Az n-dimenziós tér lineáris geometriája

Elméleti matematika

A sorozat kötetei:

Bagyinszky János – György Anna: Diszkrét matematika főiskolásoknak
Bolla Marianna – Krámli András: Statisztikai következtetések elmélete
G. Horváth Ákos – Szirmai Jenő: Nemeuklideszi geometriák modelljei
Járai Antal: Modern alkalmazott analízis
Kiss Emil: Bevezetés az algebrába
Komornik Vilmos: Valós analízis előadások I–II.
Kostrikin, A. I. – Safarevics, I. R.: Algebra
Lovász László: Kombinatorikai problémák és feladatok
Lovász László – Pelikán József – Vesztergombi Katalin: Diszkrét matematika
Praszolov, Viktor Vasziljevics: Lineáris algebra
Stoyan Gisbert: Numerikus matematika mérnököknek és programozóknak
Tóth János – Simon L. Péter: Differenciálegyenletek
Weeks, Jeffrey R.: A tér alakja

Bártfai Pál

AZ n-DIMENZIÓS TÉR LINEÁRIS GEOMETRIÁJA

TYPOTEX

A könyv a Magyar Tudományos Akadémia támogatásával készült.

ISBN 978 963 279 823 3
ISSN 1788-1811

Témakör: *matematika, lineáris algebra*

Kiadja a Typotex Elektronikus Kiadó Kft.
Felelős vezető: Votisky Zsuzsa
A kötetet gondozta: Gerner József
Borítóterv: Szalay Éva
Készítette: B-Group Kft., Pellérd
Felelős vezető: Borbély Zsolt

Tartalomjegyzék

Előszó

A könyv, vagy inkább füzet, bizonyos szempontból rendhagyó. Kinek ajánlható az elolvasása vagy a tanulmányozása? Mindenkinek, akinek lineáris algebrai tanulmányai, vagy oktatómunkája során miért kérdések merülnek fel. Mivel az oktatói tevékenység során, ha az oktatóban nem fogalmazódna meg ilyen kérdés, akkor a hallgatók teszik fel. Állítom, hogy mindenkinek, aki lineáris algebrát oktat, e füzetnek kötelező olvasmányává kell válnia, akkor is, ha ismeri a benne foglaltakat, és még inkább, ha nem.

Voltaképpen, amit lineáris algebra címen a gazdasági szakok oktatnak, az egy tudományágnak csak egy „egydimenziós vetülete", ami a gazdasági élet képleteihez és az operációkutatás módszereihez szükséges. A tárgyon belül maradva csupán árnyékokat üldözünk, a valódi folyamatokat nem látjuk. A valódi jelenségek az n-dimenziós térben játszódnak le: az algebrának geometriai háttere van.

Kívülállók úgy vélhetik, hogy nagyképűség n-dimenziós geometriáról beszélni, hiszen ki tudja azt áttekinteni, felfogni, látni! Ezt a téveszmét igyekszem mindjárt a könyv elején eloszlatni, igyekszem bevezetni az olvasót az n-dimenziós szemléletmódba. Az n-dimenziós kockával való játszadozás nem öncélú, a szemlélet kialakításán túlmenően a későbbiekben sokszor szerepet kap, pl. a parallelepipedon, a gúla, vagy a szimplex esetében. Az első fejezetek könnyebb, meseszerű stílusa fokozatosan vált át egyre komolyabb, nehezebb mondanivalókra, és a végére – azt mondhatjuk – eléggé eldurvul.

A könyv tartalmazza a lineáris algebra szokásos anyagát, a sokéves oktatási tapasztalatot felhasználva tálalja, remélhetőleg élvezetesen, röviden. A geometriai szemléleti mód e köré épül fel, és az anyag nagy részét ez utóbbi szolgáltatja. Ezen túlmenően, egyes alkalmazott irányokba is betekintést nyújt (pl. a vetítés felhasználásai, a lineáris transzformáció normája, a lineáris egyenletrendszer stabilitása, kvadratikus alakok vizsgálata).

Sajátos szerepe van a beépített kérdéseknek. A kérdések általában nem az ún. ellenőrző kérdéseknek felelnek meg, hanem többnyire továbbgondolásra serkentenek. Van, ahol tényleges számítás elvégzésére, annak gyakorlására szólítanak fel, van, ahol további elméleti anyagot tartalmaznak, van, ahol a kérdés anyaga később felhasználásra kerül, sőt egy esetben a válasz egy tételre is vezet. Javasoljuk, hogy a fejezetek végén lévő válaszokat is tüzetesen nézzék át.

Vezérfonalat képezett Dancs István és Puskás Csaba *Vektorterek* című könyve, akkor is, ha nem is hasonlít rá ez a mű. Ez a könyv jóval konkrétabb területre összpontosít, így a mondanivaló is konkrétabb, jól illusztrálható, könnyebben érthető, tényleges számításokkal követhető, és a tételei az n-dimenziós térre vonatkozó elképzelésünkbe jobban beilleszthetők. Valójában nem a térszemléletet növeli, hanem a térre vonatkozó ismereteinket bővíti. Ezúton is szeretnék köszönetet mondani Puskás Csabának, a Corvinus Egyetem docensének, az előbbi könyv társszerzőjének hasznos észrevételeiért és a kézirat átnézéséért.

Az n-dimenziós tér tekintetében itt végig valós n-dimenziós térről lesz szó. A komplex számokat csak ritkán és segédeszközként használjuk, viszont beszélünk a komplex sajátérték és sajátvektor valós térben játszott szerepéről.

Olvasásnál figyeljünk a félkövér betűtípusra. A vektorokat és a mátrixokat mindig félkövérrel jelöljük, és a számjeggyel jelölt vektorok és mátrixok kivételével dőlt jelölést használunk. Az $\boldsymbol{Ax}$ lineáris leképezés jelölésében az argumentumot, az $\boldsymbol{x}$-et gyakran elhagyjuk, ekkor a lineáris leképezés jelölése megegyezik a leképezést létrehozó mátrix jelölésével. Ez bonyodalom helyett inkább egyszerűsítést fog jelenteni. A számok és halmazok (alterek, síkok) jelölésére normál, dőlt betűtípust alkalmazunk. Az n-dimenziós tér jelölése, a szokásostól kissé eltérően, vastagított és álló: $\mathbf{R}^n$, az n és m betűket fenntartjuk a tér dimenziójának a jelölésére.

A könyv elkészítése során igyekeztem olvasmányos, élvezetes művet alkotni. Élvezetessé talán a rácsodálkozás kiváltása teszi: jé, ez is igaz!, vagy a szemléletformálás: ezt másképpen kell elképzelni! Matematikai írásokban ezt a hatást kiváltani más területen sokkal nehezebb, mint itt, a „legkézzelfoghatóbb” témakörben, a véges dimenziós terek tárgyalásánál. Ezek után nem marad más hátra, mint hogy élvezetes, örömteli olvasást kívánjak.

1. fejezet

Az n-dimenziós tér

1.1. Vektor és az n-dimenziós terek fogalma

Mivel a sík vagy a tér pontjait, egy koordináta-rendszer felvétele után, a koordinátáikkal jellemezhetjük, a sík azonosítható a számpárok, a tér a számhármasok halmazával. Az n-dimenziós teret a szám-n-esek halmaza képezi.

Míg az elérhető dimenziókban, tehát $n = 1, 2, 3$ esetén, a vektorokat nyíllal szemléltethetjük, ez a szemléltetés magasabb dimenziókban nem túl meggyőző. Itt is el lehet képzelni az origóból $(3, 2, -1, 5)$-be mutató nyilat, de sok haszna nincs. A vektorokat tehát a tér pontjaival azonosítjuk.

1.1.1. Definíció. Az n-dimenziós vektor n darab valós számból álló rendezett számcsoport, ún. szám-n-es. Jelölése $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$. A vektor megadásában szereplő számokat a vektor koordinátáinak, vagy elemeinek nevezzük. Az n-dimenziós vektorok az n-dimenziós (euklideszi) tér, $\mathbf{R}^n$ pontjai.

Hallgatóim többször megkérdezték már tőlem, hogy látok-e a négydimenziós térben. Előre elkészített válaszom a következő:

– Természetesen! Például a négydimenziós kockának 16 csúcsa van, és 8 darab háromdimenziós kocka határolja.

Harsány derültség fogadja, pedig csak a hallgatók látásmódja hiányos.

Mi is a négydimenziós egységkocka? Azon (a_1, a_2, a_3, a_4) vektorok összessége, melyek minden koordinátája 0 és 1 közé esik, azaz $0 \leq a_i \leq 1$ minden i-re. Melyek ezek közül a csúcsok? Azon vektorok, melyek minden koordinátája 0 vagy 1, és hány ilyen van: 2^4. Melyek a háromdimenziós oldallapok? Minden egyes oldallapnál egy koordináta értéke rögzítetten vagy 0, vagy 1, a többi koordináta tetszőleges 0 és 1 közötti érték. Hány lap van? Ahányféleképpen a koordináta értékét rögzíteni tudom, 4 koordináta közül mindegyiket kétféleképpen választhatom meg, ez 8 lehetőség. Ez tényleg kocka, hiszen rögzítsük, mondjuk, az első koordinátát 0-nak, akkor a $(0, a_2, a_3, a_4)$ alakú vektorok, ahol a_2, a_3 és a_4 értéke 0 és 1 között változik, 3-dimenziós kockát alkotnak. Ha az első koordinátát 1-nek rögzítjük, akkor az előbbi $(1, 0, 0, 0)$ vektorral képezett eltoltját kapjuk. Természetesen van a kockánknak kétdimenziós oldallapja is (négyzet), és egydimenziós éle is.

Felbátorodva, nézzük meg, hogy hogyan „néz ki” az n-dimenziós egységkocka. Az előző okfejtés alapján 2^n csúcsa van, és $2n$ darab $(n-1)$-dimenziós oldallap határolja. Hány k-dimenziós kocka határolja? Most $n-k$ koordinátát kell 0-nak, vagy 1-nek rögzíteni, a többi koordináta szabadon változhat $[0, 1]$-ben. Az $n-k$ rögzítendő koordinátát $\binom{n}{k}$-féleképpen választhatom ki, majd a rögzített értékeket 2^{n-k}-féle módon oszthatom ki, így $2^{n-k}\binom{n}{k}$ darab k-dimenziós kocka határolja.

A binomiális együtthatók közismert módon Pascal-háromszögbe írhatók fel:

$$
\begin{array}{ccccccc}
 & & & 1 & & & \\
 & & 1 & & 1 & & \\
 & 1 & & 2 & & 1 & \\
1 & & 3 & & 3 & & 1
\end{array}
$$

és így tovább, minden szám a felette lévő kettő összege. Az $(n+1)$-edik sor elemei az $\binom{n}{k}$ binomiális együtthatók. Módosítsuk a Pascal-háromszöget: az összeadásnál mindig a jobbra felette álló szám kétszeresét adjuk a bal oldalihoz:

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 2 & & 1 & & & \\
 & & 4 & & 4 & & 1 & & \\
 & 8 & & 12 & & 6 & & 1 & \\
16 & & 32 & & 24 & & 8 & & 1
\end{array}
$$

így az $(n + 1)$-edik sorban megkapjuk az n-dimenziós kocka különböző dimenziós határoló kockáinak a számát. Az „alkatrészek” száma mindig 3^n, beleszámítva magát a vizsgált objektumot is, hiszen $(2 + 1)^n$ binomiális tétel szerinti felbontásában lévő tagokat adjuk össze.

Jó példa arra, hogy sok mindent lehet „látni” a magasabb dimenzióban is, csak más szemléletre, más „látásmódra” van szükség.

Az n-dimenziós tér természetesen több, mint bizonyos objektumok halmaza. Jellemzéséhez további strukturális és geometriai tulajdonságokat fogunk rögzíteni.

1. Kérdés. Mit nevezünk az n-dimenziós kocka testátlójának? Hány testátlója van?

1.2. Műveletek vektorokkal

A két- vagy háromdimenziós vektorok fogalmát a fizikai erővektorok kezelésére alakították ki. A vektorokat nyilakkal ábrázoltuk, a nyíl mutatta az erőhatás irányát. Az ún. parallelogramma-szabály tette lehetővé erők összegzését és felbontását, az erőhatások függetlenségének megfogalmazását. A fizikai erővektorok a támadásvonaluk mentén eltolható vektorok, kimozdítva onnan azonban pl. egy merev testre kifejtett hatása megváltozik. Ezzel ellentétben a matematikai vektorok, ha nyilaknak képzeljük el ezeket, akkor ún. szabad-vektorok, a párhuzamosságot megtartva tetszőlegesen eltolhatók, a koordinátáik nem változnak meg. Ha a tér pontjaival azonosítjuk, akkor az eltolás ilyen formája értelmét veszti.

A gazdasági életben a vektorok másképpen lépnek fel: adott gazdasági jelenségre vonatkozó adathalmaz formájában, így a vektorok nem korlátozódnak két vagy három dimenzióra, sőt általában nagyon sokdimenziós vektorokról van szó. Példaként nézzünk egy raktárkészletet. Leltározásnál felmérjük, hogy melyik árucikkből mennyi van a raktáron, vagyis előállítunk egy $\boldsymbol{k}$ készletvektort (itt feltételezzük, hogy tudjuk, hogy a vektor j-edik komponense milyen árucikkre vonatkozik, azaz az árucikkek meg vannak számozva). Ha újabb szállítmány érkezik, akkor az érkező mennyiség egy $\boldsymbol{s}$ szállítmányvektorral, az új raktárkészlet pedig a $\boldsymbol{k} + \boldsymbol{s}$ vektorral adható meg.

Ha minden árucikkhez a számozásuknak megfelelő sorrendben megadjuk az egységárát, akkor egy $\boldsymbol{\acute{a}}$ árvektort kapunk. Ha minden árura vonatkozóan 10%-os árleszállítást hajtunk végre, az új árvektor $0{,}9\cdot\boldsymbol{\acute{a}}$ lesz.

A változások leírása érdekében tehát a vektorokkal műveleteket kell végezni, ezeknek a műveleteknek az értékét a felhasználhatóság szabja meg. A vektorok összegét és számszorosát ennek megfelelően definiáljuk.

1.2.1. Definíció: Az $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ és a $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ vektorok *összegét* a koordináták szerinti összeadással definiáljuk: $\boldsymbol{a} + \boldsymbol{b} = (a_1+b_1, a_2 + b_2, \ldots, a_n + b_n)$. A *$c$ valós számmal történő szorzást* is koordinátánkénti szorzással definiáljuk: $c\boldsymbol{a} = \boldsymbol{a}c = (ca_1, ca_2, \ldots, ca_n)$. Különböző dimenziójú vektorok nem adhatók össze.

Mivel mindkét műveletet koordinátánként hajtjuk végre, a valós számokra érvényes kommutatív, asszociatív és disztributív műveleti szabályok érvényben maradnak, nevezetesen:

Kommutatív szabály:	$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$	$c\boldsymbol{a} = \boldsymbol{a}c$
Asszociatív szabály:	$\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}) = (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c}$	$c(d\boldsymbol{a}) = (cd)\boldsymbol{a}$
Disztributív szabály:	$c(\boldsymbol{a} + \boldsymbol{b}) = c\boldsymbol{a} + c\boldsymbol{b}$	$(c + d)\boldsymbol{a} = c\boldsymbol{a} + d\boldsymbol{a}$

A 1.2.1. Definícióban definiált műveleteket lineáris műveleteknek nevezzük. Vektorok adott halmazát, ha az elemeire elvégzett lineáris műveletek eredménye is a halmazhoz tartozik, altérnek nevezzük.

Tetszőleges objektumok halmazát, ha elemeire a fenti tulajdonságokkal rendelkező lineáris műveleteket értelmezzük, és az elemekre elvégzett műveletek eredménye is a halmazhoz tartozik, *lineáris térnek* vagy *vektortérnek* nevezzük. Ilyen például az (a, b) intervallumon értelmezett folytonos függvények halmaza, vagy az adott eseménytéren értelmezett valószínűségi változók halmaza. Ezek általánosabb struktúrák, ezekkel itt nem foglalkozunk.

1.2.2. Definíció. $L \subset \mathbf{R}^n$ halmaz *altér*, ha bármely $\boldsymbol{a} \in L$ és $\boldsymbol{b} \in L$ vektorokra és bármely c valós számra $\boldsymbol{a} + \boldsymbol{b} \in L$ és $c\boldsymbol{a} \in L$.

Például az $\mathbf{R}^3$ alterei az origón átmenő síkok vagy egyenesek, esetleg maga az origó, mint egy pontból álló halmaz, vagy a teljes $\mathbf{R}^3$. Az altér mindig tartalmazza a $\mathbf{0}$-t, hiszen $\boldsymbol{a} \in L$ esetén

$$\boldsymbol{a} + (-1) \cdot \boldsymbol{a} = \mathbf{0} \in L.$$

Az előbb említett készletvektorból és árvektorból ki tudjuk számítani a raktárkészlet értékét: a megfelelő koordináták szorzatát kell csupán összeadni. Erre a számításra is bevezetünk egy műveletet, melynek neve a skalárszorzat. Az elnevezés azt mondja számunkra, hogy a vektorok szorzásának az eredménye egyetlen szám.

1.2.3. Definíció. Az $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ és a $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ vektorok *skalárszorzata*

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n.$$

A skalárszorzat kommutatív és disztributív:

Kommutatív szabály:	$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \langle \boldsymbol{b}, \boldsymbol{a} \rangle$
Disztributív szabály:	$\langle \boldsymbol{a}, \boldsymbol{b} + \boldsymbol{c} \rangle = \langle \boldsymbol{a}, \boldsymbol{b} \rangle + \langle \boldsymbol{a}, \boldsymbol{c} \rangle$

A konstanssal történő szorzásra asszociatív: $c\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \langle c\boldsymbol{a}, \boldsymbol{b} \rangle = \langle \boldsymbol{a}, c\boldsymbol{b} \rangle$, de különben nem asszociatív: $\langle \boldsymbol{a}, \boldsymbol{b} \rangle \boldsymbol{c}$ a $\boldsymbol{c}$ vektor többszöröse, míg $\boldsymbol{a}\langle \boldsymbol{b}, \boldsymbol{c} \rangle$ az $\boldsymbol{a}$ vektor többszöröse, és a kettő különbözhet.

Egyedül a disztributív szabály érvényessége szorul bizonyításra:

$$\langle \boldsymbol{a}, \boldsymbol{b} + \boldsymbol{c} \rangle = \sum_{i=1}^{n} a_i(b_i + c_i) = \sum_{i=1}^{n} (a_i b_i + a_i c_i) = \sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} a_i c_i =$$
$$= \langle \boldsymbol{a}, \boldsymbol{b} \rangle + \langle \boldsymbol{a}, \boldsymbol{c} \rangle.$$

2. Kérdés. n darab számszerű statisztikai megfigyelés, $a_1, a_2, \dots, a_n$, az n-dimenziós tér egy pontjával azonosítható. Jelöljük $\bar{a}$-sal a minta átlagát, azaz legyen $\bar{a} = \frac{a_1+a_2+\ldots+a_n}{n}$, és képezzük az $\boldsymbol{x} = (a_1 - \bar{a}, a_2 - \bar{a}, \ ..., a_n - \bar{a})$ vektorokat. Igazoljuk, hogy ezek a vektorok az $\mathbf{R}^n$ egy valódi alterébe esnek. Mi jellemzi ezt az alteret?

1.3. Távolság és szög (norma és ortogonalitás)

Két pont távolságánál, a vektor hosszánál és két vektor szögénél az n-dimenziós térben nem tudjuk megtenni azt, hogy odamegyünk, és mérőeszközzel lemérjük. Mégis, ha az n-dimenziós tér geometriájáról akarunk beszélni, ezek megkerülhetetlen fogalmak.

Egy hallgatóm azt javasolta, hogy fektessünk át egy kétdimenziós síkot a vektoron vagy vektorokon, és az (x, y)-koordinátasíkkal képezett metszésvonala körül forgassuk be az alapsíkba. Ekkor sétáljunk oda, és mérjük le.

A gondolat – bár logikusan hangzik – de több sebből vérzik. Először is mi az, hogy forgatás? Speciális távolságtartó transzformáció. Akkor viszont ne akarjuk a távolságot önmaga felhasználásával definiálni, ezt mindenképpen önállóan meg kell adni, és a tér geometriáját erre a definícióra kell felépíteni.

A definíció két alapkövetelménye, hogy egyrészt két és három dimenzióban a szokásos távolságfogalmat kapjuk vissza, másrészt a kiterjesztés megfeleljen a matematikában a távolságtól elvárt tulajdonságoknak. A vektor $\mathbf{R}^n$-ben definiált hosszát, jelezve, hogy ez általánosított hosszúság, a vektor normájának fogjuk nevezni, és $||\boldsymbol{a}||$-val jelöljük (szemben a két- és háromdimenziós $|\boldsymbol{a}|$ jelöléssel).

1.3.1. Definíció. Az $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$ vektor *normája*: $||\boldsymbol{a}|| = \sqrt{a_1^2 + a_2^2 + ... + a_n^2}$. Két $\mathbf{R}^n$-beli pont, $\boldsymbol{a}$ és $\boldsymbol{b}$ távolsága $d(\boldsymbol{a}, \boldsymbol{b}) = ||\boldsymbol{a} - \boldsymbol{b}||$.

Nyilvánvaló módon a képlet $n = 2$ vagy 3 esetén a vektor hosszának Pythagorasz-tétellel kiszámított értékét adja meg.

3. Kérdés. Ha az $a_1, a_2, \dots, a_n$ statisztikai mintából elkészítjük az $\boldsymbol{x} = (a_1 - \bar{a}, a_2 - \bar{a}, ..., a_n - \bar{a})$ vektort, ahol $\bar{a}$ a mintaátlagot jelöli, akkor a minta szórása és a $\boldsymbol{x}$ normája között milyen összefüggés áll fenn?

Vegyük észre, hogy a norma kifejezhető a skalárszorzattal: $||\boldsymbol{a}||^2 = \langle \boldsymbol{a}, \boldsymbol{a} \rangle$.

A távolságfogalomtól a következő tulajdonságok teljesülését várjuk el:

1. $d(\boldsymbol{a},\ \boldsymbol{b}) \geq 0$, és $d(\boldsymbol{a},\ \boldsymbol{b}) = 0$ akkor és csak akkor teljesül, ha $\boldsymbol{a} = \boldsymbol{b}$.
2. $d(\boldsymbol{a},\ \boldsymbol{b}) = d(\boldsymbol{b},\ \boldsymbol{a})$.
3. $d(\boldsymbol{a},\ \boldsymbol{b}) \leq d(\boldsymbol{a},\ \boldsymbol{c}) + d(\boldsymbol{c},\ \boldsymbol{b})$.

A harmadik tulajdonságot háromszög-egyenlőtlenségnek hívjuk.

A normával megadott távolság az első két tulajdonságot láthatóan teljesíti, a harmadikat rövidesen bebizonyítjuk.

