\frac{\partial L}{\partial y} = xz - 4\lambda = 0$$

$$\frac{\partial L}{\partial z} = xy - 2\lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = -(2x + 4y + 2z - 2) = 0$$

Innen $x = z = \frac{1}{3}$ m, $y = \frac{1}{6}$ m. A térfogat tehát legfeljebb $\frac{1}{54}\text{m}^3$. Az Olvasó bizonyítsa be, hogy itt valóban lokális feltételes maximum van.

8. A feladat egy feltételes szélsőértékfeladat: minimalizálandó a $f(x,y,z) := \frac{1}{2}xy + 2xz + 2yz$ költség a $V(x,y,z) := xyz - 16 = 0$ feltétel mellett. A Lagrange-függvény: $L(x,y,z,\lambda) = \frac{1}{2}xy + 2xz + 2yz - \lambda(xyz - 16)$. Feltételes szélsőérték ott lehet, ahol L parciális deriváltjai eltűnnek, azaz:

$$\frac{\partial L}{\partial x} = \frac{1}{2}y + 2z - \lambda yz = 0$$

$$\frac{\partial L}{\partial y} = \frac{1}{2}x + 2z - \lambda xz = 0$$

$$\frac{\partial L}{\partial x} = 2x + 2y - \lambda xy = 0$$

$$\frac{\partial L}{\partial \lambda} = -(xyz - 16) = 0$$

Az első két egyenletből: $\lambda = \frac{1}{2z} + \frac{1}{y} = \frac{1}{2z} + \frac{1}{x}$, innen $x = y$. Ezt beírva a 3. egyenletbe, kapjuk, hogy $\lambda = \frac{4}{x}$. Innen, és a 2. egyenletből: $z = \frac{x}{4}$. Végül a 4. egyenlet alapján: $x = y = 4$ m, $z = 1$ m. Az Olvasóra bízzuk annak belátását, hogy itt valóban feltételes lokális minimum van.

9. Áttérve polárkoordinátákra:

$$\int\int_{\Omega} (x+y)^2 dxdy = \int_0^{2\pi} \int_0^1 (r\cos t + r\sin t)^2 r dr dt =$$

$$= \int_0^{2\pi} \int_0^1 r^3 (\cos^2 t + 2\cos t \sin t + \sin^2 t)\, dr dt =$$

$$= \left(\int_0^{2\pi} (1 + \sin 2t) dt \right) \cdot \left(\int_0^1 r^3 dr \right) = \frac{\pi}{2}.$$

Tárgymutató

alsó, felső integrál 120
alsó, felső integrálközelítő összegek 119
altér 8

bázis, dimenzió 15

determináns 60
differenciálhatóság, gradiens 98

feltételes szélsőérték 110
folytonosság 95

kvadratikus alak 78

lineáris függetlenség 13
lineáris leképezés 49
lokális szélsőérték 104

magtér 59
mátrix 51
mátrix adjungáltja 76
mátrix definitsége 78
mátrix inverze 56
mátrixok összege és konstans-szorosa 52
mátrixszorzás 52

nilpotens mátrix 83
norma 16

önadjungált mátrix 76
ortogonális bázis, ortonormált bázis 21
ortogonális mátrix 84
ortogonális vektorok 19

projektormátrix 80

Riemann-integrál 120

sajátérték, sajátvektor 70
skaláris szorzat 16

távolság 16

vektoriális szorzat 36
vektorsorozat konvergenciája 95
vektortér 7

Ajánlott irodalom

Freud Róbert: Lineáris algebra
ELTE, Eötvös Kiadó, Budapest, 2006.

Horváth Erzsébet: Lineáris algebra
Műegyetemi Kiadó, Budapest, 1998.

Obádovics J. Gyula: Lineáris algebra példákkal
SCOLAR Bt, Budapest, 2001.

Scharnitzky Viktor: Vektorgeometria és lineáris algebra
Nemzeti Tankönyvkiadó, Budapest, 2000.