1.3.2. Tétel (Cauchy–Schwarz-egyenlőtlenség)

$$|\langle \boldsymbol{a}, \boldsymbol{b} \rangle| \leq ||\boldsymbol{a}|| \cdot ||\boldsymbol{b}||$$

Bizonyítás. Számítsuk ki a nemnegatív $||\boldsymbol{a} - \lambda\ \boldsymbol{b}||^2$ kifejezést:

$$\begin{aligned} &||\boldsymbol{a} - \lambda \boldsymbol{b}||^2 = \\ &= \langle \boldsymbol{a} - \lambda \boldsymbol{b}, \boldsymbol{a} - \lambda \boldsymbol{b} \rangle = \langle \boldsymbol{a}, \boldsymbol{a} \rangle - \langle \lambda \boldsymbol{b}, \boldsymbol{a} \rangle - \langle \boldsymbol{a}, \lambda \boldsymbol{b} \rangle + \lambda^2 \langle \boldsymbol{b}, \boldsymbol{b} \rangle = \\ &= ||\boldsymbol{a}||^2 - 2\lambda \langle \boldsymbol{a}, \boldsymbol{b} \rangle + \lambda^2 ||\boldsymbol{b}||^2 \geq 0. \end{aligned} \qquad (*)$$

A jobb oldali, a λ változót tekintve másodfokú függvény (parabola) csak úgy lehet nemnegatív, ha a tengelyt nem metszi, vagyis ha nincs két gyöke. Ekkor diszkriminánsa negatív vagy nulla, tehát $(2\langle \boldsymbol{a},\ \boldsymbol{b} \rangle)^2 - 4||\boldsymbol{a}||^2||\boldsymbol{b}||^2 \leq 0$, ami ekvivalens az állítással. $\square$

1.3.3. Tétel (háromszög-egyenlőtlenség).

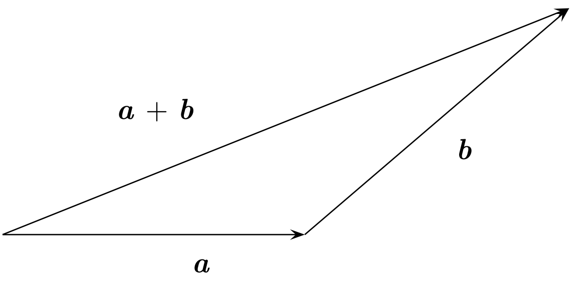
$$||\boldsymbol{a} + \boldsymbol{b}|| \leq ||\boldsymbol{a}|| + ||\boldsymbol{b}||.$$

Bizonyítás. Mivel mindkét oldal nemnegatív, az egyenlőtlenség akkor és csak akkor áll fenn, ha négyzetre emelve igaz. $||\boldsymbol{a} + \boldsymbol{b}||$ négyzete a (*) képletsorból $\lambda = -1$ helyettesítéssel megkapható, tehát csak az

$$||\boldsymbol{a}||^2 + 2\langle \boldsymbol{a}, \boldsymbol{b} \rangle + ||\boldsymbol{b}||^2 \leq ||\boldsymbol{a}||^2 + 2||\boldsymbol{a}|| \cdot ||\boldsymbol{b}|| + ||\boldsymbol{b}||^2$$

egyenlőtlenséget kell igazolni. Ez azonban a Cauchy–Schwarz-egyenlőtlenség miatt igaz. $\square$

A háromszög-egyenlőtlenség az alábbi vektorábrának megfelelően úgy interpretálható, hogy egy háromszögben két oldal összege nagyobb, mint a harmadik oldal.



A távolságra vonatkozó háromszög egyenlőtlenség $\boldsymbol{a} = \boldsymbol{x} - \boldsymbol{z}$, $\boldsymbol{b} = \boldsymbol{z} - \boldsymbol{y}$ helyettesítéssel adódik:

$$d(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}|| = ||\boldsymbol{a} + \boldsymbol{b}|| \leq ||\boldsymbol{a}|| + ||\boldsymbol{b}|| =$$
$$= ||\boldsymbol{x} - \boldsymbol{z}|| + ||\boldsymbol{z} - \boldsymbol{y}|| = d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y}).$$

4. Kérdés. Ha $\boldsymbol{a} \in \mathbf{R}^4$ és $||\boldsymbol{a}|| = 4$, továbbá $\mathbf{1} = (1, 1, 1, 1)$, akkor mennyi az $\langle \boldsymbol{a} - \mathbf{1}, \boldsymbol{a} + \mathbf{1} \rangle$ kifejezés értéke?

Két vektor szögének meghatározásánál már a bevezetőben említett gondolat alkalmazható. A forgatásról egyelőre felejtkezzünk el, beszéljünk inkább távolságtartó transzformációról. Később be fogjuk látni, hogy van olyan távolságtartó transzformáció, amely két adott vektort az (x, y)-koordinátasíkban lévő vektorokba visz át. Két vektor szögét továbbra is definiálni kell, a definíció alapgondolata az, hogy az n-dimenziós terek távolságtartó transzformációja egyben szögtartó is. Ez a tulajdonság a két és háromdimenziós terekben teljesül, így a megalkotott szögfogalom az $\mathbf{R}^2$-ben használtnak a kiterjesztése lesz.

Az (x, y)-koordinátasíkra transzformált háromszögben két oldal szöge a koszinusz-tétellel kiszámítható. A háromszög oldalai, a távolságtartó transzformáció miatt $||\boldsymbol{a}||$, $||\boldsymbol{b}||$ és $||\boldsymbol{a} - \boldsymbol{b}||$, tehát

$$||\boldsymbol{a} - \boldsymbol{b}||^2 = ||\boldsymbol{a}||^2 + ||\boldsymbol{b}||^2 - 2||\boldsymbol{a}|| \cdot ||\boldsymbol{b}|| \cos \gamma \ .$$

Ugyanakkor a (*) összefüggésből $\lambda \ = 1$ helyettesítéssel

$$||\boldsymbol{a} - \boldsymbol{b}||^2 = ||\boldsymbol{a}||^2 + ||\boldsymbol{b}||^2 - 2\langle \boldsymbol{a}, \boldsymbol{b} \rangle,$$

és a kettő összevetéséből

$$\cos \gamma = \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{||\boldsymbol{a}|| \cdot ||\boldsymbol{b}||},$$

ha $\boldsymbol{a} \neq \mathbf{0}$ és $\boldsymbol{b} \neq \mathbf{0}$ (ahol $\mathbf{0} = (0, 0, 0, \ldots, 0)$ a nullvektort jelöli).

1.3.4. Definíció. Ha $\boldsymbol{a} \neq \boldsymbol{0}$ és $\boldsymbol{b} \neq \boldsymbol{0}$, akkor a vektorok által bezárt γ *szöget* a

$$\cos\gamma = \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{||\boldsymbol{a}|| \cdot ||\boldsymbol{b}||}$$

képlet határozza meg. Ha $\boldsymbol{a} = \boldsymbol{0}$ vagy $\boldsymbol{b} = \boldsymbol{0}$, akkor a bezárt szöget nem értelmezzük.

5. Kérdés. Legyen $a_1, a_2, \ldots, a_n$ és $b_1, b_2, \ldots, b_n$ két statisztikai minta, vonjuk le a mintaelemekből az átlagukat, a kapott vektorokat jelöljük $\boldsymbol{x}$-szel ill. $\boldsymbol{y}$-nal. Mi az összefüggés a $\boldsymbol{x}$ és az $\boldsymbol{y}$ korrelációs együtthatója és a szöge között?

Speciálisan, ha $\boldsymbol{a}$ és $\boldsymbol{b}$ merőleges, vagyis $\gamma = 90°$, $\cos\gamma = 0$, akkor ez csak úgy lehet, ha $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 0$. Megfordítva: ha $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 0$, akkor a két vektor vagy merőleges, vagy valamelyik nullvektor.

1.3.5. Definíció. Az $\boldsymbol{a}$ és a $\boldsymbol{b}$ vektorokat *ortogonálisaknak* nevezzük, ha $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 0$.

Ahogy a norma az általánosított távolság, az ortogonalitás az általánosított merőlegesség, kiegészítve azzal, hogy a nullvektor minden vektorra ortogonális.

6. Kérdés. Az n-dimenziós kockának vannak-e ortogonális testátlói?

1.4. Lineáris függetlenség

A könnyebb megértés kedvéért először a lineáris összefüggést mutatjuk be. A vektorok egy véges rendszere lineárisan összefüggő, ha valamelyik vektor lineáris egyenlettel kifejezhető a többi segítségével. Nézzünk egy példát. Legyen $\boldsymbol{a} = (1, 2, 1, 3)$, $\boldsymbol{b} = (1, 1, 1, 1)$, $\boldsymbol{c} = (-1, 2, -2, 4)$, $\boldsymbol{d} = (3, 2, 4, 2)$, akkor mivel fennáll, hogy $\boldsymbol{d} = 2\boldsymbol{a} - \boldsymbol{c}$, a vektorrendszer lineárisan összefüggő vektorokból áll.

A lineáris kifejezés „konstans vektort” nem tartalmazhat. Például, ha $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{1}$, akkor $\boldsymbol{a}$ és $\boldsymbol{b}$ még lineárisan független lehet: példa erre az $\boldsymbol{a} = (2, -1)$ és a $\boldsymbol{b} = (-1, 2)$. Az $\boldsymbol{1}$ vektor itt „idegen” vektornak számít, az $\boldsymbol{a}$, $\boldsymbol{b}$ és $\boldsymbol{1}$ három vektorból álló rendszer természetesen már lineárisan összefüggő.

Rendezzük a fennálló összefüggést 0-ra. Ha $\boldsymbol{d} = 2\boldsymbol{a} - \boldsymbol{c}$, akkor $2\boldsymbol{a} - \boldsymbol{c} - \boldsymbol{d} = \boldsymbol{0}$. A jobb oldalon álló $\boldsymbol{0}$ a nullvektort jelöli, melynek minden koordinátája 0: $\boldsymbol{0} = (0, 0, 0, 0)$. Részletesebben felírva: $2 \cdot \boldsymbol{a} + 0 \cdot \boldsymbol{b} + (-1) \cdot \boldsymbol{c} + (-1) \cdot \boldsymbol{d} = \boldsymbol{0}$. Akkor mondjuk, hogy az $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ és $\boldsymbol{d}$ vektorok lineárisan összefüggőek, ha fennáll a $c_1\boldsymbol{a} + c_2\boldsymbol{b} + c_3\boldsymbol{c} + c_4\boldsymbol{d} = \boldsymbol{0}$ összefüggés olyan c_1, c_2, c_3 és c_4 számokra, amelyek nem mind nullák. Amelyik vektor együtthatója nem nulla, az kifejezhető a többivel. Ha a c_1, c_2, c_3, és c_4 számok mind nullák, akkor egyik sem fejezhető ki, akkor az egyenlet nem jelent lineáris összefüggést.

Ha egy vektorrendszer egyik eleme nullvektor, akkor a rendszer mindig lineárisan összefüggő (még akkor is, ha egyelemű, csak a **0**-ból álló vektorrendszerről van szó). Ha pl. az $\boldsymbol{a}$ nullvektor, akkor a lineáris összefüggés: $\boldsymbol{a} = \mathbf{0}$, tehát $c_1 = 1$, a többi együttható nulla, de nem minden együttható nulla.

Definíciónak a gyakrabban használt ellenkező esetet mondjuk ki: ha nem lineárisan összefüggők a vektorok, akkor lineárisan függetleneknek nevezzük. Ekkor a lineáris összefüggés létét tagadni kell.

1.4.1. Definíció. Az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_k$ vektorok ($\boldsymbol{a}_i \in \mathbf{R}^n$) *lineárisan függetlenek*, ha a

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_k\boldsymbol{a}_k = 0$$

összefüggés *csak úgy* teljesülhet, ha $c_1 = c_2 = ... = c_k = 0$. A $c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + + ... + c_k\boldsymbol{a}_k$ kifejezést az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_k$ vektorok *lineáris kombinációjának* nevezzük.

A lineáris függetlenség és az ortogonalitás kapcsolatát mutatja be a következő állítás.

1.4.2. Tétel. Ha $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_k$ vektorok lineárisan függetlenek, és $\boldsymbol{b} \neq \mathbf{0}$ ortogonális az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_k$ vektorokra, akkor az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_k$, $\boldsymbol{b}$ vektorrendszer is lineárisan független.

Bizonyítás. Tegyük fel, hogy teljesül a $c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_k\boldsymbol{a}_k + d\boldsymbol{b} = \mathbf{0}$ összefüggés, és szorozzuk meg mindkét oldalát skalárisan a $\boldsymbol{b}$ vektorral, akkor az ortogonalitás miatt $d||\boldsymbol{b}||^2 = 0$ egyenletet kapjuk, amiből $\boldsymbol{b} \neq \mathbf{0}$ miatt $d = 0$ adódik. Visszatérve a kiinduláshoz, $c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_k\boldsymbol{a}_k = \mathbf{0}$, ami a lineáris függetlenség miatt csak $c_1 = c_2 = ... = c_k = 0$ esetében lehetséges. A kiindulási reláció tehát csak úgy teljesülhet, ha az összes együttható nulla, vagyis a vektorrendszer lineárisan független. □

1.4.3. Következmény. Ha az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_k$ vektorok egyike sem nullvektor és bármely kettő ortogonális, akkor lineárisan függetlenek.

Bizonyítás. Ha a második vektor ortogonális az elsőre, akkor az 1.4.2. Tétel miatt lineárisan függetlenek. Ha az első kettő lineárisan független, és a harmadik ortogonális rájuk, akkor a három vektor is lineárisan független. A lineárisan független rendszer így lépésenként bővíthető. □

7. Kérdés. Igaz-e a vektorrendszer lineáris függetlensége akkor, ha bármely két vektor szöge $60°$?

Többször, és sokoldalúan használt tétel következik, melyet több fogalom megalapozásához felhasználunk. Lényege a következő: Ha adott néhány vektor, és van ennél több lineárisan független vektor, akkor a térben (altérben) van olyan vektor is, amelyik nem fejezhető ki a megadottak lineáris kombinációjaként, sőt ami ennél több, még merőleges is rájuk. Ha a megadott vektorok lineárisan függetlenek, akkor a tétel szerint a lineárisan független rendszer a maximumig bővíthető (kibővítési tétel néven ismert állítás). A merőlegességi állítás miatt ortogonális rendszerek konstrukciójához is felhasználható (Gram–Schmidt-féle ortogonalizálási eljárás).

1.4.4. Tétel. Legyen $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, ..., $\boldsymbol{x}_k$ tetszőleges és $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ..., $\boldsymbol{a}_{k+1}$ lineárisan független vektorrendszer, akkor $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ..., $\boldsymbol{a}_{k+1}$ lineáris kombinációjaként előállítható olyan $\boldsymbol{y} \neq \boldsymbol{0}$ vektor mely ortogonális az $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, ..., $\boldsymbol{x}_k$ vektorokra.

Bizonyítás. Első lépésben megadunk egy $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, ..., $\boldsymbol{b}_k$ vektorrendszert, melynek minden eleme ortogonális $\boldsymbol{x}_1$-re. Ha az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ..., $\boldsymbol{a}_{k+1}$ rendszerben vannak $\boldsymbol{x}_1$-re ortogonális vektorok, akkor átszámozással hozzuk ezeket előre. Legyen

$$\boldsymbol{b}_i = \boldsymbol{a}_i + \alpha_i \boldsymbol{a}_{i+1} \quad (i = 1, 2, \ldots, k),$$

ahol $\alpha_i = -\frac{\langle \boldsymbol{a}_i, \boldsymbol{x}_1\rangle}{\langle \boldsymbol{a}_{i+1}, \boldsymbol{x}_1\rangle}$, ha $\langle \boldsymbol{a}_{i+1}, \boldsymbol{x}_1\rangle \neq 0$, és $\alpha_i = 0$, ha $\boldsymbol{a}_{i+1}$ ortogonális $\boldsymbol{x}_1$-re.

A kapott vektorok mind ortogonálisak $\boldsymbol{x}_1$-re, mert

$$\langle \boldsymbol{b}_i, \boldsymbol{x}_1\rangle = \langle \boldsymbol{a}_i, \boldsymbol{x}_1\rangle - \frac{\langle \boldsymbol{a}_i, \boldsymbol{x}_1\rangle}{\langle \boldsymbol{a}_{i+1}, \boldsymbol{x}_1\rangle}\langle \boldsymbol{a}_{i+1}, \boldsymbol{x}_1\rangle = 0,$$

ha $\langle \boldsymbol{a}_{i+1}, \boldsymbol{x}_1\rangle \neq 0$, és ha $\langle \boldsymbol{a}_{i+1}, \boldsymbol{x}_1\rangle = 0$, akkor az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ..., $\boldsymbol{a}_{k+1}$ rendszer átrendezése miatt $\langle \boldsymbol{a}_i, \boldsymbol{x}_1\rangle = 0$ és $\boldsymbol{b}_i = \boldsymbol{a}_i$.

A kapott $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, ..., $\boldsymbol{b}_k$ rendszer lineárisan független. Jobban látjuk, ha $k = 3$ esetén részletesen felírjuk, tetszőleges k-ra ugyanígy megy. Vegyük a rendszer egy lineáris kombinációját:

$$c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + c_3\boldsymbol{b}_3 = c_1(\boldsymbol{a}_1 + \alpha_1\boldsymbol{a}_2) + c_2(\boldsymbol{a}_2 + \alpha_2\boldsymbol{a}_3) + c_3(\boldsymbol{a}_3 + \alpha_3\boldsymbol{a}_4) =$$

$$= c_1\boldsymbol{a}_1 + (c_1\alpha_1 + c_2)\boldsymbol{a}_2 + (c_2\alpha_2 + c_3)\boldsymbol{a}_3 + c_3\alpha_3\boldsymbol{a}_4.$$

Az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$, $\boldsymbol{a}_4$ vektorok lineáris függetlenségéből következik, hogy a fenti kifejezés csak úgy lehet nulla, ha $c_1 = 0$, $c_1\alpha_1 + c_2 = 0$, $c_2\alpha_2 + c_3 = = 0$, $c_3\alpha_3 = 0$. Ez pedig csak úgy lehetséges, ha $c_1 = 0$, $c_2 = 0$, $c_3 = 0$, ami a $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$ vektorok lineáris függetlenségét jelenti.

A második lépésben a $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, ..., $\boldsymbol{b}_k$ vektorokból ugyanígy állítsunk elő $\boldsymbol{x}_2$-re ortogonális $\boldsymbol{c}_1$, $\boldsymbol{c}_2$, ..., $\boldsymbol{c}_{k-1}$ vektorokat. Ezek is lineárisan függetlenek lesznek és $\boldsymbol{x}_1$-re is ortogonálisok, mert $\boldsymbol{x}_1$-re ortogonális vektorok lineáris kombinációi.

Az eljárást folytatva végül egy $\boldsymbol{y}$ vektort kapunk, ami már minden $\boldsymbol{x}_i$ vektorra ortogonális, ugyanakkor $\boldsymbol{y} = \mathbf{0}$ nem lehet, mert lineárisan független vektorok lineáris kombinációjaként áll elő. □

1.4.5. Következmény. $\mathbf{R}^n$-ben legfeljebb n elemű lineárisan független vektorrendszer létezik.

Bizonyítás. n elemű mindenesetre létezik, hiszen a koordináta-egységvektorok,

$$\boldsymbol{e}_1 = (1,\, 0,\, 0,\, 0,\, \dots,\, 0),$$

$$\boldsymbol{e}_2 = (0,\, 1,\, 0,\, 0,\, \dots,\, 0),$$

$$\boldsymbol{e}_3 = (0,\, 0,\, 1,\, 0,\, \dots,\, 0),$$

$$\boldsymbol{e}_4 = (0,\, 0,\, 0,\, 1,\, \dots,\, 0),$$

$$\vdots$$

$$\boldsymbol{e}_n = (0,\, 0,\, 0,\, 0,\, \dots,\, 1),$$

páronként ortogonálisak, tehát lineárisan függetlenek.

Ha létezne az $\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_{n+1}$ $n+1$ elemű lineárisan független vektorrendszer $\mathbf{R}^n$-ben, akkor a 4. Tétel miatt létezne $\boldsymbol{y} = (y_1,\, y_2,\, \dots, y_n) \neq$ $\neq \mathbf{0}$ mindegyikre ortogonális vekor. Erre a vektorra számoljuk ki a skalárszorzatokat: $\langle \boldsymbol{e}_k,\ \boldsymbol{y}\rangle = y_k = 0$, vagyis $\boldsymbol{y}$ minden koordinátája nulla, ami ellentmondás. □

8. Kérdés. Lehetséges-e, hogy az n-dimenziós kocka összes testátlóvektora lineárisan független rendszert alkot?

1.5. Altér, dimenzió

Az 1.2. fejezetből idézzük fel az altér definícióját.

1.2.2. Definíció. $L \subset \mathbf{R}^n$ halmaz *altér*, ha bármely $\boldsymbol{a} \in L$ és $\boldsymbol{b} \in L$ vektorokra és bármely c valós számra $\boldsymbol{a} + \boldsymbol{b} \in L$ és $c\boldsymbol{a} \in L$.

9. Kérdés. Az $(x_1,\, x_2,\, x_3)$ vektor legyen az

$$\begin{aligned} a_1x_1 + b_1x_2 + c_1x_3 &= 0\\ a_2x_1 + b_2x_2 + c_2x_3 &= 0\\ a_3x_1 + b_3x_2 + c_3x_3 &= 0 \end{aligned}$$

ún. homogén lineáris egyenletrendszer megoldása. Igaz-e, hogy a megoldások halmaza $\mathbf{R}^3$ altere? Igaz-e ugyanez, ha az egyenletrendszer nem homogén, azaz a jobb oldalon nem csupa nulla áll?

1.5.1. Definíció. Az L altér *dimenzióján* az L-be tartozó lineárisan független vektorrendszerek maximális elemszámát értjük.

Az L dimenziója véges, hiszen az 1.4.5. Következmény alapján maximálisan n lehet.

1.5.2. Következmény (kibővítési tétel). Az L-ben felvett tetszőleges lineárisan független vektorrendszer kibővíthető maximális elemszámúvá.

Bizonyítás. Ha a felvett vektorrendszer az L dimenziójánál kisebb elemszámú, akkor van olyan $\boldsymbol{y} \neq \mathbf{0}$ vektor, hogy $\boldsymbol{y}$ a rendszer minden elemére ortogonális. Az 1.4.2. Tétel miatt ez a rendszer kibővítését jelenti. Az eljárás, ha szükséges, akkor ismételhető. □

1.5.3. Következmény (Gram–Schmidt-féle ortogonalizálás). A k-dimenziós L altérben felvehető k darab egymásra ortogonális vektor.

Bizonyítás. A dimenzió definíciója értelmében a k-dimenziós L-ben felvehető k elemű, $\boldsymbol{a}_1$, $\boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ lineárisan független vektorrendszer. A 1.4.4. Tétel konstrukciójával páronként ortogonális k elemű rendszer is megalkotható. $\boldsymbol{a}_1$-hez készítsük el az erre ortogonális $\boldsymbol{b}_1$, $\boldsymbol{b}_2, \ldots, \boldsymbol{b}_{k-1}$ rendszert, majd $\boldsymbol{a}_1$ és $\boldsymbol{b}_1$-hez készítsük el az ezekre ortogonális $\boldsymbol{c}_1$, $\boldsymbol{c}_2, \ldots, \boldsymbol{c}_{k-2}$ rendszert, ezután $\boldsymbol{a}_1$, $\boldsymbol{b}_1$ és $\boldsymbol{c}_1$-gyel folytassuk az eljárást. □

Viszonylag könnyű megadni azt a legszűkebb, L alteret, amely bizonyos, adott vektorokat tartalmaz. Legyenek megadva tetszőlegesen az $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\ldots$, $\boldsymbol{x}_k$ n-dimenziós vektorok. Mivel az L altér tartalmazza ezeket a vektorokat, tartalmazza ezek lineáris kombinációit is. Másrészt az összes lineáris kombináció halmaza zárt a lineáris vektorműveletekre nézve, tehát alteret alkot.

1.5.4. Definíció. Az $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ vektorokat tartalmazó legszűkebb alteret, ami egyben a vektorokból képezhető lineáris kombinációk halmaza, az $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ vektorok által *generált altérnek* nevezzük.

1.5.5. Definíció. Az $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ vektorrendszerben található lineárisan független vektorok maximális számát a vektorrendszer *rangjának* nevezzük.

1.5.6. Tétel. Az $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ vektorok által generált altér *dimenziója* megegyezik a generáló vektorrendszer rangjával. Az altérben lévő, lineárisan független vektorok maximális elemszámú rendszere mindig generáló rendszer.

Bizonyítás. Ha az $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ vektorok közül legfeljebb j lineárisan független ($j < k$), legyenek ezek $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_j$, akkor $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_{j+1}$ már nem lineárisan független, így $c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + ... + c_{j+1}\boldsymbol{x}_{j+1} = \mathbf{0}$ teljesül nem csupa nulla szorzóval. c_{j+1} azonban nem lehet nulla, mert $c_1\boldsymbol{x}_1 + +c_2\boldsymbol{x}_2 + \cdots + c_j\boldsymbol{x}_j = \mathbf{0}$ csak 0 együtthatókkal teljesülhet. Ha $c_{j+1} \neq 0$, akkor $\boldsymbol{x}_{j+1}$ és ugyanígy a további vektorok is, már kifejezhetők az $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_j$ vektorokkal, és a lineáris kombinációkból ezek a vektorok kiküszöbölhetők. Így $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_j$ vektorok is generálják L-et.

Az $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_j$ rendszer nem bővíthető, hiszen minden $\boldsymbol{x} \in L$-re $\boldsymbol{x} = \sum_{i=1}^{j} c_i\boldsymbol{x}_i$, ami lineáris összefüggést jelent az $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_j, \boldsymbol{x}$ vektorok között. Ha nem bővíthető, akkor az 1.5.2. (kibővítési) Tétel miatt L-ben sem lehet $j+1$ elemű lineárisan független vektorrendszer.

Legyen L j dimenziós, és $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_j$ maximális elemszámú lineárisan független rendszer. Tetszőleges $\boldsymbol{x} \in L$-et hozzávéve a rendszer összefüggővé válik, azaz

$$c_1\boldsymbol{y}_1 + c_2\boldsymbol{y}_2 + ... + c_j\boldsymbol{y}_j + c_{j+1}\boldsymbol{x} = 0$$

teljesül nem csupa nulla együtthatóval. A $c_{j+1} \neq 0$, mert $c_1\boldsymbol{y}_1 + c_2\boldsymbol{y}_2 + \cdots + +c_j\boldsymbol{y}_j = \mathbf{0}$ nem teljesülhet, tehát az egyenletből $\boldsymbol{x}$ kifejezhető, vagyis bármely $\boldsymbol{x} \in L$ benne van az $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_j$ által generált altérben. □

10. Kérdés. Az $x_1 + x_2 + ... + x_n = 0$ egyenletnek eleget tevő $\boldsymbol{x} = = (x_1, x_2, ..., x_n)$ vektorok a 9. Kérdés szerint alteret alkotnak. Generálják-e ezt az alteret az

$$\boldsymbol{a}_1 = (1,\ 0,\ 0,\ ...,\ 0,\ -1),$$

$$\boldsymbol{a}_2 = (0,\ 1,\ 0,\ ...,\ 0,\ -1),$$

$$\boldsymbol{a}_3 = (0,\ 0,\ 1,\ ...,\ 0,\ -1),$$

$$\ldots$$

$$\boldsymbol{a}_{n-1} = (0,\ 0,\ 0,\ ...,\ 1,\ -1)$$

vektorok? Mennyi az altér dimenziója?

1.6. Bázis

Mit nevezünk az n-dimenziós térben derékszögű koordináta-rendszernek? n darab egységvektorból álló vektorrendszert, melyek páronként merőlegesek. Például:

$$\boldsymbol{e}_1 = (1,\ 0,\ 0,\ 0,\ \ldots,\ 0),$$

$$\boldsymbol{e}_2 = (0,\ 1,\ 0,\ 0,\ \ldots,\ 0),$$

$$\boldsymbol{e}_3 = (0, 0, 1, 0, \ldots, 0),$$

$$\boldsymbol{e}_4 = (0, 0, 0, 1, \ldots, 0),$$

$$\vdots$$

$$\boldsymbol{e}_n = (0, 0, 0, 0, \ldots, 1).$$

Ezen vektorok segítségével bármely $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ vektor előállítható $\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i$ alakban. Az $x_1, x_2, ..., x_n$ számokat az $\boldsymbol{x}$ vektor $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, …, $\boldsymbol{e}_n$ koordináta-rendszerre vonatkozó koordinátáinak nevezzük. Természetesen $\mathbf{R}^n$-ben felvehető más koordináta-rendszer is, ekkor az $\boldsymbol{x}$ koordinátái megváltoznak.

Lényeges-e, hogy koordinátavektoroknak egységvektorokat vegyünk? Nem lényeges. Már két dimenzióban is sokszor torzítva kell ábrázolnunk a függvényt, hogy „ráférjen a papírra”.

Lényeges-e, hogy ortogonális rendszer legyen? Nem lényeges, de kényelmesebb. Meglehetősen nehéz például távolságot számolni.

Elvileg tehát létezik „ferdeszögű koordináta-rendszer” is. Ez tetszőleges vektorrendszer, amitől megköveteljük, hogy:

1. Minden $\mathbf{R}^n$-be (ill. L-be) eső vektor előállítható legyen a koordinátavektorok lineáris kombinációjaként.

2. Ez az előállítás egyértelmű legyen, így minden pontnak egyértelműen megfeleltethetők legyenek az adott koordináta-rendszerre vonatkozó koordinátái.

Az első tulajdonság azt jelenti, hogy a koordinátavektorok generálják a teret (alteret), a második tulajdonság pedig akkor és csak akkor teljesül, ha a vektorrendszer lineárisan független.

1.6.1. Tétel. Az $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, …, $\boldsymbol{x}_k$ vektorok által generált L altérben a következő három tulajdonság ekvivalens:

a) Van olyan $\boldsymbol{x} \in L$, mely egyértelműen állítható elő a generáló vektorok lineáris kombinációjaként.

b) Bármely $\boldsymbol{x} \in L$ egyértelműen állítható elő a generáló vektorok lineáris kombinációjaként.

c) A generáló rendszer lineárisan független elemekből áll.

Bizonyítás. c) $\Rightarrow$ b) Tegyük fel, hogy a rendszer lineárisan független és van olyan $\boldsymbol{x} \in L$, hogy $\boldsymbol{x}$ előállítása nem egyértelmű, azaz $\boldsymbol{x} =$

$= \sum_{i=1}^{k} a_i \boldsymbol{x}_i$, és $\boldsymbol{x} = \sum_{i=1}^{k} b_i \boldsymbol{x}_i$. A két összefüggést vonjuk ki egymásból, akkor $0 = \sum_{i=1}^{k} (a_i - b_i)\boldsymbol{x}_i$, de ez csak csupa nulla együtthatókkal teljesülhet, tehát minden i-re $a_i = b_i$, így a két előállítás csak azonos lehet.

b) $\Rightarrow$ a) Az a) állítás speciális esete b)-nek.

a) $\Rightarrow$ c) Tegyük fel, hogy $\boldsymbol{x} \in L$ előállítása egyértelmű, azaz $\boldsymbol{x} = \sum_{i=1}^{k} a_i \boldsymbol{x}_i$, de a rendszer nem lineárisan független, vagyis nem csupa nulla együtthatókkal $0 = \sum_{i=1}^{k} c_i \boldsymbol{x}_i$. Adjuk össze a két összefüggést, akkor $\boldsymbol{x} + 0 = \boldsymbol{x} = \sum_{i=1}^{k} (a_i + c_i)\boldsymbol{x}_i$, vagyis $\boldsymbol{x}$ előállítása mégsem egyértelmű. □

1.6.2. Definíció. Az L altér elemeiből álló, lineárisan független generáló rendszert *bázisnak* nevezzük.

A bázis a „ferdeszögű koordináta-rendszer” többdimenziós általánosítása. Az 1.5.5. Tétel értelmében a bázisvektorok száma mindig megegyezik a dimenzióval, továbbá a dimenzióval azonos elemszámú lineárisan független vektorrendszer mindig bázis.

11. Kérdés. Az n-dimenziós kocka $(n-1)$-dimenziós lapátlójának nevezzük a kockát határoló $(n-1)$-dimenziós kockák testátlóit. A kocka egy csúcsából kiinduló $(n-1)$-dimenziós lapátló vektorok bázist alkotnak-e $\mathbf{R}^n$-ben?

1.7. Ortogonális altér

Legyen L az $\mathbf{R}^n$ k-dimenziós valódi altere, azaz legyen $k < n$. Vegyünk fel L-ben egy bázist, amely szükségképpen k vektorból áll. Mivel $\mathbf{R}^n$-ben van n elemű lineárisan független rendszer, az 1.4.4. Tétel miatt található $n-k$ darab olyan lineárisan független vektor, amelyik ortogonális az L bázisvektoraira. Ezek a vektorok $(n-k)$-dimenziós alteret generálnak, jelöljük ezt $L^\perp$-sal. $L^\perp$ minden eleme a generáló elemeinek lineáris kombinációja, tehát minden eleme ortogonális az L minden elemére.

L és $L^\perp$ bázisvektorai együttesen $\mathbf{R}^n$ bázisát alkotják, így minden $\boldsymbol{x} \in \mathbf{R}^n$ felírható a bázisvektorok lineáris kombinációjaként. A bázisvektorok csoportosításával látható, hogy minden $\boldsymbol{x} \in \mathbf{R}^n$ felbontható $\boldsymbol{x}_1 \in L$ és $\boldsymbol{x}_2 \in L^\perp$ összegére: $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$. Az is könnyen látható, hogy $\mathbf{R}^n$-ben nincs olyan, L minden elemére ortogonális vektor, mely nem tartozik $L^\perp$-hoz. Ha lenne ilyen $\boldsymbol{x}$, akkor ez $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ alakban írható fel ($\boldsymbol{x}_1 \in L$ és $\boldsymbol{x}_2 \in L^\perp$);

szorozzuk meg mindkét oldalt $\boldsymbol{x}_1$-gyel, akkor $0 = ||\boldsymbol{x}_1||^2$, azaz $\boldsymbol{x}_1 = \mathbf{0}$, vagyis $\boldsymbol{x} \in L^\perp$ adódik. Kimondhatjuk tehát a következő definíciót:

1.7.1. Definíció. Az L altér *ortogonális kiegészítő alterének* nevezzük $\mathbf{R}^n$ azon elemeinek $L^\perp$ halmazát, melyek ortogonálisak L minden elemére.

Az előbb említett $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ ($\boldsymbol{x}_1 \in L$ és $\boldsymbol{x}_2 \in L^\perp$) ortogonális felbontás egyértelmű. Ha lenne még egy felbontás, jelölje ezt $\boldsymbol{x} = \boldsymbol{y}_1 + \boldsymbol{y}_2$ ($\boldsymbol{y}_1 \in L$ és $\boldsymbol{y}_2 \in L^\perp$), akkor $\boldsymbol{x}_1 + \boldsymbol{x}_2 = \boldsymbol{y}_1 + \boldsymbol{y}_2$, vagy másképpen $\boldsymbol{x}_1 - \boldsymbol{y}_1 = \boldsymbol{y}_2 - \boldsymbol{x}_2$, ahol $\boldsymbol{x}_1 - \boldsymbol{y}_1 \in L$ és $\boldsymbol{y}_2 - \boldsymbol{x}_2 \in L^\perp$. A két altérnek közös pontja azonban csak a $\mathbf{0}$ (nincs más önmagára ortogonális vektor), de akkor a két felbontás azonos.

Természetesen az előbbi felbontásra érvényes a Pythagorasz-tétel: $||\boldsymbol{x}||^2 = ||\boldsymbol{x}_1||^2 + ||\boldsymbol{x}_2||^2$.

12. Kérdés. L legyen azon (x, y, z, w) vektorok halmaza, melyre $2x + 3y + 4z + 5w = 0$. Adjuk meg L és $L^\perp$ egy-egy lehetséges generátorrendszerét.

Ahogy a koordináta-rendszernél, itt is az ortogonalitás – bizonyos veszteségek árán – mellőzhető.

1.7.2. Definíció. Az L altér *kiegészítő alterének* nevezzük $\mathbf{R}^n$ olyan $\bar{L}$ alterét, melyre

a) a két altérnek csak a $\mathbf{0}$ a közös pontja,

b) minden $\boldsymbol{x} \in \mathbf{R}^n$ felbontható $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ alakban, ahol $\boldsymbol{x}_1 \in L$ és $\boldsymbol{x}_2 \in \bar{L}$.

Adott $\bar{L}$ mellett a felbontás egyértelműsége ugyanúgy bizonyítható. $\bar{L}$, ugyanúgy mint $L^\perp$, $(n - k)$-dimenziós, hiszen L és $\bar{L}$ generátor vektorai együttesen $\mathbf{R}^n$-et generálják.

Szemben az ortogonális kiegészítő altérrel, $\bar{L}$ nincs egyértelműen meghatározva. Háromdimenziós példaként L legyen egy sík, amely átmegy az origón, $\bar{L}$ egy origón áthaladó, de L-hez nem illeszkedő egyenes. Általánosabb példát mutat be a következő kérdés.

13. Kérdés. Adott L altérhez vegyünk fel az $\boldsymbol{a} \neq \mathbf{0}$, $\boldsymbol{a} \in L$, és a $\boldsymbol{b} \notin L$ vektorokat, és képezzük az $\bar{L} = \{\boldsymbol{y} : \boldsymbol{y} = \boldsymbol{x} + \langle \boldsymbol{x}, \boldsymbol{b} \rangle \boldsymbol{a},\ \boldsymbol{x} \in L^\perp\}$ halmazt. Igaz-e, hogy L egyik kiegészítő alterét definiáltuk?

1.8. Egyenes kieséses táblák

Sportversenyeken gyakori versenyrendezési forma az egyenes kiesés, amikor a vesztes fél kiesik a további versenyből. A viszonylag gyors lebonyolítási rendszer legfőbb hátránya az, hogy az erős ellenfelek a verseny korai szakaszában összekerülve korán kiüthetik egymást a versenyből, így a helyezések tekintetében a tábla összeállítása torzíthat. Ennek elkerülésére kiemelést szoktak alkalmazni, biztosítva, hogy a két legerősebb ellenfél csak a döntőben, a legjobb négy csak az elődöntőben (és így tovább, amíg a kiemelés terjed) találkozhasson. A kiemelt versenyzőket a tábla kijelölt helyére beírják, és a többieket a maradék helyekre sorsolással osztják be.

Ha van egy feltételezett erősorrend a teljes mezőny vonatkozásában, akkor a teljes tábla elkészíthető az erősorrend alapján (sorsolás nélkül). Mindig feltehető, hogy a mezőny létszáma 2^n, mert hiányzó versenyzőkkel feltölthetjük a táblát, ezek ellenfelei az első fordulóban játék nélkül jutnak tovább. A hiányzó versenyzők az erősorrend alapján az utolsó helyekre kerülnek. A tábla összeállításának *első alapszabálya* az, hogy a legjobb mérkőzik a leggyengébbel, a második legjobb a második leggyengébbel, és így tovább. Ez úgy fogalmazható meg, hogy az első fordulóban a mérkőző felek helyezési számainak az összege $2^n + 1$. Az első alapszabályhoz tartozzon hozzá annak öröklődése, azaz hogy a további fordulókban is érvényesüljön ez az elv, vagyis feltételezve, hogy a mindig jobb versenyző jut tovább, minden fordulóban, ahol 2^k versenyző küzd még a továbbjutásért, a mérkőző felek helyezési számainak az összege $2^k + 1$.

Számozzuk meg a versenyzőket a feltételezett erősorrend alapján 0-tól $2^n - 1$-ig (a korábbi helyezési számokat eggyel csökkentve), és írjuk fel a sorszámaikat kettes számrendszerben n számjeggyel! Ez azt jelenti, hogy a szám elejéről az esetleg hiányzó számjegyeket pótoljuk mindig 0-val. A számjegyekből készítsünk egy n-dimenziós vektort, ami egyértelműen hozzárendelhető a versenyzőhöz. Például 128 versenyző esetén a 20. versenyzőhöz a (0, 0, 1, 0, 0, 1, 1) vektor tartozik, mert $19 = 0{\cdot}2^6 + 0{\cdot}2^5 + 1{\cdot}2^4 + 0{\cdot}2^3 + 0{\cdot}2^2 + 1{\cdot}2 + 1$.

A fenti konstrukcióval minden versenyzőnek kölcsönösen egyértelműen megfelel az n-dimenziós térben egy pont, mégpedig az n-dimenziós egységkocka egy csúcsa. Az első alapszabály azt jelenti, hogy minden versenyző a kocka testátlójának másik végpontjával mérkőzik. Ha a kocka $x_n = 0$ síkban fekvő lapját alaplapnak, az $x_n = 1$ síkban fekvő lapját fedőlapnak nevezzük, akkor minden alaplapon lévő versenyző egy fedőlapon lévővel játszik. Ha a fedőlapon lévő veszt, kiesik, ha nyer, akkor ültessük be az általa megvert versenyző helyére az alaplapba, és kapja meg az ő rangszámát. A fedőlap így kiürül, és az alaplapon lévő versenyzők folytatják a küzdelmet. Mivel az alaplap $(n-1)$-dimenziós kocka, itt ugyanígy a testátlók mentén jelöljük ki a küzdő párokat, és a módszer folytatható.

A jobb áttekinthetőség kedvéért a fenti konstrukciót síkban kellene elvégezni. E célból alkotják meg a kieséses táblákat. Példaként bemutatunk egy 16-os kieséses táblát, melyben az első alapszabály érvényesül. A tábla sorait is 0-tól 15-ig számozzuk meg. A lebonyolítás során itt feltételeztük, hogy mindig a jobban rangsorolt versenyző nyer.

0.	0	0	0	0	0
1.	15				
2.	8	7			
3.	7				
4.	4	4	3		
5.	11				
6.	12	3			
7.	3				
8.	2	2	2	1	
9.	13				
10.	10	5			
11.	5				
12.	6	6	1		
13.	9				
14.	14	1			
15.	1				

Az első alapszabály betartása mellett a versenyzők párosítása egyértelmű, ez a térbeli modell alapján nyilvánvaló. A kieséses tábla – ennek síkbeli ábrázolása – azonban nem, itt bármely ágat tükrözhetünk, és ez többször is megtehető. Ahhoz, hogy a táblát egyértelművé tegyük, szükség van egy *második alapszabályra*, feltételezzük, hogy páros sorszámú versenyzők páros sorokba kerülnek, továbbá ha mindig az erősebbnek feltételezett nyer, akkor páros sorokban is maradnak (amíg ki nem esnek a küzdelmekből). Így a tábla már egyértelmű. Ez visszafelé történő indukcióval látszik: a döntőben a 0-s és az 1-es küzd, és a 0-s van felül. Az elődöntőkben a 0, 1, 2, 3 küzd, az első alapszabály szerint $0 - 3$ és $1 - 2$ párosításban. Így a 0 van legfelül (az első alapszabály miatt), 3 alatta, a 2-es páros sorba kerül, tehát a 2. sorba, az 1-es a tábla aljára. Az indukció így folytatható.

A két alapszabálynak megfelelő kieséses tábla így visszafelé mindig megkonstruálható, de kérdéses még a közvetlen előállítás. A kérdés így hangzik: az erősorrendben k-adik versenyző a tábla hányadik sorába kerül? Keressük tehát a $0, 1, 2, \ldots, 2^n - 1$ számoknak olyan permutációját, ami teljesíti a két alapszabályt.

Definiáljuk először a 0, 1, 2, ..., $2^n - 1$ számoknak két egyszerű permutációját π_1-et és π_2-t a $\pi_1(k)$ és a $\pi_2(k)$ függvényekkel, mely a $0 \leq k < 2^n$ számokat kölcsönösen egyértelműen ugyanezekre a számokra képezi le. Írjuk fel a k-t kettes számrendszerben n számjeggyel, ennek a számnak visszafelé történő olvasata legyen a $\pi_1(k)$. Ha k-t kettes számrendszerben az ε_n, ε_{n-1}, ..., ε_2, ε_1 számjegyek állítják elő, akkor k az $(\varepsilon_n, \varepsilon_{n-1}, \ldots, \varepsilon_2, \varepsilon_1)$ vektorral adható meg. Az előbbi definíció alapján $\pi_1(k)$ ugyanígy az $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ vektorral jellemezhető.

A másik permutáció, a $\pi_2(k)$ a k számhoz a $2k$ számot rendeli, ha $2k < 2^n$, és a $2^{n+1} - 1 - 2k$ számot különben. Ez a leképezés is elmondható számjegyekkel: az $(\varepsilon_n, \varepsilon_{n-1}, \ldots, \varepsilon_2, \varepsilon_1)$ vektorhoz a $(\varepsilon_{n-1}, \ldots, \varepsilon_1, 0)$ vektort rendeli, ha $\varepsilon_n = 0$, és $(\bar{\varepsilon}_{n-1}, ..., \bar{\varepsilon}_1, 1)$ vektort, ha $\varepsilon_n = 1$. Itt az $\bar{\varepsilon}_i = 1 - \varepsilon_i$ jelölést használtuk.

Könnyen látható, hogy mindkét leképezés kölcsönösen egyértelműen képezi le a 0 és $2^n - 1$ közötti egész számok halmazát önmagára.

1.8.1. Tétel. A két alapszabálynak megfelelő egyenes kieséses táblát a $\pi = \pi_2\pi_1$ permutáció hozza létre.

Bizonyítás. A kettes számrendszerbeli számjegyekből képezett vektorokkal megadva a $\pi_2\pi_1$ permutáció a $k = (\varepsilon_n, \varepsilon_{n-1}, \ldots, \varepsilon_2, \varepsilon_1)$ vektorhoz az $(\varepsilon_2, \ldots, \varepsilon_n, 0)$ vektort rendeli, ha $\varepsilon_1 = 0$, és az $(\bar{\varepsilon}_2, ..., \bar{\varepsilon}_n, 1)$ vektort rendeli, ha $\varepsilon_1 = 1$. Ebből látható a második alapszabály teljesülése: ha k páros, azaz, ha $\varepsilon_1 = 0$, akkor $(\varepsilon_2, \ldots, \varepsilon_n, 0)$ is páros számnak felel meg, ha k páratlan, azaz $\varepsilon_1 = 1$, akkor $(\bar{\varepsilon}_2, ..., \bar{\varepsilon}_n, 1)$ is páratlannak.

Legyen $(\varepsilon_n, \varepsilon_{n-1}, \ldots, \varepsilon_2, 0)$ páros k sorszámhoz tartozó vektor, ekkor a $k_1 = 2^n - 1 - k$ számhoz tartozó vektor $(\bar{\varepsilon}_n, ..., \bar{\varepsilon}_2, 1)$. A k sorszámú versenyző sora a táblázatban $(\varepsilon_2, \ldots, \varepsilon_n, 0)$, míg a k_1 sorszámúé $(\varepsilon_2, \ldots, \varepsilon_n, 1)$, vagyis egymással fognak küzdeni. Megállapíthatjuk tehát, hogy az első fordulóban egymással szembekerülő versenyzők sorszámainak az összege $2^n - 1$.

Bizonyítandó még, hogy ha mindig a jobb nyer, akkor a második fordulóra is öröklődnek ezek a tulajdonságok. Ha egy adott versenyző győztes, akkor a sorszáma $2^n - 1$-nél kisebb, tehát a hozzá tartozó vektor $(0, \varepsilon_{n.-.1}, \ldots, \varepsilon_2, \varepsilon_1)$ alakú. A helye a táblán az $(\varepsilon_2, \ldots, \varepsilon_{n-1}, 0, 0)$ vektorral jellemezhető, ha $\varepsilon_1 = 0$, és az $(\bar{\varepsilon}_2, ..., \bar{\varepsilon}_{n-1}, 1, 1)$ vektorral, ha $\varepsilon_1 = 1$. A versenyző második fordulóbeli sorszámát az eredeti sorszám felének egészrésze adja, amihez az $(\varepsilon_2, \ldots, \varepsilon_{n-1}, 0)$, vagy az $(\bar{\varepsilon}_2, ..., \bar{\varepsilon}_{n-1}, 1)$ vektor tartozik aszerint, hogy $\varepsilon_1 = 0$ vagy 1. Ez pontosan megfelel a 2^{n-1} méretű tábla konstrukciójának. □

További érdekességként említsük meg, hogy a permutáció ismételt alkalmazása visszaállítja az eredeti sorrendet. Ha feltételezzük, hogy $\varepsilon_1 = 0$, akkor az $(\varepsilon_2, \ldots, \varepsilon_n, 0)$ vektorra alkalmazva a transzformációt $(\varepsilon_n, \ldots, \varepsilon_2,$

$0) = (\varepsilon_n, \ldots, \varepsilon_2, \varepsilon_1)$-et kapunk, $\varepsilon_1 = 1$ esetén az $(\bar{\varepsilon}_2, ..., \bar{\varepsilon}_n, 1)$ vektorból $(\varepsilon_n, \ldots, \varepsilon_2, 1) = (\varepsilon_n, \ldots, \varepsilon_2, \varepsilon_1)$ adódik. A megjegyzésből következik, hogy ha a k sorszámú játékos a j-edik sorba kerül, akkor a j sorszámú a k-adik sorba kerül. További átfogalmazást jelent az, hogy a π permutáció inverze is π, vagyis a versenyzőnek a táblán elfoglalt helyből ugyanezen eljárással lehet az eredeti erősorrendre következtetni.

14. Kérdés. Igaz-e, hogy a megadott π permutáció páros? (A páros permutáció fogalmát lásd a 4.4. fejezet végén.)

1.9. Válaszok a kérdésekre

1. Mit nevezünk az n-dimenziós kocka testátlójának? Hány testátlója van?

A testátló két csúcsot összekötő szakasz, melynek minden pontja, a végpontoktól eltekintve belső pontja a kockának. A két csúcs tehát azonos $(n-1)$-dimenziós lapon nem lehet, ezért azonos koordinátái nem lehetnek. Mivel a csúcspontok minden koordinátája csak 0 vagy 1 lehet, ezért a másik (átellenes) csúcs már egyértelműen meghatározott. Például a (0, 1, 1, 0, 0) csúcsból kiinduló testátló csak az (1, 0, 0, 1, 1) csúcsban végződhet.

Mivel minden csúcsból egyetlen testátló indul ki, és 2^ncsúcs van, a testátlók száma 2^{n-1}.

2. n darab számszerű statisztikai megfigyelés, $a_1, a_2, \ldots, a_n$, az n-dimenziós tér egy pontjával azonosítható. Jelöljük $\bar{a}$-sal a minta átlagát, azaz legyen $\bar{a} = \frac{a_1+a_2+...+a_n}{n}$, és képezzük az $\boldsymbol{x} = (a_1 - \bar{a}, a_2 - \bar{a}, ..., a_n - \bar{a})$ vektorokat. Igazoljuk, hogy ezek a vektorok az $\mathbf{R}^n$ egy valódi alterébe esnek. Mi jellemzi ezt az alteret?

$\boldsymbol{x}$ mintaátlaga már nulla lesz, vagyis a koordinátáinak összege nulla. Ez a tulajdonság megőrződik, ha két ilyen vektort összeadunk, vagy ha számmal szorzunk, tehát az ilyen vektorok halmaza alteret alkot. Legyen $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, akkor az altér pontjait az $x_1 + x_2 + ... + x_n = 0$ egyenlet jellemzi.

3. Ha az $a_1, a_2, \ldots, a_n$ statisztikai mintából elkészítjük az $\boldsymbol{x} = (a_1 - \bar{a}, a_2 - \bar{a}, ..., a_n - \bar{a})$ vektort, ahol $\bar{a}$ a mintaátlagot jelöli, akkor a minta szórása és a $\boldsymbol{x}$ normája között milyen összefüggés áll fenn?

Mivel a minta szórása az $\frac{1}{\sqrt{n}}\sqrt{(a_1-\bar{a})^2 + (a_2-\bar{a})^2 + ... + (a_n-\bar{a})^2}$ képlettel számolható, a minta szórása $\frac{||\boldsymbol{x}||}{\sqrt{n}}$.

4. Ha $\boldsymbol{a} \in \mathbf{R}^4$ és $||\boldsymbol{a}|| = 4$, továbbá $\mathbf{1} = (1, 1, 1, 1)$, akkor mennyi az $\langle \boldsymbol{a} - \mathbf{1}, \boldsymbol{a} + \mathbf{1} \rangle$ kifejezés értéke?

$\langle \boldsymbol{a}-\mathbf{1}, \boldsymbol{a}+\mathbf{1}\rangle = ||\boldsymbol{a}||^2-\langle \mathbf{1}, \cdot\boldsymbol{a}\rangle+\langle \mathbf{1}, \cdot\boldsymbol{a}\rangle-\langle \mathbf{1}, \cdot\mathbf{1}\rangle = ||\boldsymbol{a}||^2-||\mathbf{1}||^2 = 16-4 =$ $= 12$. (Vigyázat $\langle \mathbf{1},\mathbf{1}\rangle \neq 1$, mint ahogy $\langle \mathbf{1},\boldsymbol{a}\rangle$ sem egyezik meg $\boldsymbol{a}$-val!)

5. Legyen $a_1, a_2, \ldots, a_n$ és $b_1, b_2, \ldots, b_n$ két statisztikai minta, vonjuk le a mintaelemekből az átlagukat, a kapott vektorokat jelöljük $\boldsymbol{x}$-szel ill. $\boldsymbol{y}$-nal. Mi az összefüggés a $\boldsymbol{x}$ és az $\boldsymbol{y}$ korrelációs együtthatója és a szöge között?

A két minta, és egyben az $\boldsymbol{x}$ és az $\boldsymbol{y}$ korrelációs együtthatója az

$$r(\boldsymbol{x}, \boldsymbol{y}) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n\sigma_1\sigma_2}$$

képlettel számolható ki, ahol σ_1 és σ_2 a két minta szórása. Egyszerű átalakítással $r(\boldsymbol{x},\ \boldsymbol{y}) = \frac{\langle \boldsymbol{x},\boldsymbol{y}\rangle}{||\boldsymbol{x}||\cdot||\boldsymbol{y}||} = \cos\gamma$. Ebből az összefüggésből nyilvánvaló az a statisztikai állítás, hogy

$$|r(\boldsymbol{x}, \boldsymbol{y})| \ \leq \ 1.$$

6. Az n-dimenziós kockának vannak-e ortogonális testátlói?

Tekintsük a koordináta egységvektorok által meghatározott kockát. Az 1. Kérdésnél láttuk, hogy a testátló a szemközti csúcsokat köti össze, és ha a kocka egyik csúcsa $\boldsymbol{a}$, akkor a szemközti csúcs $\mathbf{1} - \boldsymbol{a}$ ($\boldsymbol{a}$ minden koordinátája 0 vagy 1). Az összekötő vektor $\mathbf{1} - 2\boldsymbol{a}$, melynek minden koordinátája 1 vagy -1. Két ilyen vektor skalárszorzata ± 1-ek összege, mely nullát csak úgy adhat, ha páros sok összeadandó van. Páros dimenziós térben tehát lehet, páratlanban nem!

Például $\mathbf{R}^4$-ben:

$$\begin{aligned} \boldsymbol{a} &= (1,0,0,1), \quad \mathbf{1} - \boldsymbol{a} = (0,1,1,0), \quad \mathbf{1} - 2\boldsymbol{a} = (-1,1,1,-1) \\ \boldsymbol{b} &= (0,0,1,1), \quad \mathbf{1} - \boldsymbol{b} = (1,1,0,0), \quad \mathbf{1} - 2\boldsymbol{b} = (1,1,-1,-1), \end{aligned}$$

és az $\mathbf{1} - 2\boldsymbol{a}$ és $\mathbf{1} - 2\boldsymbol{b}$ vektorok ortogonálisak.

7. Igaz-e a vektorrendszer lineáris függetlensége akkor, ha bármely két vektor szöge 60°?

Feltehető, hogy a lineárisan független rendszert alkotó vektorok egységvektorok, a feltétel szerint ekkor bármely kettő skalárszorzata $\frac{1}{2}$. Tegyük fel, hogy

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_k\boldsymbol{a}_k = 0,$$

és szorozzuk meg az egyenlőség mindkét oldalát $\boldsymbol{a}_j$-vel, akkor

$$\frac{1}{2}(c_1 + c_2 + ... + 2c_j + ... + c_k) = 0$$

adódik minden $1 \leq j \leq k$-ra. Az összes egyenletet összeadva

$$\frac{k+1}{2}(c_1 + c_2 + ... + c_j + ... + c_k) = 0,$$

amiből $c_j = 0$-t kapunk minden j-re. Ez a lineáris függetlenséget jelenti.

8. Lehetséges-e, hogy az n-dimenziós kocka összes testátló-vektora lineárisan független rendszert alkot?

A testátlók száma az 1. Kérdés szerint 2^{n-1}, de $\mathbf{R}^n$-ben legfeljebb n lineárisan független vektor van. Mivel $2^{n-1} > n$, ez nem lehetséges, kivéve az $n = 2$ esetet. A négyzet átlóvektorai valóban linárisan függetlenek.

9. Az $(x_1,\, x_2,\, x_3)$ vektor legyen az

$$\begin{aligned} a_1x_1 + b_1x_2 + c_1x_3 &= 0 \\ a_2x_1 + b_2x_2 + c_2x_3 &= 0 \\ a_3x_1 + b_3x_2 + c_3x_3 &= 0 \end{aligned}$$

ún. homogén lineáris egyenletrendszer megoldása. Igaz-e, hogy a megoldások halmaza $\mathbf{R}^3$ altere? Igaz-e ugyanez, ha az egyenletrendszer nem homogén, azaz a jobb oldalon nem csupa nulla áll?

Könnyen látható, hogy ha $(x_1,\, x_2,\, x_3)$ és $(y_1,\, y_2,\, y_3)$ megoldások, akkor $(x_1+y_1,\, x_2 + y_2,\, x_3+y_3)$ és $(cx_1,\, cx_2,\, cx_3)$ is megoldás, tehát a megoldások tényleg alteret alkotnak. Itt sem az egyenletek száma, sem az ismeretlenek száma nem számít, egyenlőknek sem kell lenniük.

Ha az egyenletrendszer nem homogén, akkor a $\mathbf{0}$ nem megoldás, tehát altérről nem eshet szó.

10. A $x_1 + x_2 + ... + x_n = 0$ egyenletnek eleget tevő $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ vektorok a 9. Kérdés szerint alteret alkotnak. Generálják-e ezt az alteret az

$$\begin{aligned} \boldsymbol{a}_1 &= (1,\, 0,\, 0,\, ...,\, 0,\, -1), \\ \boldsymbol{a}_2 &= (0,\, 1,\, 0,\, ...,\, 0,\, -1), \\ \boldsymbol{a}_3 &= (0,\, 0,\, 1,\, ...,\, 0,\, -1), \\ &\cdots \\ \boldsymbol{a}_{n-1} &= (0,\, 0,\, 0,\, ...,\, 1,\, -1) \end{aligned}$$

vektorok? Mennyi az altér dimenziója?

Ellenőrizhető, hogy a vektorok az altérhez tartoznak. Másrészt vegyünk egy, az altérhez tartozó $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ vektort, akkor $\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 +$

$+\ldots+x_{n-1}\boldsymbol{a}_{n-1}$, hiszen első koordinátája x_1, a második $x_2, \ldots$, az $(n-1)$-edik x_{n-1}, és az n-edik koordináta $-(x_1+x_2+\ \ldots\ +x_{n-1})$ lesz, ami az egyenlet alapján x_n. Vagyis a vektorok generálják az alteret.

A vektorok lineárisan függetlenek, mert bármely kettő szöge 60^o, és a 7. Kérdésben éppen ezt bizonyítottuk be. (A lineáris függetlenség másképpen is belátható. Felírva a $c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2+\ldots+c_{n-1}\boldsymbol{a}_{n-1}=0$ relációt koordinátánként, az i-edik koordinátára $c_i=0$ adódik, ahol i tetszőleges, $1\le i\le n-1$.) Az altér dimenziója az elmondottak szerint $n-1$.

11. Az n-dimenziós kocka $(n-1)$-dimenziós lapátlójának nevezzük a kockát határoló $(n-1)$-dimenziós kockák testátlóit. A kocka egy csúcsából kiinduló $(n-1)$-dimenziós lapátló vektorok bázist alkotnak-e $\mathbf{R}^n$-ben?

Tekintsük ismét a koordináta egységvektorok által meghatározott kockát, és a kiválasztott csúcs legyen az origó. A kockának n olyan lapja van, melynek pontja az origó, mégpedig egy-egy lapot azon pontok alkotnak, amelyeknél egy koordinátát rögzítünk nullának, a többi 0 és 1 között szabadon változhat. A rögzített koordinátát pedig n-féleképpen tudjuk kiválasztani. Ebből azt is megkapjuk, hogy a lapon szemközti csúcsok, és egyben a lapátló vektorok koordinátái:

$$\begin{aligned}\boldsymbol{a}_1 &= (0,\ 1,\ 1,\ ...,\ 1,\ 1),\\ \boldsymbol{a}_2 &= (1,\ 0,\ 1,\ ...,1,\ 1),\\ \boldsymbol{a}_3 &= (1,\ 1,\ 0,\ ...,\ 1,\ 1),\\ &\ \ \cdots\\ \boldsymbol{a}_n &= (1,\ 1,\ 1,\ ...,\ 1,\ 0).\end{aligned}$$

A lapátlókkal előállítható az origóból kiinduló $\boldsymbol{b}$ testátló: $\boldsymbol{b}=\mathbf{1}=\frac{1}{n-1}(\boldsymbol{a}_1+\boldsymbol{a}_2+\ldots+\boldsymbol{a}_n)$, és előállítható tetszőleges koorditáta egységvektor is: $\boldsymbol{e}_i=(0,0,0,\ldots,1,\ldots,0)=\boldsymbol{b}$ - $\boldsymbol{a}_i$. A lapátlók tehát generálják a koordináta egységvektorokat, így a teljes teret is. Ebből következik, hogy az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ vektorok lineárisan függetlenek, mert összefüggés esetén n-nél kisebb méretű generátorrendszer létezne.

12. L legyen azon $(x,\ y,\ z,\ w)$ vektorok halmaza, melyre $2x+3y+4z+5w=0$. Adjuk meg L és $L^{\perp}$ egy-egy lehetséges generátorrendszerét.

A 10. Kérdéshez hasonlóan a generátorrendszer:

$$\begin{aligned}\boldsymbol{a}_1 &= (5,\ 0,\ 0,\ -2),\\ \boldsymbol{a}_2 &= (0,\ 5,\ 0,\ -3),\\ \boldsymbol{a}_3 &= (0,\ 0,\ 5,\ -4).\end{aligned}$$

Könnyen látható, hogy az $(x,\ y,\ z,\ w)$ koordinátájú vektor előáll $\frac{x}{5}\boldsymbol{a}_1+$ $+\frac{y}{5}\boldsymbol{a}_2+\frac{z}{5}\boldsymbol{a}_3$ alakban. Az altér dimenziója 3, ezért az ortogonális kiegészítő altér egydimenziós. Generáló vektora

$$\boldsymbol{a}_4=(2,\ 3,\ 4,\ 5).$$

13. Adott L altérhez vegyünk fel az $\boldsymbol{a}\neq\boldsymbol{0}$, $\boldsymbol{a}\in L$, és a $\boldsymbol{b}\notin L$ vektorokat, és képezzük az $\bar{L}=\{\boldsymbol{y}:\boldsymbol{y}=\boldsymbol{x}+\langle\boldsymbol{x},\boldsymbol{b}\rangle\boldsymbol{a},\ \ \boldsymbol{x}\in L^{\perp}\}$ halmazt. Igaz-e, hogy L egyik kiegészítő alterét definiáltuk?

Igazoljuk, hogy $\bar{L}$ altér. Ha $\boldsymbol{y}_i=\boldsymbol{x}_i+\langle\boldsymbol{x}_i,\boldsymbol{b}\rangle\boldsymbol{a}\quad(\boldsymbol{x}_i\in L^{\perp};\ i=1,\ 2)$, akkor

$$\boldsymbol{y}_1+\boldsymbol{y}_2=\boldsymbol{x}_1+\langle\boldsymbol{x}_1,\boldsymbol{b}\rangle\boldsymbol{a}+\boldsymbol{x}_2+\langle\boldsymbol{x}_2,\boldsymbol{b}\rangle\boldsymbol{a}=\boldsymbol{x}_1+\boldsymbol{x}_2+\langle\boldsymbol{x}_1+\boldsymbol{x}_2,\boldsymbol{b}\rangle\boldsymbol{a}$$

és $\boldsymbol{x}_1+\boldsymbol{x}_2\in L^{\perp}$, tehát $\boldsymbol{y}_1+\boldsymbol{y}_2\in\bar{L}$. Hasonlóan $c\boldsymbol{y}_1=c\boldsymbol{x}_1+\langle c\boldsymbol{x}_1,\boldsymbol{b}\rangle\boldsymbol{a}$ és $c\boldsymbol{x}_1\in L^{\perp}$, tehát $c\boldsymbol{y}_1\in\bar{L}$.

Igazoljuk, hogy a két altér közös része $\{\boldsymbol{0}\}$. Tegyük fel, hogy $\boldsymbol{y}=\boldsymbol{x}+$ $+\langle\boldsymbol{x},\ \boldsymbol{b}\rangle\boldsymbol{a}=\boldsymbol{z}$ és $\boldsymbol{z}\in L$, akkor $\boldsymbol{x}=\boldsymbol{z}-\langle\boldsymbol{x},\ \boldsymbol{b}\rangle\boldsymbol{a}$, ahol $\boldsymbol{x}\in L^{\perp}$, és $\boldsymbol{z}-\langle\boldsymbol{x},\ \boldsymbol{b}\rangle\boldsymbol{a}\in L$. Ez csak úgy lehetséges, ha $\boldsymbol{x}=\boldsymbol{0}$, de akkor $\boldsymbol{y}=\boldsymbol{z}=\boldsymbol{0}$.

Igazoljuk a generálást. Legyen $\boldsymbol{x}\in\mathbf{R}^n$, akkor $\boldsymbol{x}=\boldsymbol{u}+\boldsymbol{v}$, ahol $\boldsymbol{u}\in L$ és $\boldsymbol{v}\in L^{\perp}$. Átalakítva:

$$\boldsymbol{x}=(\boldsymbol{u}-\langle\boldsymbol{v},\boldsymbol{b}\rangle\boldsymbol{a})+(\boldsymbol{v}+\langle\boldsymbol{v},\boldsymbol{b}\rangle\boldsymbol{a}),$$

ahol $\boldsymbol{u}-\langle\boldsymbol{v},\boldsymbol{b}\rangle\boldsymbol{a}\in L$ és $\boldsymbol{v}+\langle\boldsymbol{v},\ \boldsymbol{b}\rangle\boldsymbol{a}\in\bar{L}$, ami a kívánt előállítást adja.

Megjegyzés. Az adott k-dimenziós L altérhez vegyük fel az $\boldsymbol{a}_i\neq\boldsymbol{0}$, $\boldsymbol{a}_i\in L$, és a $\boldsymbol{b}_i$ vektorokat $(i=1,\ 2,\ \ldots,\ k)$, és képezzük az $L_1=\{\boldsymbol{y}:\boldsymbol{y}=\boldsymbol{x}+\sum\limits_{i=1}^{k}\langle\boldsymbol{x},\boldsymbol{b}_i\rangle\boldsymbol{a}_i,\ \ \boldsymbol{x}\in L^{\perp}\}$ halmazt. Ugyanúgy megmutatható, hogy L kiegészítő alterét kapjuk. Igaz-e, hogy L minden kiegészítő altere ilyen alakú?

14. Igaz-e, hogy a megadott π permutáció páros? (A páros permutáció fogalmát lásd a 4.4. fejezet végén.)

Először határozzuk meg, hogy hány elem marad a helyén a permutáció során. Mint említettük, az $(\varepsilon_n,\varepsilon_{n-1},\ldots,\varepsilon_2,0)$ vektor átmegy az $(\varepsilon_2,\ldots,\varepsilon_n,$ $0)$ vektorba, míg $(\varepsilon_n,\varepsilon_{n-1},\ldots,\varepsilon_2,1)$ az $(\bar{\varepsilon}_2,\ldots,\ \bar{\varepsilon}_n,\ 1)$-be megy át. Másképp fogalmazva, ha $(\varepsilon_n,\ \varepsilon_{n-1},\ \ldots,\ \varepsilon_2)$ szimmetrikus, vagy antiszimmetrikus, akkor generál egy fix elemet. Ha $j=\left[\frac{n+1}{2}\right]$, akkor 2^j ilyen vektor van, tehát ennyi fix eleme van a permutációnak (egyszerű kombinatorikai feladat).

A permutáció során tehát $2^n - 2^j$ elem helyet változtat. Mivel $\pi = \pi^{-1}$, ezek az elemek párokba sorolhatók, és a párok elemei helyet cserélnek a permutáció során. Az eredeti sorrend tehát $\frac{1}{2}(2^n - 2^j)$ pár cseréjével helyreállítható. Minden pár cseréje szomszédos elemek páratlan számú cseréjével elvégezhető, így az összes szomszédos párcsere paritása megegyezik $2^{n-1} - 2^{j-1}$ paritásával, a π permutáció tehát páros, ha $2^{n-1} - 2^{j-1}$ páros szám.

$n = 0$, vagy $n = 1$ esetén a π permutáció az identitás, tehát páros. Ha $n \geq 3$, akkor $j \geq 2$, vagyis $2^{n-1} - 2^{j-1}$ két páros szám különbsége, azaz páros, tehát π is páros. Ha $n = 2$, akkor $j = 1$, vagyis $2^{n-1} - 2^{j-1} = 1$, vagyis egyedül ebben az egy esetben nem lesz páros a π permutáció.

2. fejezet

Lineáris leképezések

2.1. Lineáris függvény

Idézzük fel a függvény definícióját. Adott két halmaz, X és Y, az f az X halmaz minden eleméhez egyértelműen hozzárendel egy Y-ba eső elemet, akkor ezt a hozzárendelést nevezzük $f\colon X \to Y$ függvénynek.

Példaként legyen $X = \mathbf{R}^2$, $Y = \mathbf{R}$. Akkor $f\colon X \to Y$ a sík (x, y) pontjaihoz rendel egy z számértéket, amit szokásosan kétváltozós függvénynek nevezünk. A kétváltozós függvény lineáris, ha a két változónak elsőfokú polinomjáról van szó, azaz, ha $f(x, y) = ax + by + c$ alakú. A függvény homogén lineáris, ha $c = 0$, azaz csak elsőfokú tagokból áll, nulladfokú konstans nincs benne: $z = ax + by$. A továbbiakban ezt feltételezzük, nem jelent túl lényeges korlátozást, de a tárgyalást megkönnyíti. Mi a jellemző a (homogén) lineáris függvényre? A $c{\cdot}(x, y)$-hoz hozzárendelt érték, bármely c szám esetén, cz lesz, továbbá ha (x_1, y_1)-hez a z_1 tartozik és (x_2, y_2)-höz a z_2, akkor $(x_1 + x_2, y_1 + y_2)$-höz a függvény a $z_1 + z_2$-t rendeli hozzá. A függvény jelölésmódján a továbbiakban többször is változtatunk: A-val jelöljük a lineáris függvényt, és a változót közrefogó zárójeleket is mellőzzük ($A\boldsymbol{x}$ olvasata: A alkalmazva $\boldsymbol{x}$-re). Ennek megfelelően definiáljuk a lineáris leképezést.

2.1.1. Definíció. Az $A\colon \mathbf{R}^m \to \mathbf{R}^n$ függvényt *lineáris leképezésnek* nevezzük, ha bármely $\boldsymbol{x}$, $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in \mathbf{R}^m$-re és bármely c számra

a) $Ac\boldsymbol{x} = c \cdot A\boldsymbol{x}$,
b) $A(\boldsymbol{x}_1 + \boldsymbol{x}_2) = A\boldsymbol{x}_1 + A\boldsymbol{x}_2$.

Természetesen A értelmezési tartománya $\mathbf{R}^m$, értékkészlete, amit képtérnek is nevezünk, R_A. R_A-ról egyelőre csak annyit mondhatunk, hogy

$R_A \subset \mathbf{R}^n$. Az előbbi a) és b) tulajdonságok biztosítják, hogy R_A mindig altér.

1. Kérdés. Lineáris leképezés-e az A: $\mathbf{R}^2 \to \mathbf{R}^3$, ha $A(x, y) = (x + y, 2x - y, 0)$? Lineáris leképezés-e az A: $\mathbf{R}^2 \to \mathbf{R}^3$, ha $A(x, y) = (x + y, 2x - y, 1)$?

Hogyan adható meg egy lineáris leképezés? Természetesen a fenti kérdésnek megfelelően képlettel, de ennél célszerűbb megadási módot is fogunk találni. Ha megadjuk azt, hogy az

$$\boldsymbol{e}_1 = (1,\, 0,\, 0,\, 0,\, \dots,\, 0),$$

$$\boldsymbol{e}_2 = (0,\, 1,\, 0,\, 0,\, \dots,\, 0),$$

$$\boldsymbol{e}_3 = (0,\, 0,\, 1,\, 0,\, \dots,\, 0),$$

$$\vdots$$

$$\boldsymbol{e}_m = (0,\, 0,\, 0,\, 0,\, \dots,\, 1),$$

koordináta egységvektorok a leképezés során mely vektorokba mennek át, akkor a linearitás miatt a teljes leképezést megadtuk. Legyen ugyanis $\boldsymbol{a}_i = A\boldsymbol{e}_i$ $(i = 1,\, 2,\, \dots,\, m)$, akkor az

$$\boldsymbol{x} = (x_1,\ x_2,\ \dots,\ x_m) = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_m\boldsymbol{e}_m$$

vektor képe

$$A\boldsymbol{x} = A(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_m\boldsymbol{e}_m) = x_1A\boldsymbol{e}_1 + x_2A\boldsymbol{e}_2 + \ldots + x_mA\boldsymbol{e}_m = \\ = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_m\boldsymbol{a}_m.$$

Az A leképezés tehát paraméterekkel megadható, a paraméterek ($n \times m$ darab szám) m darab vektorként sorolhatók fel, ha az A megadásánál ezt fel akarjuk tüntetni, akkor A helyett bővebben $A(\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m)$-et írunk. Itt $A(\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m)$ maga a leképezés, ha ezt alkalmazzuk $\boldsymbol{x}$-re, akkor $A(\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m)\boldsymbol{x}$-et kell írni.

A jelölésmódot tovább egyszerűsítjük: $A(\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m)$ helyett egyszerűen $(\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m)$-t írunk, és az $(\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m)$ leképezést nevezzük $\boldsymbol{A}$-nak, azaz legyen $\boldsymbol{A} = (\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m)$. Az $\boldsymbol{A} = (\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m)$ kifejezést *mátrixnak* nevezzük.

Ha részletesen ki akarjuk írni az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ... és az $\boldsymbol{a}_m$ vektorokat, mondjuk, legyen

$$\begin{aligned}\boldsymbol{a}_1 &= (1, 0, 3, 4),\\ \boldsymbol{a}_2 &= (-2, 3, 0, -1),\\ \boldsymbol{a}_3 &= (1, 1, -1, 2),\end{aligned}$$

akkor ezt az

$$\boldsymbol{A} = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 3 & 1 \\ 3 & 0 & -1 \\ 4 & -1 & 2 \end{pmatrix}$$

formában tehetjük. A vektorokat tehát függőlegesen írjuk fel elválasztó jeleket nem alkalmazva. Ezt úgy mondjuk, hogy ebben a témakörben a vektoroknál szokásos (vízszintesen felírt) sorvektorok helyett (függőleges elrendezésű) ún. oszlopvektorokat használunk. Innen már csak egy lépés az, hogy $\boldsymbol{A}$ alkalmazva az $\boldsymbol{x}$-re helyett $\boldsymbol{A}$-szor $\boldsymbol{x}$-et mondunk, ahol a mátrix és a vektor szorzatát az előbbieknek megfelelően, alkalmasan definiáljuk. Ha $\boldsymbol{x} = (x, y, z)$, akkor

$$\boldsymbol{A}\boldsymbol{x} = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 3 & 1 \\ 3 & 0 & -1 \\ 4 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - 2y + z \\ 3y + z \\ 3x - z \\ 4x - y + 2z \end{pmatrix}.$$

A szorzás szabályait később majd pontosan megtanuljuk.

A továbbiakban az $\boldsymbol{A}$ mátrix által létrehozott lineáris leképezést és az $\boldsymbol{A}$ mátrixot, azonos jelölést alkalmazva párhuzamosan használjuk. (Függvényeknél is gyakran $f(x)$ helyett f-et írunk.)

Említettük, hogy az R_A képtér mindig altér. Rögtön látszik, hogy ha az $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$ mátrixról van szó, akkor $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ..., $\boldsymbol{a}_m$ eleme az R_A-nak, hiszen $\boldsymbol{a}_i = \boldsymbol{A}\boldsymbol{e}_i$. Ráadásul $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ..., $\boldsymbol{a}_m$ generáló rendszere R_A-nak az $\boldsymbol{A}\boldsymbol{x}$-re fent megadott formula alapján. Így R_A dimenziója a generáló rendszer rangja (ld. 1.5.5. Definíció). Ezt a számot egyben az $\boldsymbol{A}$ mátrix rangjának is nevezzük.

2.1.2. Definíció. Az $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$ *mátrix rangja* az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ..., $\boldsymbol{a}_m$ vektorok között található lineárisan független vektorok maximális száma. Ez egyben a lineáris leképezés rangja is. Jelölése: $\operatorname{rang}(\boldsymbol{A})$.

2. Kérdés. Jelöljük az $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ minta átlagát $\bar{x}$-sal, és képezzük az $\boldsymbol{y} = (x_1 - \bar{x}, x_2 - \bar{x}, \ldots, x_n - \bar{x})$ vektort. Adjuk meg annak a lineáris transzformációnak a mátrixát, mely az $\boldsymbol{x}$-et az $\boldsymbol{y}$-ba képezi le! Mennyi a mátrix rangja?

2.2. Lineáris műveletek mátrixokkal

Ismerkedjünk meg a mátrixok írásmódjával, jelöléseivel és a mátrixszal kapcsolatos fogalmakkal.

A mátrixot, általános alakban, az alábbi formában szoktuk felírni:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{pmatrix}$$

Az elemek jelölése dupla indexszel történik (tehát a_{12} olvasata a egy kettő és nem a tizenkettő!), az indexek közé írásjelet, szóközt nem teszünk. A mátrix vízszintesen elhelyezkedő sorokból, vagy sorvektorokból, és függőlegesen elhelyezkedő oszlopokból, oszlopvektorokból áll. A két index közül mindig a sorindex az első, az oszlopindex a második:

$$a_{23}$$

sorindex oszlopindex

A sorindex tünteti fel, hogy hányadik sor, az oszlopindex, hogy hányadik oszlop eleméről van szó.

A mátrixról röviden azt mondjuk, hogy $n \times m$-es méretű, ez azt jelenti, hogy n sora és m oszlopa van. Az első szám mindig a sorméret (hány sora van, és nem a sornak a mérete), a második az oszlopméret. A mátrixban lévő számokat elemeknek nevezzük. A mátrix elemeit nagyméretű zárójellel fogjuk össze, a zárójel lehet kerek vagy szögletes (kapcsos nem). Rövid jelölésként használjuk az $\boldsymbol{A} = (a_{ij})$ jelölést is, ez nem mond mást, mint, hogy az $\boldsymbol{A}$ mátrix az a_{ij} elemekből áll, a méretét itt nem tüntetjük fel.

A mátrixok használatának elsődleges célja a lineáris leképezések megadása, kezelése. Minden, amit felépítünk, elsődlegesen ezt a célt szolgálja. A későbbi alkalmazásokban ez a cél többnyire a háttérben lappang, esetleg észre sem vesszük, hogy lineáris leképezésekről van szó.

Lineáris leképezésekkel lehet műveleteket végezni. Egy lineáris leképezést lehet a c számmal szorozni: az eredetileg hozzárendelt vektor c-szeresét rendeljük hozzá. Lehet két lineáris leképezést összeadni: az eredetileg hozzárendelt vektorok összegét rendeljük hozzá.

Az első esetben az $\boldsymbol{e}_k$ vektorokhoz nem az $\boldsymbol{a}_k$, hanem a $c\boldsymbol{a}_k$ vektorokat rendeli hozzá, vagyis a leképezés mátrixa $c\boldsymbol{A} = (c\boldsymbol{a}_1,\ c\boldsymbol{a}_2,\ \ldots,\ c\boldsymbol{a}_m)$ lesz. A szabály nagyon egyszerű: mátrixot számmal úgy szorzunk, hogy minden elemét megszorozzuk.

A második esetben adott az $\boldsymbol{A} = (\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \ldots,\ \boldsymbol{a}_m)$ és a $\boldsymbol{B} = (\boldsymbol{b}_1,\ \boldsymbol{b}_2,\ \ldots,\ \boldsymbol{b}_m)$ mátrix. Az $\boldsymbol{A} + \boldsymbol{B}$ az $\boldsymbol{e}_k$ vektorokhoz az $\boldsymbol{a}_k + \boldsymbol{b}_k$ vektorokat rendeli hozzá: $\boldsymbol{A} + \boldsymbol{B} = (\boldsymbol{a}_1 + \boldsymbol{b}_1,\ \boldsymbol{a}_2 + \boldsymbol{b}_2,\ \ldots,\ \boldsymbol{a}_m + \boldsymbol{b}_m)$, tehát mátrixokat úgy adunk össze, hogy a megfelelő oszlopvektoraikat összeadjuk, ami megfelel annak, hogy az azonos helyen álló elemeiket összeadjuk. Természetes az is, hogy két lineáris leképezést csak akkor tudunk összeadni, ha mindkettő $\mathbf{R}^m \to \mathbf{R}^n$ leképezés, és ekkor az összeg is ilyen lesz. Mátrixok vonatkozásában elmondva: két mátrixot csak akkor tudunk összeadni, ha méreteik megegyeznek.

2.2.1. Definíció. Az $\boldsymbol{A}$ mátrix *c-szerese*

$$c\boldsymbol{A} = \begin{pmatrix} ca_{11} & ca_{12} & ca_{13} & \cdots & ca_{1m} \\ ca_{21} & ca_{22} & ca_{23} & \cdots & ca_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & ca_{n3} & \cdots & ca_{nm} \end{pmatrix}.$$

Az $\boldsymbol{A}$ és a $\boldsymbol{B} = (b_{ij})$ mátrixokat *összeadni* csak akkor lehet, ha méreteik megegyeznek, ekkor

$$\boldsymbol{A} + \boldsymbol{B} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} & \cdots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} & \cdots & a_{2m}+b_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & a_{n3}+b_{n3} & \cdots & a_{nm}+b_{nm} \end{pmatrix}$$

A mátrixok összeadására és számmal történő szorzására vonatkozóan a valós számokra ismert összes – kommutatív, asszociatív és disztributív – tulajdonság fennáll, hiszen elemenkénti, valós számokra elvégzett műveletekről van szó. Például, $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$, vagy $c(\boldsymbol{A} + \boldsymbol{B}) = c\boldsymbol{A} + c\boldsymbol{B}$.

Beszéljünk külön az $\mathbf{R}^m \to \mathbf{R}$ lineáris leképezésről, amikor az $\mathbf{R}^m$-en értelmezett függvény (valós) szám. Az ilyen leképezést *lineáris funkcionálnak* is nevezik. Ez esetben is előállítható az $\boldsymbol{A} = (\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \ldots,\ \boldsymbol{a}_m)$ mátrix, csak minden $\boldsymbol{a}_i$ egydimenziós vektor, vagyis szám. Ez azt jelenti, hogy $\boldsymbol{A}$ $1 \times m$-es mátrix, azaz sorvektor. Kimondhatjuk az alábbi tételt:

2.2.2. Tétel. Az $\mathbf{R}^m$-en értelmezett tetszőleges $\boldsymbol{A}$ lineáris funkcionálhoz található olyan $\boldsymbol{a} \in \mathbf{R}^m$ sorvektor, hogy $\boldsymbol{A}\boldsymbol{x} = \langle \boldsymbol{a},\ \boldsymbol{x} \rangle$, minden $\boldsymbol{x} \in$

$\in \mathbf{R}^m$-re. Röviden: $\mathbf{R}^m$-ben minden lineáris funkcionál skalárszorzattal adható meg.

(A tétel jóval általánosabb körülmények között is érvényes, és Riesz-tétel néven vált ismertté.)

3. Kérdés. Hány dimenziós teret alkot az $\mathbf{R}^m \to \mathbf{R}^n$ lineáris leképezések halmaza?

2.3. Szorzás mátrixszal

Mátrix és vektor szorzatát a 2.1. fejezetben már bevezettük. Az $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$ és az $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$ vektor szorzata a fentiek szerint $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \ldots + \boldsymbol{a}_m x_m$. Az $\boldsymbol{y}$ vektor i-edik koordinátája részletesen kiírva: $y_i = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{im}x_m$. Tanuljuk meg: mátrixot vektorral akkor szorozhatunk össze, ha a mátrix oszlopmérete megegyezik a vektor méretével, és az eredményvektor i-edik koordinátáját a mátrix i-edik sorvektorának és az adott vektornak a skalárszorzata adja.

A mátrix-aritmertikában – hacsak mást nem mondunk – a vektorokat mindig oszlopvektoroknak írjuk fel. Ennek jelentőségét már az $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$ mátrixkonstrukciónál is láttuk.

4. Kérdés. Gyakran szükségessé válik az $\boldsymbol{a} = (a_1, a_2, \ldots, a_m)$ és az $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$ vektor $(a_1x_1, a_2x_2, \ldots, a_mx_m)$ alakú „szorzatát” előállítani (pl. raktárkészlet értéke árufajtánként). Hogyan tehetjük meg az eddigi műveletekkel?

$\boldsymbol{BA}$ – a lineáris transzformációk vagy mátrixok szorzata – jelentse a lineáris leképezések egymás utáni elvégzését, először az $\boldsymbol{A}$-t, majd a $\boldsymbol{B}$-t elvégezve. Ha $\boldsymbol{A}$ az $\mathbf{R}^m$-et $\mathbf{R}^p$-be viszi át, akkor a $\boldsymbol{B}$ csak akkor alkalmazható, ha az $\mathbf{R}^p$-t képezi le, mondjuk $\mathbf{R}^n$-be. A $\boldsymbol{BA}$ leképezés tehát $\mathbf{R}^m \to \mathbf{R}^n$ leképezés, és nyilván továbbra is lineáris. Ha $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$, akkor az $\boldsymbol{e}_j$ koordinátavektort az $\boldsymbol{A}$ az $\boldsymbol{a}_j$ vektorba viszi át, $\boldsymbol{BA}$ pedig a $\boldsymbol{Ba}_j$-be:

$$\boldsymbol{e}_j \xrightarrow{\boldsymbol{A}} \boldsymbol{A}\mathbf{e}_j = \boldsymbol{a}_j \xrightarrow{\boldsymbol{B}} \boldsymbol{Ba}_j.$$

A $\boldsymbol{BA}$ lineáris leképezés mátrixa tehát $\boldsymbol{BA} = (\boldsymbol{Ba}_1, \boldsymbol{Ba}_2, \ldots, \boldsymbol{Ba}_m)$, mérete: $n \times m$.

Két mátrix szorzata tehát úgy képezendő, hogy az első mátrixot szorzom a második mátrix egyes oszlopvektoraival, és az eredménymátrixot a kapott oszlopvektorokból állítom össze. Vonjuk össze a műveleti szabályt a mátrix és vektor szorzására vonatkozó szabállyal, és rögzítsük végleges alakban a mátrixok szorzási szabályát.

A $\boldsymbol{B}$ mátrix és az $\boldsymbol{A}$ mátrix akkor és csak akkor szorozható össze, ha $\boldsymbol{B}$ oszlopmérete megegyezik $\boldsymbol{A}$ sorméretével:

$$\begin{array}{cc} \boldsymbol{A}\cdot\boldsymbol{B} & \boldsymbol{AB} \\ n\times \underbrace{p\quad q}_{p=q}\times m & n\times m \end{array}$$

(az „érintkező” méretek megegyeznek, és kiesnek). Az eredménymátrix i-edik sorának j-edik eleme pedig az első mátrix i-edik sorának és a második mátrix j-edik oszlopának skalárszorzata („sor-oszlop szorzás”).

A mátrix-vektor szorzást külön megtanulni nem kell, tekintsük az oszlopvektort $n \times 1$-es mátrixnak, és alkalmazzuk a mátrixszorzást.

Ha figyelmesen olvastuk a lineáris leképezések mátrixelőállítását, akkor azt láttuk, hogy minden lineáris leképezés felírható alkalmas mátrixszal történő szorzással. A fordított állítást azonban nem vizsgáltuk: az $\boldsymbol{Ax}$ alakú leképezés tényleg lineáris leképezés-e? Az $\boldsymbol{A}c\boldsymbol{x} = c\boldsymbol{Ax}$ összefüggés bizonyítása nem okoz gondot, hiszen a skalárszorzatból a konstans mindig kiemelhető. Az $\boldsymbol{A}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \boldsymbol{Ax}_1 + \boldsymbol{Ax}_2$ bizonyítása a skalárszorzat disztributív tulajdonságán alapszik. Legyenek $\boldsymbol{A}$ sorvektorai $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\ldots$, $\boldsymbol{b}_n$, akkor

$$\boldsymbol{A}(\boldsymbol{x}_1+\boldsymbol{x}_2) = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{pmatrix} (\boldsymbol{x}_1+\boldsymbol{x}_2) = \begin{pmatrix} \langle \boldsymbol{b}_1, \boldsymbol{x}_1+\boldsymbol{x}_2\rangle \\ \langle \boldsymbol{b}_2, \boldsymbol{x}_1+\boldsymbol{x}_2\rangle \\ \vdots \\ \langle \boldsymbol{b}_n, \boldsymbol{x}_1+\boldsymbol{x}_2\rangle \end{pmatrix} =$$

$$= \begin{pmatrix} \langle \boldsymbol{b}_1, \boldsymbol{x}_1\rangle + \langle \boldsymbol{b}_1, \boldsymbol{x}_2\rangle \\ \langle \boldsymbol{b}_2, \boldsymbol{x}_1\rangle + \langle \boldsymbol{b}_2, \boldsymbol{x}_2\rangle \\ \vdots \\ \langle \boldsymbol{b}_n, \boldsymbol{x}_1\rangle + \langle \boldsymbol{b}_n, \boldsymbol{x}_2\rangle \end{pmatrix} =$$

$$= \begin{pmatrix} \langle \boldsymbol{b}_1, \boldsymbol{x}_1\rangle \\ \langle \boldsymbol{b}_2, \boldsymbol{x}_1\rangle \\ \vdots \\ \langle \boldsymbol{b}_n, \boldsymbol{x}_1\rangle \end{pmatrix} + \begin{pmatrix} \langle \boldsymbol{b}_1, \boldsymbol{x}_2\rangle \\ \langle \boldsymbol{b}_2, \boldsymbol{x}_2\rangle \\ \vdots \\ \langle \boldsymbol{b}_n, \boldsymbol{x}_2\rangle \end{pmatrix} = \boldsymbol{Ax}_1 + \boldsymbol{Ax}_2.$$

Ezzel a gondolatmenettel egyben bebizonyítottuk a mátrixszorzás $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$ alakú disztributív tulajdonságát is.

Van azonban egy bökkenő: a mátrixszorzás kommutatív tulajdonsága nem igaz, még akkor sem, ha mindkét szorzás elvégezhető, és még akkor sem, ha mindkét mátrix $n\times n$-es mátrix. Erre nagyon könnyű példát mutatni (lásd pl. az 5. Kérdést). Nem kommutatív szorzás esetén azonban a $(\boldsymbol{B} + \boldsymbol{C})\boldsymbol{A} =$

$= \boldsymbol{BA} + \boldsymbol{CA}$ alakú disztributív tulajdonságot külön meg kell mutatni. Elég a $(\boldsymbol{B} + \boldsymbol{C})\boldsymbol{x} = \boldsymbol{Bx} + \boldsymbol{Cx}$ összefüggést bizonyítani. Legyen $\boldsymbol{B} = (b_{ij})$ és $\boldsymbol{C} = (c_{ij})$, és $(\boldsymbol{B} + \boldsymbol{C})\boldsymbol{x}$ i-edik koordinátája

$$\sum_j (b_{ij} + c_{ij})x_j = \sum_j b_{ij}x_j + \sum_j c_{ij}x_j,$$

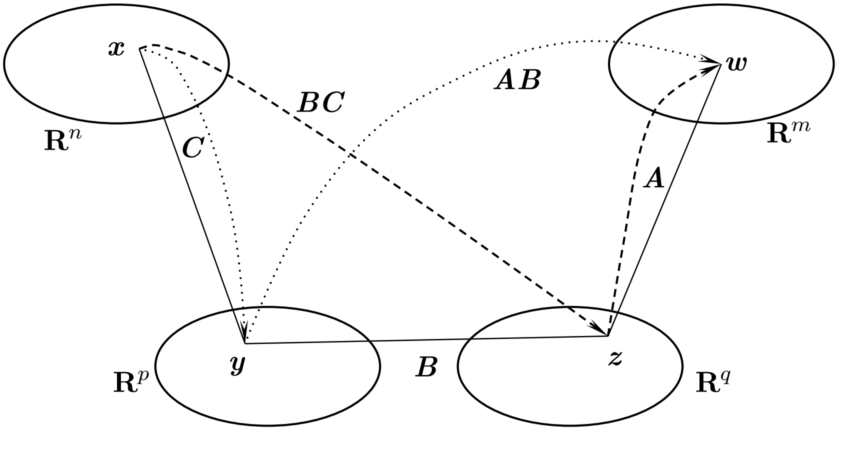
ami $\boldsymbol{Bx}$ és $\boldsymbol{Cx}$ i-edik koordinátájának az összege.

Szabad-e egyáltalán szorzásnak nevezni a mátrixszorzást, ha nem kommutatív? A skalárszorzatnál is előfordult hasonló, ott az asszociativitás hiányzott. Elfogadott nézet az, hogy egy tulajdonság hiányát el lehet nézni. Az asszociatív tulajdonság teljesülését azonban elvárjuk, és mint látni fogjuk, teljesül is.

5. Kérdés. Mint említettük, általában $\boldsymbol{AB} \neq \boldsymbol{BA}$. Előfordulhat-e, hogy $\boldsymbol{AB} = \boldsymbol{C} \neq \mathbf{0}$, és $\boldsymbol{BA} = 2\boldsymbol{C}$? Vizsgáljuk meg a kérdést az $\boldsymbol{A} = \begin{pmatrix} 6 & -4 \\ 9 & -6 \end{pmatrix}$ és a $\boldsymbol{B} = \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix}$ mátrixok vonatkozásában! Előfordulhat-e, hogy $\boldsymbol{AB} = \mathbf{0}$, de $\boldsymbol{BA} \neq \mathbf{0}$? Adjunk rá példát 2×2-es mátrixokkal! Könnyű példát adni arra is, hogy $\mathbf{0}$-tól különböző mátrix négyzete $\mathbf{0}$. ($\mathbf{0}$ a megfelelő méretű nullmátrixot jelöli, mely csupa nullából áll.)

A mátrixszorzás asszociatív tulajdonsága a leképezésekből azonnal látható. Legyen

$$\boldsymbol{C} : \mathbf{R}^n \to \mathbf{R}^p,$$
$$\boldsymbol{B} : \mathbf{R}^p \to \mathbf{R}^q,$$
$$\text{és } \boldsymbol{A} : \mathbf{R}^q \to \mathbf{R}^m.$$



Az $\boldsymbol{ABCx}$ kifejezés ($\boldsymbol{x} \in \mathbf{R}^n$) a lineáris leképezések egymás utáni elvégzését jelenti. Az $\boldsymbol{A}(\boldsymbol{BC})\boldsymbol{x}$ a $\boldsymbol{B}$ és a $\boldsymbol{C}$ összevonásával, az $(\boldsymbol{AB})\boldsymbol{Cx}$ az $\boldsymbol{A}$ és a $\boldsymbol{B}$

összevonásával keletkező leképezéseket jelenti, de az eredmény ugyanaz: az ábra szerint a szaggatott vonal ugyanoda vezet, mint a pontozott vonal.

Megállapítható tehát, hogy $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$.

2.4. Transzponálás

A mátrix transzponáltja a sorok és oszlopok felcseréléséből adódó új mátrix. Képlettel megadva, ha $\boldsymbol{A} = (a_{ij})$ $n \times m$-es mátrix, akkor $\boldsymbol{A}^* = (b_{ij})$, az $\boldsymbol{A}$ transzponáltja, olyan $m \times n$-es mátrix, melyre $b_{ij} = a_{ji}$ (i= 1, 2, ..., m; $j = 1, 2, \ldots, n$). Nyilván

$$(\boldsymbol{A}^*)^* = \boldsymbol{A}.$$

Szemléletesen, ha a mátrixot olyan táblázatnak képzeljük el, amelyhez sorfejléc és oszlopfejléc tartozik, akkor a két fejlécet felcserélve az adatok is helyet változtatnak: a mátrixból a transzponáltját kapjuk.

Nem szorulnak bizonyításra az alábbi, lineáris műveletekre vonatkozó, műveleti szabályok, hiszen a transzponált képzése alapján nyilvánvalóak:

$$(\boldsymbol{A} + \boldsymbol{B})^* = \boldsymbol{A}^* + \boldsymbol{B}^* \text{ és } (c\boldsymbol{A})^* = c\boldsymbol{A}^*.$$

Azt gondolnánk, hogy a transzponálás a mátrix külalakjával történő játszadozás, és nem sok köze van a lineáris leképezésekhez. Ez tévedés, éppoly szerves része a transzformációkkal végzett műveleteknek, mint a többi, korábban megismert művelet.

2.4.1. Definíció. Ha $\boldsymbol{Ax} : \mathbf{R}^m \to \mathbf{R}^n$ leképezés, akkor $\boldsymbol{A}^*\boldsymbol{y} : \mathbf{R}^n \to \to \mathbf{R}^m$ lineáris leképezést létesít, amit a leképezés *transzponáltjának* nevezünk.

A transzponálás egyik fontos, meghatározó tulajdonsága, hogy a skalárszorzatot – bizonyos értelemben megtartja.

2.4.2. Tétel. Legyen $\boldsymbol{A}$ $n \times m$-es mátrix, $\boldsymbol{x} \in \mathbf{R}^m$ és $\boldsymbol{y} \in \mathbf{R}^n$, akkor $\langle \boldsymbol{Ax}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{A}^*\boldsymbol{y} \rangle$. Továbbá, ha minden $\boldsymbol{x} \in \mathbf{R}^m$-re és minden $\boldsymbol{y} \in \mathbf{R}^n$-re fennáll, hogy $\langle \boldsymbol{Ax}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{By} \rangle$, akkor $\boldsymbol{B} = \boldsymbol{A}^*$.

Bizonyítás. 1. Jelölje $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ az $\mathbf{R}^m$, $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n$ az $\mathbf{R}^n$ koordináta egységvektorait, és legyen $\boldsymbol{A}^* = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n) = (b_{ij})$. Az $\langle \boldsymbol{Ax}, \boldsymbol{y} \rangle$ és az $\langle \boldsymbol{x}, \boldsymbol{A}^*\boldsymbol{y} \rangle$ kifejezés $\boldsymbol{x}$-ben is és $\boldsymbol{y}$-ban is lineáris függvény, ezért elég az $\boldsymbol{e}_i$ ill. $\boldsymbol{f}_j$ vektorokra bizonyítani az összefüggést, az $\boldsymbol{x}$-re és $\boldsymbol{y}$-ra vonatkozó állítás ebből már felépíthető.

$$\langle \boldsymbol{Ae}_i, \boldsymbol{f}_j \rangle = \langle \boldsymbol{a}_i, \boldsymbol{f}_j \rangle = a_{ji}, \quad \text{másrészt} \quad \langle \boldsymbol{e}_i, \boldsymbol{A}^*\boldsymbol{f}_j \rangle = \langle \boldsymbol{e}_i, \boldsymbol{b}_j \rangle = b_{ij}, \qquad (*)$$

és $a_{ji} = b_{ij}$ teljesül a transzponálás miatt.

2. Ha $\langle \boldsymbol{Ax},\ \boldsymbol{y}\rangle = \langle \boldsymbol{x},\ \boldsymbol{By}\rangle$ minden koordináta egységvektorra teljesül, akkor (*) miatt $\boldsymbol{B} = \boldsymbol{A}^*$. $\square$

Írjuk fel az $\langle \boldsymbol{ABx},\ \boldsymbol{y}\rangle$ kifejezést (tegyük fel, hogy a méretek megengedik a szorzások elvégzését), és a 2.4.2. Tétel kétszeri alkalmazásával alakítsuk át a kifejezést:

$$\langle \boldsymbol{ABx}, \boldsymbol{y}\rangle = \langle \boldsymbol{Bx}, \boldsymbol{A}^*\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{B}^*\boldsymbol{A}^*\boldsymbol{y}\rangle,$$

majd a 2.4.2. Tétel második része alapján vonjuk le a következtetést:

$$(\boldsymbol{AB})^* = \boldsymbol{B}^*\boldsymbol{A}^*.$$

Ez a szorzat transzponáltjára vonatkozó műveleti szabály.

6. Kérdés. Adott $\boldsymbol{A}$ $n \times m$-es mátrix és $\boldsymbol{a} \in \mathbf{R}^n$ esetén az $\langle \boldsymbol{a},\ \boldsymbol{Ax}\rangle$ skalárszorzat lineáris funkcionál $\mathbf{R}^m$-en. Mint minden lineáris funkcionál, ez is felírható egy adott vektorral képezett skalárszorzatként (ld. 2.2.2. Tétel). Hogyan adható meg jelen esetben ez a vektor?

2.5. Dimenzió-tétel, rangszám-tétel

Ebben a fejezetben – a címtől eltérően – a lineáris leképezések szürjektív, injektív és bijektív tulajdonságait vizsgáljuk.

Emlékeztetünk arra, hogy az $f\colon X \to Y$ függvény (ahol X és Y tetszőleges halmazok) szürjektív, ha értékkészlete a teljes Y; injektív, ha különböző X-beli pontok képe különböző, bijektív, ha szürjektív és injektív is.

Hogy jobban rögzüljön a három fogalom, végezzünk el egy gondolatkísérletet: Y minden pontjához írjuk oda, hogy hány X-beli ponthoz van hozzárendelve. Ez a jelzés lehet akár 0 is, vagy végtelen is. Szürjektív a függvény, ha a számok között nincs nulla, injektív, ha csak 0 vagy 1 szerepel, bijektív esetben minden ponthoz 1-et írtunk. A bijektív függvény nyilván kölcsönösen egyértelmű leképezést valósít meg.

A szürjektív tulajdonság a 2.1. fejezetben mondottak értelmében egyszerűen kezelhető: az R_A értékkészlet, mint altér, dimenziója $\operatorname{rang}(\boldsymbol{A})$, tehát $R_A = \mathbf{R}^n$ akkor és csak akkor valósul meg, ha $\operatorname{rang}(\boldsymbol{A}) = n$. Ez a szürjektivitás szükséges és elégséges feltétele.

Például $\boldsymbol{Ax}\colon \mathbf{R}^m \to \mathbf{R}^n$ $m < n$ esetén nem lehet szürjektív, mert

$$\operatorname{rang}(\boldsymbol{A}) = \operatorname{rang}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m) \le m < n.$$

Az injektivitás vizsgálatához definiáljuk a magtér fogalmát.

2.5.1. Definíció. Az $\boldsymbol{Ax}\colon \mathbf{R}^m \to \mathbf{R}^n$ leképezés M_A *magtere* azon $\boldsymbol{x}$ vektorok halmaza, melyek a $\mathbf{0}$-ra képződnek le: $M_A = \{\boldsymbol{x}\colon\ \boldsymbol{Ax} = \mathbf{0}\}$.

Az M_A magtér nyilván altere $\mathbf{R}^m$-nek.

2.5.2. Tétel. Az $\boldsymbol{Ax}$: $\mathbf{R}^m \to \mathbf{R}^n$ leképezés akkor és csak akkor injektív, ha $\boldsymbol{A}$ magtere egy elemű, vagyis ha $M_A = \{\mathbf{0}\}$.

Bizonyítás. $\Rightarrow$ Tegyük fel, hogy $\boldsymbol{a} \in M_A$ és $\boldsymbol{a} \neq \mathbf{0}$. Legyen $\boldsymbol{y} = \boldsymbol{Ax}$ valamely adott $\boldsymbol{x}$-re, akkor $\boldsymbol{A}(\boldsymbol{x} + \boldsymbol{a}) = \boldsymbol{y} + \mathbf{0} = \boldsymbol{y}$, tehát nem lehet injektív.

$\Leftarrow$ Ha nem lenne injektív, akkor $\boldsymbol{y} = \boldsymbol{Ax}_1 = \boldsymbol{Ax}_2$ állna fenn, de $\mathbf{0} = \boldsymbol{Ax}_1 - \boldsymbol{Ax}_2 = \boldsymbol{A}(\boldsymbol{x}_1 - \boldsymbol{x}_2)$, vagyis $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in M_A$, de $M_A = \{\mathbf{0}\}$, tehát $\boldsymbol{x}_1 = \boldsymbol{x}_2$. □

2.5.3. Tétel (dimenzió tétel). Jelöljük M_A dimenzióját s-sel, R_A dimenzióját r-rel, akkor

$$s + r = m.$$

Bizonyítás. Vegyünk fel M_A-ban egy bázist, jelöljük $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s$-sel, és ezt egészítsük ki $\mathbf{R}^m$-ben bázissá (1.5.2. kibővítési tétel), legyen ez $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s, \ldots, \boldsymbol{x}_m$. R_A minden eleme

$$\boldsymbol{A} \cdot \sum_{k=1}^{m} a_k \boldsymbol{x}_k = \sum_{k=s+1}^{m} a_k \ \boldsymbol{Ax}_k$$

alakba írható, tehát az $\boldsymbol{Ax}_{s+1}, \boldsymbol{Ax}_{s+2}, ..., \boldsymbol{Ax}_m$ vektorok generálják R_A-t. Ugyanakkor lineárisan függetlenek is, mert

$$\begin{aligned} c_{s+1}\boldsymbol{Ax}_{s+1} + c_{s+2}\boldsymbol{Ax}_{s+2} + ... + c_m\boldsymbol{Ax}_m = \\ = \boldsymbol{A}(c_{s+1}\boldsymbol{x}_{s+1} + c_{s+2}\boldsymbol{x}_{s+2} + ... + c_m\boldsymbol{x}_m) = 0 \end{aligned}$$

alapján $c_{s+1}\boldsymbol{x}_{s+1} + c_{s+2}\boldsymbol{x}_{s+2} + ... + c_m\boldsymbol{x}_m \in M_A$. Az $\boldsymbol{x}_{s+1}, \boldsymbol{x}_{s+2}, ..., \boldsymbol{x}_m$ vektorok által generált altér az M_A kiegészítő altere, tehát közös elemük csak a $\mathbf{0}$ lehet, így $c_{s+1}\boldsymbol{x}_{s+1} + c_{s+2}\boldsymbol{x}_{s+2} + ... + c_m\boldsymbol{x}_m = 0$, ami csak $c_{s+1} = c_{s+2} = ... = c_m = 0$ esetén lehetséges.

Mivel megadtuk R_A-nak $m - s$ darab bázisvektorát, R_A dimenziója $r = m - s$. □

A tételnek érdekes következménye van $m = n$ esetén. A lineáris leképezések speciális esetét, amikor $m = n$, vagyis amikor az $\mathbf{R}^n$-et önmagára képezzük le, külön névvel illetjük: *lineáris transzformációnak* nevezzük.

2.5.4. Következmény. Az $\boldsymbol{Ax}$: $\mathbf{R}^n \to \mathbf{R}^n$ lineáris transzformációra a szürjektivitás, injektivitás és a bijektivitás ekvivalens fogalmak.

Bizonyítás. A dimenzió tétel miatt, ha injektív, azaz $s = 0$, akkor $r = n$, tehát szürjektív is. Ha szürjektív, azaz $r = n$, akkor $s = 0$, tehát injektív is. □

2.5.5. Tétel (rangszám-tétel). Bármely lineáris leképezés és a transzponáltjának a rangja megegyezik. Mátrixokra kimondva: az $\boldsymbol{A}$ és $\boldsymbol{A}^*$ mátrix rangja egyenlő.

Bizonyítás. Legyen $\boldsymbol{x} \in M_A$, akkor $\boldsymbol{Ax} = \boldsymbol{0}$. Bármely $\boldsymbol{y} \in \mathbf{R}^n$-re $\langle \boldsymbol{y}, \boldsymbol{Ax} \rangle = \langle \boldsymbol{x}, \boldsymbol{A}^*\boldsymbol{y} \rangle = 0$, tehát $\boldsymbol{x}$ ortogonális R_{A^*} minden elemére. R_{A^*} ezért az M_A ortogonális kiegészítő alterébe esik, vagyis $r_1 = \mathrm{rang}(\boldsymbol{A}^*) \leq$ $\leq m - s = r$.

A gondolatmenetet elismételve úgy, hogy $\boldsymbol{A}^*\boldsymbol{x}$ leképezésből indulunk ki, kapjuk, hogy $r \leq r_1$. A két egyenlőtlenség alapján $r = r_1$. □

2.5.6. Következmény. $\boldsymbol{A}^*$ képtere megegyezik $\boldsymbol{A}$ magterének ortogonális kiegészítő alterével. Ha M_A ortogonális kiegészítő alterét N_A -val jelöljük, akkor $N_A = R_{A^*}$.

Bizonyítás. A 2.5.5. Tétel bizonyításában láttuk, hogy $N_A \supset R_{A^*}$, de dimenzióban megegyeznek, tehát $N_A = R_{A^*}$. □

A rangszám-tétel következménye, hogy az $n \times m$-es mátrix rangja legfeljebb $\min(n, m)$ lehet. Azokat a mátrixokat, melyekre itt egyenlőség teljesül, *teljesrangú mátrixoknak* nevezzük.

2.5.7. Tétel. $\mathrm{rang}(\boldsymbol{AB}) \leq \min(\mathrm{rang}(\boldsymbol{A}), \mathrm{rang}(\boldsymbol{B}))$.

Bizonyítás. Mivel $R_{AB} \subset R_A$, R_{AB} dimenziója nem lehet nagyobb, mint R_A dimenziója, ezért $\mathrm{rang}(\boldsymbol{AB}) \leq \mathrm{rang}(\boldsymbol{A})$. Másrészt a rangszám-tétel felhasználásával $\mathrm{rang}(\boldsymbol{AB}) = \mathrm{rang}((\boldsymbol{AB})^*) = \mathrm{rang}(\boldsymbol{B}^*\boldsymbol{A}^*) \leq \mathrm{rang}(\boldsymbol{B}^*) =$ $\mathrm{rang}(\boldsymbol{B})$. □

7. Kérdés. Adjuk meg a 2. Kérdésben szereplő lineáris transzformáció magterét!

2.6. Rangszámítás

A lineáris leképezések egyik fő jellemzője a rang. A rang kiszámítása azonban jól használható a lineáris függetlenség eldöntésére és a lineárisan független vektorok számának meghatározására is. Az alapgondolat az, hogy az oszlopokkal bizonyos műveleteket végezve megmutatható, hogy a mátrix rangja nem változik, és közben a mátrix áttekinthetőbbé válik. Mivel $\mathrm{rang}(\boldsymbol{A}) =$ $= \mathrm{rang}(\boldsymbol{A}^*)$, természetesen ugyanezen műveleteket sorokra is elvégezhetjük.

2.6.1. Segédtétel. Ha $\boldsymbol{A}$ $n \times m$-es, r rangú mátrix, $\boldsymbol{B}$ $m \times p$-s mátrix, melynek a rangja m, akkor $\mathrm{rang}(\boldsymbol{AB}) = r$.

Bizonyítás.. Mivel $\boldsymbol{Bz}$ értékkészlete $\mathbf{R}^m$, az $\boldsymbol{ABz}$ értékkészlete R_A, ezért $R_{AB} = R_A$, vagyis $\mathrm{rang}(\boldsymbol{AB}) = r$. □

Adjuk meg az $m \times m$-es $\boldsymbol{B}$ mátrixot a következőképpen:

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ c & 1 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Az oszlopvektorok lineáris függetlensége a definíció alapján igazolható, tehát $\operatorname{rang}(\boldsymbol{B}) = m$. Ebből következően $\operatorname{rang}(\boldsymbol{AB}) = \operatorname{rang}(\boldsymbol{A})$, vagyis az $\boldsymbol{A}$ mátrix rangja nem változik, ha $\boldsymbol{B}$-vel jobbról megszorozzuk. Legyen $\boldsymbol{A} = (\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \boldsymbol{a}_3,\ \dots,\ \boldsymbol{a}_m)$, akkor

$$\boldsymbol{AB} = \begin{pmatrix} a_{11}+ca_{12} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21}+ca_{22} & a_{22} & a_{23} & \cdots & a_{2m} \\ \vdots & & & \ddots & \vdots \\ a_{n1}+ca_{n2} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{pmatrix} = (\boldsymbol{a}_1 + c\boldsymbol{a}_2, \boldsymbol{a}_2, \boldsymbol{a}_3, ..., \boldsymbol{a}_m)$$

rangja megegyezik $\boldsymbol{A}$ rangjával. Ezzel bebizonyítottuk a következő tételt.

2.6.2. Tétel. A mátrix rangja nem változik, ha valamelyik oszlopához egy másik oszlop c-szeresét, vagy ha egy sorához egy másik sor c-szeresét adjuk hozzá.

Ezzel a tétellel megnyílik az út a mátrixok rangjának numerikus kiszámítására. A sor- és oszlopműveletek sorozatos alkalmazásával annyi nullát „gyártunk" a mátrixban, amennyit csak lehet. Ezt azonban célszerű módszeresen végezni, mert különben néhány lépés után a korábban előállított nullák elromlanak, és a nullák száma nem növekszik. Általában a következő eljárás javasolható.

Válasszunk ki egy nem nulla elemet, a_{ij}-t, amelynek mondjuk a sorát nullázni akarjuk. A j-edik oszlop megfelelő többszörösét hozzáadva az első oszlophoz elérhető, hogy a_{i1} nullává váljon, és ugyanezt megtehetjük a többi oszloppal is a j-edik kivételével. Egy elemet sorban magányosnak nevezünk, ha rajta kívül a sor minden eleme nulla, oszlopban magányosnak, ha az oszlop többi eleme nulla. Az eljárással tehát a_{ij} sorban magányossá vált. Ez után a_{ij}-vel az oszlopának többi eleme is nullázható, közben a többi elem nem változik, mert 0-t adunk hozzájuk. Foglaljuk össze: sorban magányos elem oszlopban is magányos, oszlopban magányos elem sorban is magányos, mondjuk úgy, hogy teljesen magányos. Az eljárást addig lehet folytatni, amíg minden nem nulla elem teljesen magányossá válik. Ez a végállás. Ekkor közvetlenül látható, hogy a magányos elemeket tartalmazó oszlopok lineárisan függetlenek, míg a csupa nulla oszlopok mindig lineárisan függővé teszik a rendszert, tehát a kapott mátrix rangja, ami megegyezik az eredeti mátrix rangjával, a magányos elemek száma.

Az eljárás értelemszerűen felhasználható egy adott vektorrendszer rangjának a kiszámítására is, csak a vektorokból, mint oszlopvektorokból mátrixot kell készíteni. Vannak azonban olyan esetek, amikor nem elég a lineárisan független vektorok számát megállapítani, hanem ezeket ki is kell választani a vektorrendszerből. Az eljárás – bizonyos szabályok beiktatásával – erre is lehetőséget ad.

Nevezzük az oszlopműveletek kapcsán aktív oszlopnak azt, amelyik többszörösét képezve hozzáadjuk egy másikhoz. Az aktív oszlop a művelet során változatlan marad. A másik oszlopot, amelyikhez ezt hozzáadjuk, amelyik „eltűri" a hozzáadást, nevezzük passzív oszlopnak. Vezessük be az alábbi korlátozásokat az oszlopműveletek tekintetében (a sorműveletekre itt nincs korlátozás):

1. Oszlopcserét nem hajtunk végre.
2. Egyik aktív oszlopvektor sem válik a későbbiek során nullvektorrá.

Ha ez a két szabály nem sérül, akkor a teljesen magányos elemek kijelölik a lineárisan független vektorokat az eredeti mátrixban is.

Bizonyítás. A rangszámolás végállásának ismeretében az eredeti mátrixból töröljük a végállásban csupa nullává váló oszlopokat. A rangszámolás korábbi lépései erre a mátrixra is megismételhetők (a törölt passzív oszlopokkal nem kell törődni). A végeredmény megadja a lineárisan független vektorok számát, de minden oszlopban lesz magányos elem, így ezek a vektorok lineárisan függetlenek. □

A fent javasolt eljárás ezt a szabályt betartja. Természetesen a sorvektorokra is hasonló állítás érvényes.

8. Kérdés. Lineárisan független-e az $\boldsymbol{a} = (1, 4, 0, 0)$, $\boldsymbol{b} = (0, 2, 3, 0)$, $\boldsymbol{c} = (0, 0, 3, 2)$, $\boldsymbol{d} = (1, 0, 0, 4)$ vektorrendszer? Ha nem, akkor adjuk meg egy maximális lineárisan független részrendszerét!

2.7. Inverz leképezés, inverz mátrix

Az $f\colon X \to Y$ függvénynek akkor és csak akkor van inverze, ha bijektív. Az $\boldsymbol{Ax}\colon \mathbf{R}^m \to \mathbf{R}^n$ lineáris leképezés rangját jelöljük r-rel, magterének a dimenzióját s-el. Ha az $\boldsymbol{A}$ bijektív, akkor a szürjektivitás miatt $n =$ $= r$, az injektivitás miatt $s = 0$, tehát a 2.5.3. dimenzió-tétel (mely szerint $s+r = m$) alapján $n = m$ áll fenn. Ebben a fejezetben tehát a továbbiakban feltételezzük, hogy mindig $\boldsymbol{Ax}\colon \mathbf{R}^n \to \mathbf{R}^n$ *lineáris transzformációról* van szó, mert más esetben az inverz nem létezhet. Az $\boldsymbol{Ax}\colon \mathbf{R}^n \to \mathbf{R}^n$ lineáris transzformáció inverze akkor és csak akkor létezik, ha $n = r$, vagyis, ha az $\boldsymbol{A}$ teljesrangú.

Az $\boldsymbol{A}$ mátrix tehát $n \times n$-es, ún. négyzetes mátrix, rangja $r = n$, vagyis teljesrangú mátrix. Ezen feltételek mellett a bijektivitás biztosítva van, tehát az inverz leképezés létezik. Könnyen látható, hogy az inverz transzformáció is lineáris. Legyen ugyanis $\boldsymbol{y}_1$ és $\boldsymbol{y}_2$ a tér két tetszőleges eleme, akkor $\boldsymbol{y}_1 = \boldsymbol{A}\boldsymbol{x}_1$ és $\boldsymbol{y}_2 = \boldsymbol{A}\boldsymbol{x}_2$, és az inverz transzformáció $\boldsymbol{y}_1$-hez $\boldsymbol{x}_1$-et, $\boldsymbol{y}_2$-höz $\boldsymbol{x}_2$-t rendeli hozzá. $\boldsymbol{y}_1 + \boldsymbol{y}_2 = \boldsymbol{A}\boldsymbol{x}_1 + \boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{A}(\boldsymbol{x}_1 + \boldsymbol{x}_2)$, vagyis $\boldsymbol{y}_1 + \boldsymbol{y}_2$-höz $\boldsymbol{x}_1 + \boldsymbol{x}_2$ tartozik, míg $c\boldsymbol{y}_1 = c\boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{A}c\boldsymbol{x}_1$, vagyis $c\boldsymbol{y}_1$-hez $c\boldsymbol{x}_1$ tartozik. Jelöljük az inverz lineáris transzformációt és a mátrixát $\boldsymbol{A}^{-1}$-gyel, akkor $\boldsymbol{A}^{-1}$: $\mathbf{R}^n \to \mathbf{R}^n$, és a mátrixa szintén $n \times n$-es négyzetes mátrix.

9. Kérdés. Milyen a-ra nem létezik az $\boldsymbol{A}$ mátrixú transzformáció inverze, ahol

$$\boldsymbol{A} = \begin{pmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{pmatrix}?$$

Ha az f függvénynek létezik inverze, jelöljük szokásos módon f^{-1}-gyel, akkor $f^{-1}(f(x)) = x$ és $f(f^{-1}(y)) = y$, hiszen az oda-vissza történő leképezés minden elemet helyben hagy. Inverz leképezésre felírva ugyanezt:
$\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}$, ill. $\boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{x} = \boldsymbol{x}$, mátrixokra felírva $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$, ill. $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$, ahol $\boldsymbol{I}$ az identikus leképezés mátrixa,

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Az $\boldsymbol{I}$-t *egységmátrixnak* nevezzük.

A valós számok bevezetésénél szó esik a összeadás és a szorzás műveleti szabályairól. A kommutatív, asszociatív és disztributív szabályokon túlmenően megemlítjük, hogy mindkét műveletnek van neutrális eleme: $a + 0 = a$, és $a \cdot 1 = a$, tehát az összeadásnál a 0, a szorzásnál az 1 a másik operandust nem változtatja meg. A neutrális elem ismeretében létezik mindkét műveletre az inverz elem, olyan x szám, melyre $a + x = 0$, ill. ha $a \neq 0$, akkor $ax = a$. Ezeket jelöljük $-a$-val ill. $1/a$-val. Hasonlóan a mátrixszorzásnál $\boldsymbol{I}$ a neutrális elem, egységelem, mert $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$, és $\boldsymbol{A}^{-1}$ az inverz elem, mert $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$.

Az inverz transzformáció és az inverz mátrix fogalma az inverz függvény definíciójából következik. Megmutatjuk azonban, hogy az $\boldsymbol{AX} = \boldsymbol{XA} = \boldsymbol{I}$ összefüggés jellemző tulajdonság, azonban a definíció egyszerűbb alakra hozható.

2.7.1. Definíció. Ha négyzetes mátrixokra $\boldsymbol{AB} = \boldsymbol{I}$, akkor mindkét mátrix inverze létezik és $\boldsymbol{B}$-t az $\boldsymbol{A}$ *inverzének* és $\boldsymbol{A}$-t a $\boldsymbol{B}$ *inverzének* nevezzük. Jelöléssel: $\boldsymbol{B} = \boldsymbol{A}^{-1}$ és $\boldsymbol{A} = \boldsymbol{B}^{-1}$.

Bizonyítás. Bizonyítani kell a definíció egyértelműségét. Ha $\boldsymbol{AB} = \boldsymbol{I}$, akkor $\boldsymbol{A}$ is és $\boldsymbol{B}$ is teljesrangú mátrix. $\boldsymbol{I}$ ugyanis teljesrangú, képtere $\mathbf{R}^n$, ez csak úgy állhat elő, ha R_B is és R_A is megegyezik $\mathbf{R}^n$-nel, ekkor viszont mindkét mátrix teljesrangú. Teljesrangú négyzetes mátrixokra az inverz viszont létezik. Szorozzuk meg az $\boldsymbol{AB} = \boldsymbol{I}$ egyenletet balról $\boldsymbol{A}^{-1}$-gyel, akkor $\boldsymbol{A}^{-1}(\boldsymbol{AB}) = (\boldsymbol{A}^{-1}\boldsymbol{A})\ \boldsymbol{B} = \boldsymbol{B} = \boldsymbol{A}^{-1}$, adódik. Jobb oldali $\boldsymbol{B}^{-1}$ szorzással ugyanígy adódik a másik állítás. □

Még egyszer mondjuk el a bizonyított állítás lényegét: négyzetes mátrixok vonatkozásában elég az $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$, ill. $\boldsymbol{AA}^{-1} = \boldsymbol{I}$ egyenlőségek egyikét igazolni ahhoz, hogy belássuk, hogy $\boldsymbol{A}^{-1}$ tényleg az inverz mátrix. Ha az egyik egyenlőséget tudjuk, a másik már biztosan teljesül.

10. Kérdés. A fenti definícióban lényeges-e a négyzetes mátrixokra vonatkozó kikötés? Ha $\boldsymbol{A}$ 2×3-as és $\boldsymbol{B}$ 3×2-es akkor előfordulhat-e, hogy $\boldsymbol{AB} = \boldsymbol{I}$? Ha előfordulhat, akkor $\boldsymbol{A}$ egyértelműen meghatározza-e $\boldsymbol{B}$-t? Mi a helyzet, ha $\boldsymbol{A}$ 3×2-es?

2.7.2. Tétel. Ha $\boldsymbol{A}$ és $\boldsymbol{B}$ inverze létezik, akkor $\boldsymbol{AB}$ inverze is létezik és $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, továbbá $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$ $(c \neq 0)$.

Bizonyítás. $\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ teljesíti az $\boldsymbol{AB}$ inverz mátrixának a definícióját, ugyanis

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

$\frac{1}{c}\boldsymbol{A}^{-1}$ ugyancsak teljesíti $c\boldsymbol{A}$ inverzének a definícióját:

$$(\frac{1}{c}\boldsymbol{A}^{-1})(c\boldsymbol{A}) = \frac{1}{c}c\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}. \quad \square$$

2.8. Válaszok a kérdésekre

1. Lineáris leképezés-e az A: $\mathbf{R}^2 \to \mathbf{R}^3$, ha $A(x,y) = (x+y, 2x-y, 0)$? Lineáris leképezés-e az A: $\mathbf{R}^2 \to \mathbf{R}^3$, ha $A(x,\ y) = (x+y,\ 2x-y,\ 1)$?

A lineáris leképezést definiáló relációkat kell ellenőrizni:

$$A(cx, cy) = (cx+cy,\ 2cx-cy,\ 0) = c(x+y,\ 2x-y,\ 0) = cA(x,y),$$

és

$$A(x_1+x_2, y_1+y_2) = (x_1+x_2+y_1+y_2,,\ 2(x_1+x_2)-(y_1+y_2),\ 0) =$$
$$= (x_1+y_1,\ 2x_1-y_1,\ 0) + (x_2+y_2,\ 2x_2-y_2,\ 0) = A(x_1,y_1) + A(x_2,y_2).$$

A második esetben $A(0, 0) = (0, 0, 1) \neq \mathbf{0}$, tehát nem lineáris.

2. Jelöljük az $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ minta átlagát $\bar{x}$-sal, és képezzük az $\boldsymbol{y} = (x_1 - \bar{x}, x_2 - \bar{x}, \ldots, x_n - \bar{x})$ vektort. Adjuk meg annak a lineáris transzformációnak a mátrixát, mely az $\boldsymbol{x}$-et az $\boldsymbol{y}$-ba képezi le! Mennyi a mátrix rangja?

Legyen $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$, akkor $y_k = x_k - \bar{x} = -\frac{1}{n}x_1 - \frac{1}{n}x_2 - ... + \frac{n-1}{n}x_k - ... - \frac{1}{n}x_n$. Ennek megfelelően:

$$\boldsymbol{A} = \begin{pmatrix} \frac{n-1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & \frac{n-1}{n} & & -\frac{1}{n} \\ \vdots & & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & \frac{n-1}{n} \end{pmatrix}.$$

Az R_A altérben lévő $\boldsymbol{y}$ vektorokat az $y_1 + y_2 + \ldots + y_n = 0$ egyenlet jellemzi, az 1. fejezet 10. kérdése alapján ennek dimenziója $n - 1$, azaz $\operatorname{rang}(\boldsymbol{A}) = n - 1$.

3. Hány dimenziós teret alkot az $\mathbf{R}^m \to \mathbf{R}^n$ lineáris leképezések halmaza?

A leképezések halmaza ekvivalens (izomorf) az $n \times m$-es mátrixok halmazával. Ennek bázisát alkotják azok a mátrixok, amelyeknek egyetlen elemük 1, a többi 0. Mivel nm ilyen mátrix van, a tér dimenziója nm.

4. Gyakran szükségessé válik az $\boldsymbol{a} = (a_1, a_2, \ldots, a_m)$ és az $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$ vektor $(a_1x_1, a_2x_2, \ldots, a_mx_m)$ alakú „szorzatát" előállítani (pl. raktárkészlet értéke árufajtánként). Hogyan tehetjük meg az eddigi műveletekkel?

Írjuk fel $\boldsymbol{a}$-t mátrix alakban a következő formában:

$$\boldsymbol{A} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix},$$

akkor $(a_1x_1, a_2x_2, \ldots, a_mx_m) = \boldsymbol{Ax}$.

Felmerül a kérdés, hogy a vektorok szorzatát miért nem így definiáljuk, hiszen sokkal egyszerűbb lenne. Valójában ez lineáris vektor-vektor hozzárendelés, tehát „valódi énje" lineáris transzformáció, ezért célszerűbb az ennek megfelelő $\boldsymbol{Ax}$ jelölést használni.

5. Mint említettük, általában $\boldsymbol{AB} \neq \boldsymbol{BA}$. Előfordulhat-e, hogy $\boldsymbol{AB} = \boldsymbol{C} \neq \mathbf{0}$, és $\boldsymbol{BA} = 2\boldsymbol{C}$? Vizsgáljuk meg a kérdést az $\boldsymbol{A} = \begin{pmatrix} 6 & -4 \\ 9 & -6 \end{pmatrix}$ és a $\boldsymbol{B} = \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix}$ mátrixok vonatkozásában! Előfordulhat-e, hogy $\boldsymbol{AB} = \mathbf{0}$, de $\boldsymbol{BA} \neq \mathbf{0}$? Adjunk rá példát 2×2-es mátrixokkal! Könnyű példát adni arra is, hogy $\mathbf{0}$-tól különböző mátrix négyzete $\mathbf{0}$.

$$\boldsymbol{AB} = \begin{pmatrix} 6 & -4 \\ 9 & -6 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} -6 & 4 \\ -9 & 6 \end{pmatrix} = \boldsymbol{C} \neq 0$$

és

$$\boldsymbol{BA} = \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} 6 & -4 \\ 9 & -6 \end{pmatrix} = \begin{pmatrix} -12 & 8 \\ -18 & 12 \end{pmatrix} = 2\boldsymbol{C}.$$

Legyen most $\boldsymbol{A} = \begin{pmatrix} 1 & b \\ c & d \end{pmatrix}$ és $\boldsymbol{B} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, akkor $\boldsymbol{AB} = \mathbf{0}$ a következő egyenletek teljesülését kívánja meg:

$$\begin{aligned} A + bC &= 0 \\ B + bD &= 0 \\ cA + dC &= 0 \\ cB + dD &= 0. \end{aligned}$$

Az első két egyenlet alapján $A =$ -bC és $B =$ - bD. Ezt behelyettesítve a másik kettőbe, $C \neq 0$-t és $D \neq 0$-t feltételezve mindkettőből $d = bc$ adódik. E három egyenletnek kell eleget tennie a hét számnak. Könnyen találunk megoldást, pl. $b = 2$, $c = 3$, $C = 1$ és $D = 1$ már meghatározza a többi számot. Valóban:

$$\boldsymbol{AB} = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} -2 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

és

$$\boldsymbol{BA} = \begin{pmatrix} -2 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} -8 & -16 \\ 4 & 8 \end{pmatrix} \neq 0.$$

Ha $\boldsymbol{A} = \begin{pmatrix} 0 & 0 \\ a & 0 \end{pmatrix}$, akkor $\boldsymbol{A}^2 = \boldsymbol{AA} = \mathbf{0}$.

6. Adott $\boldsymbol{A}$ $n \times m$-es mátrix és $\boldsymbol{a} \in \mathbf{R}^n$ esetén az $\langle \boldsymbol{a}, \boldsymbol{Ax} \rangle$ skalárszorzat lineáris funkcionál $\mathbf{R}^m$-en. Mint minden lineáris funkcionál, ez is felírható egy adott vektorral képezett skalárszorzatként (ld. 2.2.2. Tétel). Hogyan adható meg jelen esetben ez a vektor?

Mivel $\langle \boldsymbol{a}, \boldsymbol{A}\boldsymbol{x}\rangle = \langle \boldsymbol{A}^*\boldsymbol{a}, \boldsymbol{x}\rangle$ a transzponálás tulajdonsága miatt, a kérdéses vektor $\boldsymbol{A}^*\boldsymbol{a}$.

7. Adjuk meg a 2. Kérdésben szereplő lineáris transzformáció magterét!

$\boldsymbol{y} = \mathbf{0}$ akkor és csak akkor, ha minden i-re $x_i = \bar{x}$, vagyis, ha minden mintaelem egyenlő. Az altér egydimenziós, és generáló eleme $\mathbf{1} = (1, 1, \dots, 1)$.

Más meggondolással a 2. Kérdésben megadott mátrix alapján látjuk, hogy $\boldsymbol{A} = \boldsymbol{A}^*$, tehát $R_A = R_{A^*}$. R_A dimenzióját és generáló elemeit az 1. fejezet 10. Kérdésében meghatároztuk. R_A ortogonális kiegészítő altere egydimenziós, generáló eleme láthatóan az $\mathbf{1}$ vektor.

8. Lineárisan független-e az $\boldsymbol{a} = (1, 4, 0, 0)$, $\boldsymbol{b} = (0, 2, 3, 0)$, $\boldsymbol{c} = (0, 0, 3, 2)$, $\boldsymbol{d} = (1, 0, 0, 4)$ vektorrendszer? Ha nem, akkor adjuk meg egy maximális lineárisan független részrendszerét!

Mátrixok rangban egyenlőségét itt $\sim$-lal jelöljük. Készítsük el az oszlopvektorokból álló mátrixot, és végezzük el az előírt sor- és oszlopműveleteket. Az első oszlop (-1)-szeresét adjuk hozzá az utolsó oszlophoz, ezáltal az első sor első eleme sorban magányossá válik. A sorban magányos elem oszlopában is magányos. A második sort adjuk hozzá a negyedikhez, ekkor a második sor negyedik eleme válik előbb oszlopában, majd teljesen magányossá. A harmadik oszlopból vonjuk ki a másodikat, ekkor a harmadik oszlop nullává válik, majd a harmadik sor második elemét tegyük teljesen magányossá. Ezzel elérkeztünk a végálláshoz: minden nem nulla elem teljesen magányos.

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 4 & 2 & 0 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 2 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 4 & 2 & 0 & -4 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 2 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & -4 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 2 & 4 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & -4 \\ 0 & 3 & 3 & 0 \\ 0 & 2 & 2 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \\ 0 & 3 & 3 & 0 \\ 0 & 2 & 2 & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \\ 0 & 3 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Olvassuk le a rangot: három nem nulla elem maradt, a rang tehát 3, a lineárisan független oszlopvektorok száma (és a sorvektoroké is) három. A megadott vektorok nem lineárisan függetlenek, de – mivel a kiegészítő szabályokat nem hágtuk át, a harmadik oszlop aktív oszlopként nem szerepelt – az $\boldsymbol{a}$, $\boldsymbol{b}$ és $\boldsymbol{d}$ vektorok biztosan lineárisan függetlenek.

A levezetésből közvetlenül nem látszik, de ellenőrizhetően igaz, hogy a vektorok között az $\boldsymbol{a} - 2\boldsymbol{b} + 2\boldsymbol{c} - \boldsymbol{d} = \boldsymbol{0}$ összefüggés áll fenn, ebből az is látszik, hogy bármely három vektor lineárisan független. A sorok közötti összefüggés bonyolultabb: $12\boldsymbol{a} - 3\boldsymbol{b} + 2\boldsymbol{c} - 3\boldsymbol{d} = \boldsymbol{0}$.

9. Milyen a-ra nem létezik az $\boldsymbol{A}$ mátrixú transzformáció inverze, ahol

$$\boldsymbol{A} = \begin{pmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{pmatrix}?$$

Azt kell eldönteni, hogy milyen a-ra igaz, hogy $\operatorname{rang}(\boldsymbol{A}) < 4$. $a = 1$ esetén $\operatorname{rang}(\boldsymbol{A}) = 1$, ezt az esetet mellőzhetjük a további vizsgálatnál. Az első sor (-1)-szeresét adjuk hozzá a második és a harmadik sorhoz, majd $(-a)$-szorosát adjuk hozzá az utolsó sorhoz. $(1-a) \neq 0$ az oszlopokból kiemelhető.

$$\begin{pmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{pmatrix} \sim \begin{pmatrix} a & 1 & 1 & 1 \\ 1-a & a-1 & 0 & 0 \\ 1-a & 0 & a-1 & 0 \\ 1-a^2 & 1-a & 1-a & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1-a & a-1 & 0 & 0 \\ 1-a & 0 & a-1 & 0 \\ 1-a^2 & 1-a & 1-a & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1+a & 1 & 1 & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 3+a & 1 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 3+a & 0 & 0 & 0 \end{pmatrix}.$$

Az utolsó mátrixból látszik, hogy $a =$ -3 esetén is $\operatorname{rang}(\boldsymbol{A}) = 3 < 4$. Más esetben az inverz létezik.

10. A 2.7.1. Definícióban lényeges-e a négyzetes mátrixokra vonatkozó kikötés? Ha $\boldsymbol{A}$ 2×3-as és $\boldsymbol{B}$ 3×2-es akkor előfordulhat-e, hogy $\boldsymbol{AB} = \boldsymbol{I}$?

Ha előfordulhat, akkor $\boldsymbol{A}$ egyértelműen meghatározza-e $\boldsymbol{B}$-t? Mi a helyzet, ha $\boldsymbol{A}$ 3×2-es?

Hasonló módszerrel, mint az 5. Kérdésnél, ellenpéldát lehet gyártani:

$$\begin{pmatrix} 0 & 0,5 & 0,5 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 3 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A $\boldsymbol{B}$ nem egyértelmű, hiszen

$$\begin{pmatrix} 0 & 0,5 & 0,5 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 5 \\ 1 & 2 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Ha $\boldsymbol{A}$ 3×2-es akkor $\boldsymbol{B}$ csak 2×3-as lehet, az eredménytől 3×3-as egységmátrixot várnánk el. $\boldsymbol{A}$ rangja legfeljebb 2, ezért $\boldsymbol{AB}$ képtere legfeljebb kétdimenziós, $\boldsymbol{AB}$ rangja legfeljebb 2, de $\boldsymbol{I}$ rangja 3. Ilyen ellenpélda nincs.

3. fejezet

Lineáris geometria

3.1. Síkok

Az 1. fejezet 12. Kérdése nyomán láttuk, hogy $ax_1 + bx_2 + cx_3 + dx_4 = 0$ egyenletnek eleget tevő $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$ vektorok (feltételezve, hogy az a, b, c és d számok nem mind nullák) az $\mathbf{R}^4$ valódi alterét alkotják, és az altér 3-dimenziós. Az állítás m dimenzióra is kiterjeszthető. Azon $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$ vektorok, melyekre $\sum_{i=1}^{m} a_i x_i = \langle a, x\rangle = 0$ ($\boldsymbol{a} = (a_1, a_2, \ldots, a_m) \neq \mathbf{0}$) teljesül, az $\mathbf{R}^m$ $(m-1)$-dimenziós alterét alkotják. Válasszunk egy nem nulla koordinátát, mondjuk $a_m \neq 0$, akkor látható ugyanis, hogy a $\boldsymbol{v}_k = (0, 0, \ldots, 0, a_m, 0, \ldots, -a_k)$ $(k = 1, 2, \ldots, m-1)$ vektorrendszer az altér bázisát képezi.

Másrészt egy tetszőleges $(m-1)$-dimenziós altér ortogonális kiegészítő alterét, amely egydimenziós altér, egy olyan $\boldsymbol{a}$ vektor generál, mely az altér minden elemére ortogonális. Az altér minden $\boldsymbol{x}$ elemére tehát fennáll az $\langle \boldsymbol{a}, \boldsymbol{x}\rangle = 0$ egyenlet. Úgy is mondhatjuk, hogy minden $(m-1)$-dimenziós altér egy lineáris funkcionál magtere.

Az m-dimenziós kocka kapcsán beszéltünk a kockát határoló $(m-1)$-dimenziós síkokról is. Ezek a síkok (az oldallapok síkjai) az $\langle \boldsymbol{e}_k, \boldsymbol{x}\rangle = 0$ vagy az $\langle \boldsymbol{e}_k, \boldsymbol{x}\rangle = 1$ egyenletnek tettek eleget, ahol $\boldsymbol{e}_k$ a koordináta egységvektort jelöli. Az $\langle \boldsymbol{e}_k, \boldsymbol{x}\rangle = 0$ megoldáshalmaza természetesen egyben altér is, de az $\langle \boldsymbol{e}_k, \boldsymbol{x}\rangle = 1$ megoldáshalmaza nem, mert a $\mathbf{0}$-t nem tartalmazza. Általában az $(m-1)$-dimenziós sík egyenletének a

$$\sum_{i=1}^{m} a_i x_i = b \quad (a = (a_1, a_2, \ldots, a_m) \neq \mathbf{0})$$

egyenletet tekintjük. Az egyenletnek eleget tevő vektorok halmaza $b = 0$ esetén altér, különben nem.

A $(m - 1)$-dimenziós sík és az $(m - 1)$-dimenziós altér között szoros összefüggés van. Ha az $\langle \boldsymbol{a}, \boldsymbol{x} \rangle = b$ síkot nézzük, és $\boldsymbol{x}_0$ egy pontja a síknak, azaz $\langle \boldsymbol{a}, \boldsymbol{x}_0 \rangle = b$, akkor $\langle \boldsymbol{a}, \boldsymbol{x} - \boldsymbol{x}_0 \rangle = 0$, azaz $\boldsymbol{x} - \boldsymbol{x}_0$ pontja az $\langle \boldsymbol{a}, \boldsymbol{x} \rangle = 0$ altérnek. Másrészt, ha $\boldsymbol{x}$ pontja az $\langle \boldsymbol{a}, \boldsymbol{x} \rangle = 0$ altérnek, és $\boldsymbol{x}_0 \in$ $\in \mathbf{R}^m$ úgy, hogy $\langle \boldsymbol{a}, \boldsymbol{x}_0 \rangle = b$, akkor $\langle \boldsymbol{a}, \boldsymbol{x} + \boldsymbol{x}_0 \rangle = b$, vagyis $\boldsymbol{x} + \boldsymbol{x}_0$ eleme az $\langle \boldsymbol{a}, \boldsymbol{x} \rangle = b$ síknak. Az $(m - 1)$-dimenziós sík tehát az $(m - 1)$-dimenziós altér eltolásával származtatható.

A k-dimenziós síkok $(1 \leq k < m)$ ugyanígy származtathatók k-dimenziós alterekből eltolással. Foglalkozzunk először a k-dimenziós altér megadásával. Alteret eddig leggyakrabban bázisával adtunk meg. Most használjunk, ezen túlmenően, még két megadási módot, valamely alkalmasan megadott lineáris leképezés képterével ill. magterével történő megadást!

Az L altér megadásának egyik legkézenfekvőbb módja generáló elemeivel (vagy bázisával) történő megadás. Ha adottak az L altér $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ generáló elemei ($\boldsymbol{v}_i \in \mathbf{R}^m$), akkor képezzük az $\boldsymbol{A} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k)$ mátrixot, és az $\boldsymbol{Ay}$: $\mathbf{R}^k \to \mathbf{R}^m$ leképezés képtere az L altér. Az L alteret ugyanis a bázisvektorok lineáris kombinációi, az $\sum_{i=1}^{k} y_i \boldsymbol{v}_i = \boldsymbol{Ay}$ vektorok alkotják, ahol $\boldsymbol{y} = (y_1, y_2, \ldots, y_k)$. Ez az ún. *képteres megadás*. Láttuk, hogy a generáló elemekkel történő és a képteres megadási mód lényegében ekvivalens.

Mivel a generáló elemek redukálhatók a bázisra, az $\boldsymbol{A}$ mátrixból is elhagyhatók a lineárisan összefüggő oszlopok, tehát feltehető, hogy az $\boldsymbol{A}$ mátrix k oszlopmérete megegyezik L dimenziójával.

Képezzük az $L^\perp$ altér, az ortogonális kiegészítő altér $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{m-k}$ generáló rendszerét (az 1.4.4. Tétel eljárást ad erre), és írjuk fel a $\boldsymbol{B}$ mátrixot az $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{m-k}$ sorvektorokból:

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_{m-k} \end{pmatrix}.$$

A $\boldsymbol{Bx}$: $\mathbf{R}^m \to \mathbf{R}^{m-k}$ leképezés magtere azon $\boldsymbol{x}$ vektorok halmaza, melyre $\boldsymbol{Bx} = \boldsymbol{0}$, ami azt jelenti, hogy $\boldsymbol{x}$ minden $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{m-k}$ vektorra ortogonális, vagyis $\boldsymbol{x} \in L$. A magtér és L egyezését a dimenzió-tétel (2.5.3. Tétel) biztosítja. Ez az ún. *magteres* (egyenlettel történő) *megadási mód.* A képteres megadáshoz hasonlóan, a magteres is redukálható, a lineárisan függő sorok a mátrixból elhagyhatók, és ekkor elérhető, hogy a mátrix $m - k$ sorméretében k az altér dimenziója.

1. Kérdés. $\mathbf{R}^4$-ben legyen az L altér bázisa $\boldsymbol{a} = (1,\ 2,\ 3,\ 0)$ és $\boldsymbol{b} = (1,\ 3,\ 2,\ 1)$. Adjuk meg azt a lineáris leképezést, amelynek L a magtere. Alkalmazzuk az 1.4.4. Tétel konstrukciós eljárását!

Természetesen sem a magteres, sem a képteres megadás nem egyértelmű, hiszen a mátrix sorait ill. oszlopait képező bázisvektorok tetszőlegesen választhatók. A fenti $\boldsymbol{A}$ és $\boldsymbol{B}$ mátrixokra azonban $\boldsymbol{BA} = \mathbf{0}$ mindig fennáll, hiszen minden $\boldsymbol{x}$-re $\boldsymbol{BAx} = \mathbf{0}$, mert $\boldsymbol{Ax}$ a $\boldsymbol{B}$ magterének eleme. Ebből is látható, hogy a $\boldsymbol{B}$ mátrix sorvektorai és az $\boldsymbol{A}$ mátrix oszlopvektorai ortogonális altereket generálnak.

3.1.1. Definíció. Az $\boldsymbol{B}$ és az $\boldsymbol{A}$ mátrixok (ebben a sorrendben) *anullálják* egymást, ha $\boldsymbol{BA} = \mathbf{0}$, és a rangok összege megegyezik $\boldsymbol{A}$ sorméretével (ill. $\boldsymbol{B}$ oszlopméretével).

Ha $\boldsymbol{B}$ $k \times n$-es r_2 rangú, $\boldsymbol{A}$ $n \times m$-es r_1 rangú mátrix és annullálják egymást, akkor $\boldsymbol{A}$ képtere $\boldsymbol{B}$ magterébe esik. $\boldsymbol{B}$ magtere azonban a dimenziótétel (2.5.3.Tétel) miatt $n - r_2 = r_1$ dimenziós, $\boldsymbol{A}$ képtere szintén r_1 dimenziós, így $\boldsymbol{A}$ képtere megegyezik $\boldsymbol{B}$ magterével. $\boldsymbol{B}$ és $\boldsymbol{A}$ akkor és csak akkor határozzák meg ugyanazt az alteret magtérként ill. képtérként, ha anullálják egymást.

Ha $\boldsymbol{B}$ és $\boldsymbol{A}$ anullálják egymást, akkor $\boldsymbol{A}^*$ és $\boldsymbol{B}^*$ is anullálják egymást (ebben a sorrendben), és az általuk meghatározott altér az előbbi ortogonális kiegészítője (ld. 2.5.6. Következmény).

3.1.2. Definíció. Legyen $L \subset \mathbf{R}^m$ tetszőleges k-dimenziós altér ($0 \leq k \leq m$) és $\boldsymbol{a} \in \mathbf{R}^m$ tetszőleges vektor. $\mathbf{R}^m$ k-dimenziós *síkjának* általános alakja: $S = \{\boldsymbol{x} + \boldsymbol{a}\colon\ \boldsymbol{x} \in L\}$. $\boldsymbol{a}$ az S sík bármely pontja lehet.

A definíció megfelel annak a megállapításnak, hogy a sík az altér eltolásával kapható meg.

Nézzük a definíció következményeit a sík különböző megadási módjaira vonatkozóan.

A k-dimenziós síkot általában $k + 1$ pontja határozza meg, jelöljük ezeket $\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \ldots,\ \boldsymbol{a}_{k+1}$-gyel. Vegyük $\boldsymbol{a}$-t $\boldsymbol{a} = \boldsymbol{a}_{k+1}$-nek, és képezzük az $\boldsymbol{a}_1 - \boldsymbol{a}_{k+1},\ \boldsymbol{a}_2 - \boldsymbol{a}_{k+1},\ \ldots,\ \boldsymbol{a}_k - \boldsymbol{a}_{k+1}$ vektorokat. Ha ezek a vektorok lineárisan függetlenek, akkor generálják L-et, és L-et $\boldsymbol{a}$-val eltolva megkapjuk a síkot.

Az $\boldsymbol{a}_1 - \boldsymbol{a}_{k+1},\ \boldsymbol{a}_2 - \boldsymbol{a}_{k+1},\ \ldots,\ \boldsymbol{a}_k - \boldsymbol{a}_{k+1}$ vektorok lineáris függetlensége azt jelenti, hogy a

$$c_1(\boldsymbol{a}_1 - \boldsymbol{a}_{k+1}) + c_2(\boldsymbol{a}_2 - \boldsymbol{a}_{k+1}) + ... + c_k(\boldsymbol{a}_k - \boldsymbol{a}_{k+1}) = 0$$

összefüggés csak úgy következhet be, ha $c_1 = c_2 = \ldots = c_k = 0$. Rendezzük át, és vezessük be a $c_{k+1} = -(c_1 + c_2 + \ldots + c_k)$ jelölést, akkor

$$c_1(\boldsymbol{a}_1 - \boldsymbol{a}_{k+1}) + c_2(\boldsymbol{a}_2 - \boldsymbol{a}_{k+1}) + ... + c_k(\boldsymbol{a}_k - \boldsymbol{a}_{k+1}) =$$

$$= \sum_{i=1}^{k} c_i \boldsymbol{a}_i - a_{k+1} \sum_{i=1}^{k} \boldsymbol{c}_i = \sum_{i=1}^{k+1} c_i \boldsymbol{a}_i = 0,$$

ahol $\sum_{i=1}^{k+1} c_i = 0$. Mondjuk ki a definíciót:

3.1.3. Definíció. Az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{k+1}$ vektorok *affin függetlenek*, ha

$$\sum_{i=1}^{k+1} c_i \boldsymbol{a}_i = 0 \quad \text{és} \quad \sum_{i=1}^{k+1} c_i = 0$$

egyidejűleg csak úgy teljesülhet, ha $c_1 = c_2 = \ldots = c_k = 0$.

Megjegyzés. Az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{k+1}$ vektorok affin függetlenségének ellenőrzése célszerű módon az $\boldsymbol{a}_1 - \boldsymbol{a}_{k+1}$, $\boldsymbol{a}_2 - \boldsymbol{a}_{k+1}, \ldots, \boldsymbol{a}_k - \boldsymbol{a}_{k+1}$ vektorok lineáris függetlenségének ellenőrzésével történhet, ami rangszámítással hajtható végre.

3.1.4. Tétel. $k+1$ darab affin független vektor egyértelműen meghatároz egy olyan k-dimenziós síkot, melynek ezek a vektorok a pontjai. Az ilyen vektorrendszert a *sík affin bázisának* nevezzük.

A sík vektorai (pontjai) az affin bázis segítségével könnyen előállíthatók. Az L altér pontjai $\sum_{i=1}^{k} c_k(\boldsymbol{a}_i - \boldsymbol{a}_{k+1})$ alakúak, ezért a sík pontjai a

$$\sum_{i=1}^{k} c_i(\boldsymbol{a}_i - \boldsymbol{a}_{k+1}) + \boldsymbol{a}_{k+1} = \sum_{i=1}^{k} c_i \boldsymbol{a}_i + \left(1 - \sum_{i=1}^{k} c_i\right) \boldsymbol{a}_{k+1}$$

formulával jellemezhetők. Ha bevezetjük a $c_{k+1} = 1 - (c_1 + c_2 + \ldots + c_k)$ jelölést, akkor a sík tetszőleges $\boldsymbol{x}$ pontja előállítható az $\boldsymbol{x} = \sum_{i=1}^{k+1} c_i \boldsymbol{a}_i$ alakban, ahol $\sum_{i=1}^{k+1} c_i = 1$. A c_i számokat az $\boldsymbol{x}$ pont $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{k+1}$ rendszerre vonatkozó baricentrikus koordinátáinak nevezzük.

A baricentrikus koordináták fizikai interpretációja a következő: az $\boldsymbol{a}_1$, $\boldsymbol{a}_2, \ldots, \boldsymbol{a}_{k+1}$ pontokban rendre $c_1, c_2, \ldots, c_{k+1}$ tömegpontokat helyezünk el, akkor a rendszer súlypontja $\boldsymbol{x}$. Negatív súlyokat – ellenkező irányú erőhatást – is megengedve a súlypont a sík bármely pontjába áthelyezhető. Affin független rendszer esetén $\boldsymbol{x}$ a súlyokat egyértelműen jellemzi, így azok koordinátáknak is felfoghatók.

Az affin független $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{k+1}$ vektorok, mint láttuk, meghatározzák az L alteret, amelynek eltolásával kapjuk a vektorok által meghatározott síkot. $L^\perp$ egy bázisát véve a bázisvektorokból, *mint sorvektorokból*, elkészíthető a $\boldsymbol{B}$ mátrix, melynek magtere L: $\boldsymbol{x} \in L$ akkor és csak akkor, ha $\boldsymbol{Bx} = \boldsymbol{0}$. Ez az L altér magteres megadási módja. Legyen $\boldsymbol{a} \in S$, akkor az S sík pontjai az $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{a}$ pontok, melyekre $\boldsymbol{By} = \boldsymbol{Bx} + \boldsymbol{Ba} = \boldsymbol{Ba} = \boldsymbol{b}$. A síkot tehát a $\boldsymbol{By} = \boldsymbol{b}$ lineáris egyenletrendszer megoldásai adják. k-dimenziós sík esetén $\boldsymbol{B}$ itt $(m-k) \times m$-es teljesrangú mátrix, $\boldsymbol{b} \in \mathbf{R}^{m-k}$ tetszőleges vektor. Megfordítva: Tetszőleges $\boldsymbol{By} = \boldsymbol{b}$ lineáris egyenletrendszer megoldásai mindig síkot alkotnak, vagy pontot, vagy üres halmazt, hiszen $\boldsymbol{Bx} = \boldsymbol{0}$ megoldásai alteret képeznek, és ha van olyan $\boldsymbol{a}$, hogy $\boldsymbol{Ba} = \boldsymbol{b}$, akkor ennek az altérnek $\boldsymbol{a}$-val történő eltoltjáról van szó.

A sík megadható képtérként is. A síkot meghatározó $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{k+1}$ vektorokból készítsük el az $\boldsymbol{a}_1 - \boldsymbol{a}_{k+1}, \boldsymbol{a}_2 - \boldsymbol{a}_{k+1}, \ldots, \boldsymbol{a}_k - \boldsymbol{a}_{k+1}$ vektorokat, ezekből az

$$\boldsymbol{A} = (\boldsymbol{a}_1 - \boldsymbol{a}_{k+1},\ \boldsymbol{a}_2 - \boldsymbol{a}_{k+1}, \ldots, \boldsymbol{a}_k - \boldsymbol{a}_{k+1})$$

mátrixot. Az $\boldsymbol{Ax}$: $\mathbf{R}^k \to \mathbf{R}^m$ leképezés képtere az L altér (képteres megadási mód), és ha a sík egy pontja $\boldsymbol{a}$ (pl. $\boldsymbol{a}_i$), akkor az $\boldsymbol{Ax} + \boldsymbol{a}$ függvény képtere (értékkészlete) az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{k+1}$ vektorok által meghatározott sík.

3.2. Lineáris egyenletrendszerek

A lineáris egyenletrendszer általános alakja,

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m = b_2$$

$$\ldots$$

$$a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m = b_n,$$

a mátrixaritmetikát felhasználva az

$$\boldsymbol{Ax} = \boldsymbol{b}$$

tömör formában írható fel. Itt $\boldsymbol{A}$ $n \times m$-es mátrix, $\boldsymbol{b} \in \mathbf{R}^n$ konstansokból álló vektor, $\boldsymbol{x} \in \mathbf{R}^m$ az ismeretleneket összefogó vektor. A $\boldsymbol{b} = \boldsymbol{0}$ speciális esetet homogén lineáris egyenletrendszernek nevezzük. Ebben a részben a megoldhatóságot, a megoldások struktúráját vizsgáljuk, a megoldások módszereivel itt nem foglalkozunk. Az 1. Kérdésben elvégzettek mintájára, az $\boldsymbol{A}$ mátrixot írjuk fel $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$ alakban, akkor a megoldás létezése azt jelenti, hogy valamilyen x_1, x_2, $\ldots$, x_m számokra $\sum_{i=1}^{m} \boldsymbol{a}_i x_i = \boldsymbol{b}$, vagyis $\boldsymbol{b}$ nem növeli az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_m$ vektorrendszer rangját. Lényegében ezen múlik a megoldhatóság feltétele.

3.2.1. Tétel. $\boldsymbol{Ax} = \boldsymbol{b}$ akkor és csak akkor oldható meg, ha $\boldsymbol{A}$ rangja megegyezik az $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m, \boldsymbol{b})$ ún. *kibővített mátrix* rangjával.

Bizonyítás. $\Rightarrow$ Ha $\boldsymbol{Ax} = \boldsymbol{b}$ megoldható, akkor a rangszámítás oszlopműveletei alapján

$$\begin{gathered}\operatorname{rang}(\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m) = \operatorname{rang}(\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m) = \\ = \operatorname{rang}(\boldsymbol{a}_1 x_1, \boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m) = \operatorname{rang}(\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2, \boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m) = ... = \\ = \operatorname{rang}(\sum_{i=1}^{m} \boldsymbol{a}_i x_i, \boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m) = \operatorname{rang}(\boldsymbol{b}, \boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m).\end{gathered}$$

$\Leftarrow$ Ha $r=$ $\operatorname{rang}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$, akkor feltehető, hogy $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_r$ lineárisan független rendszer. $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_r$, $\boldsymbol{b}$ már nem lehet lineárisan független, mert $\operatorname{rang}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m, \boldsymbol{b}) = r$. Létezik tehát nem csupa nulla együtthatós lineáris összefüggés a vektorok között:

$$\sum_{i=1}^{r} c_i \boldsymbol{a}_i + d\boldsymbol{b} = 0,$$

ahol $d \neq 0$, hiszen az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_r$ vektorok között nincs ilyen lineáris összefüggés. $\boldsymbol{b}$ tehát kifejezhető,

$$\boldsymbol{b} = \sum_{i=1}^{r} -\frac{c_i}{d} \boldsymbol{a}_i,$$

és ezzel a lineáris egyenletrendszer megoldását előállítottuk: $x_i = -\frac{c_i}{d}$, ha $i = 1, 2, \ldots, r$ és $x_i = 0$, ha $i = r + 1, \ldots, m$. $\square$

2. Kérdés. Megoldható-e az

$$\begin{gathered}x - y + 2z + t = 6 \\ 3x - 2y + z - t = 0 \\ x - 3z - 3t = 1\end{gathered}$$

egyenletrendszer?

Előző pontban láttuk, hogy tetszőleges $\boldsymbol{Ax} = \boldsymbol{b}$ lineáris egyenletrendszer megoldásai – ha léteznek – mindig síkot alkotnak, és ez a sík az $\boldsymbol{Ax}$ magterének eltoltja egy olyan $\boldsymbol{a}$-val, melyre $\boldsymbol{Aa} = \boldsymbol{b}$. Még egyszer elmondva ugyanezt az egyenletrendszerek nyelvén: az $\boldsymbol{Ax} = \boldsymbol{b}$ bármely megoldása (általános megoldása) megkapható úgy, hogy az $\boldsymbol{Ax} = \boldsymbol{b}$ egy egyedi, ún. partikuláris megoldásához az $\boldsymbol{Ax} = \mathbf{0}$ homogén egyenletrendszer valamely megoldását (ill. az általános megoldását) hozzáadjuk.

A megoldás létezése tehát a partikuláris megoldás létezésén, a megoldások „száma” pedig a magtér dimenzióján múlik. A magtér dimenziója a dimenzió-tétel szerint (2.5.3. Tétel) $s = m - r$, vagyis a létezést feltételezve a megoldás egyértelmű, ha $m = r$, és a megoldáshalmaz $s = m - r$ dimenziós síkot alkot, ha $m > r$.

3. Kérdés. Az

$$\begin{aligned} 5x + y - z &= 5 \\ x - y + z &= 1 \\ x + y - z &= 1 \end{aligned}$$

egyenletrendszernek $x = 1$, $y = 3$, $z = 3$ megoldása. Van-e más megoldás?

Speciális esetet képez az $\boldsymbol{Ax} = \mathbf{0}$ homogén egyenletrendszer. Ez mindig megoldható, hiszen $\boldsymbol{x} = \mathbf{0}$ mindig triviális megoldást jelent. Az igazi kérdést itt a nem triviális megoldás létezése jelenti. Ha $m = r$, vagyis az $\boldsymbol{A}$ oszlopvektorai lineárisan függetlenek, akkor (és csak akkor) a megoldás egyértelmű, tehát csak a triviális megoldás létezik. Az $\boldsymbol{Ax} = \mathbf{0}$ homogén lineáris egyenletrendszernek akkor és csak akkor van nem triviális megoldása, ha $m > r$. A megoldások halmaza ekkor $m - r$ dimenziós altér.

4. Kérdés. Az

$$\begin{aligned} x + y - z + 2w &= 0 \\ x + z - w &= 0 \end{aligned}$$

megoldásai hány dimenziós alteret alkotnak? Adjuk meg az általános megoldást és a megoldáshalmaz egy bázisát!

3.3. Síkok egymáshoz viszonyított helyzete

Már említettem a következő javaslatot arra, hogy egy háromszög szögét hogyan határozhatjuk meg magasabb dimenzió esetén. Fektessünk át egy kétdimenziós síkot a három ponton – láttuk, hogy ez megtehető –, majd az (x, y)-koordinátasíkkal képezett metszésvonala körül forgassuk be az alapsíkba. Ekkor sétáljunk oda, és mérjük le.

A forgatásról majd később beszélünk, nézzük meg előbb a metszésvonal kérdését.

Alterek közös része szintén altér, hiszen a lineáris műveletek minden egyes altérben külön-külön elvégezhetők, így az eredmény a metszethez is hozzátartozik. Két altér metszetét a magteres megadással jellemezhetjük a legkönnyebben, legyen az egyik egyenlete $\boldsymbol{A}_1\boldsymbol{x} = \boldsymbol{0}$, a másik $\boldsymbol{A}_2\boldsymbol{x} = \boldsymbol{0}$. Az alterek metszetére mindkét egyenlet teljesül, tehát képezzük az $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{pmatrix}$ mátrixot, és az $\boldsymbol{Ax} = \boldsymbol{0}$ egyenlet írja le a két altér közös részét képező alteret. Ha $\text{rang}(\boldsymbol{A}) = r$, akkor metszet dimenziója $m - r$.

Jóval nehezebb a feladat, ha a két alteret képtérként adjuk meg. Legyen $\boldsymbol{A}_1\boldsymbol{x}$: $\mathbf{R}^p \to \mathbf{R}^m$ és $\boldsymbol{A}_2\boldsymbol{y}$: $\mathbf{R}^q \to \mathbf{R}^m$, akkor keressük azon $\boldsymbol{b} \in \mathbf{R}^m$ vektorok halmazát, melyre $\boldsymbol{A}_1\boldsymbol{x} = \boldsymbol{b}$ is és $\boldsymbol{A}_2\boldsymbol{y}= \boldsymbol{b}$ is megoldható.

Hasonlóan képezhetjük két sík metszetét. Az $\boldsymbol{A}_1\boldsymbol{x} = \boldsymbol{b}_1$ és az $\boldsymbol{A}_2\boldsymbol{x} = \boldsymbol{b}_2$ síkok metszete azon $\boldsymbol{x}$ vektorok halmaza, amelyre mindkét egyenlet teljesül, ami összefoglalható $\boldsymbol{Ax} = \boldsymbol{b}$ alakban, ahol $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{pmatrix}$ és $\boldsymbol{b} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix}$. Ebből látható, hogy ha az egyenlet egyáltalán megoldható, akkor a megoldások halmaza sík, melynek dimenziója a korábbival megegyezően $m - r$.

Míg az alterek esetén mindig volt megoldás (a $\boldsymbol{0}$ az alterek közös pontja), itt elképzelhető, hogy nincs megoldás, vagyis a két sík nem metszi egymást: kitérők vagy párhuzamosok. Tegyük fel ebben a részben, hogy az altereket és a síkokat nem egyszerűsíthető alakban adjuk meg, vagyis az $\boldsymbol{A}_1$ sorai is és az $\boldsymbol{A}_2$ sorai is lineárisan függetlenek. Ha a magterek nem üresek, vagyis ha mindkét egyenletrendszer megoldható, akkor a lineárisan függő sorok elhagyhatók. Legyen $\text{rang}(\boldsymbol{A}_1) = r_1$, $\text{rang}(\boldsymbol{A}_2) = r_2$, akkor a mátrixok méretei $r_1 \times m$ ill. $r_2 \times m$. rang $(\boldsymbol{A}) = r$, és $\boldsymbol{A}$ mérete $(r_1 + r_2) \times m$.

3.3.1. Tétel. Adott $\boldsymbol{A}_1$ és $\boldsymbol{A}_2$ esetén az $\boldsymbol{A}_1\boldsymbol{x} = \boldsymbol{b}_1$ és az $\boldsymbol{A}_2\boldsymbol{x} = \boldsymbol{b}_2$ síkoknak akkor és csak akkor van minden $\boldsymbol{b}_1$ és $\boldsymbol{b}_2$ esetén közös pontjuk, ha $r_1 + r_2 = r$.

Bizonyítás. Használjuk az $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{pmatrix}$ és a $\boldsymbol{b} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix}$ jelöléseket.

$\Rightarrow$ Ha $\boldsymbol{Ax} = \boldsymbol{b}$ minden $\boldsymbol{b}$-re megoldható, akkor $R_A = \mathbf{R}^{r_1+r_2}$, ezért $\text{rang}(\boldsymbol{A}) = r = r_1 + r_2$.

$\Leftarrow$ Ha $r = r_1 + r_2$, akkor a kibővített mátrix rangja sem lehet több az $\boldsymbol{A}$ mérete miatt. $\square$

Nézzük azt a speciális esetet, amikor $r_1 + r_2 = r = m$. Ekkor a magtér 0 dimenziós, tehát egyetlen megoldás van. Például $\mathbf{R}^4$-ben két kétdimenziós (közönséges) sík, ha az $r_1+r_2=r$ feltétel teljesül, akkor egy pontban metszi egymást. Nem tudjuk elképzelni? Dehogynem! Vegyük az egyik síknak az $(x,\ y)$-koordinátasíkot, a sík koordinátáit az jellemzi, hogy a harmadik és a negyedik koordináta nulla, és vegyük a $(z,\ w)$-koordinátasíkot, a sík koordinátáit pedig az jellemzi, hogy az első és a második koordináta nulla. A közös pontban mind a négy koordináta 0, tehát a metszéspont csak a $\mathbf{0}$. Ráadásul ez a tipikus eset!

Nem metsző síkok párhuzamosak, vagy kitérők. Definiáljuk a párhuzamosság fogalmát, ezzel a kitérő síkok fogalma már adott: nem párhuzamos közös pont nélküli síkok.

3.3.2. Definíció. $\mathbf{R}^m$ két síkja *párhuzamos*, ha nincs közös pontjuk, és a hozzájuk tartozó alterek egyike tartalmazza a másikat.

Pont és sík távolságát a szokásos módon, két alakzat távolságának megfelelően definiáljuk, vagyis az összekötő szakaszok hosszának az infimumával.

3.3.3. Tétel. Ha az S_1 és az S_2 síkok párhuzamosak, akkor a kisebb (vagy egyenlő) dimenziós sík minden pontja egyenlő távol van a másik síktól. Ez a távolság a két sík távolsága.

Bizonyítás. Legyen a két sík egyenlete

$$\begin{aligned} S_1: &\qquad \boldsymbol{A}_1\boldsymbol{x}=\boldsymbol{b}_1 \\ S_2: &\qquad \boldsymbol{A}_2\boldsymbol{x}=\boldsymbol{b}_2, \end{aligned}$$

továbbá legyen S_2 a kisebb (vagy egyenlő) dimenziójú sík. A párhuzamosság úgy fogalmazható meg, hogy $\boldsymbol{A}_2\boldsymbol{x}=\mathbf{0}$-ból következik, hogy $\boldsymbol{A}_1\boldsymbol{x}=\mathbf{0}$. Az S_2 tetszőleges $\boldsymbol{x}_2$ pontját kössük össze az S_1 valamely $\boldsymbol{x}_1$ pontjával, a távolságuk $d=||\boldsymbol{x}_1-\boldsymbol{x}_2||$. Ha S_2-nek egy $\boldsymbol{x}_2$' pontját választjuk, akkor ehhez található olyan S_1-beli $\boldsymbol{x}_1$' pont, hogy ezek távolsága is d. Legyen ugyanis $\Delta\boldsymbol{x}=\boldsymbol{x}_2$'$-\ \boldsymbol{x}_2$ és $\boldsymbol{x}_1$'$=\boldsymbol{x}_1+\Delta\boldsymbol{x}$, akkor $\boldsymbol{A}_2\Delta\boldsymbol{x}=\boldsymbol{b}_2-\boldsymbol{b}_2=\mathbf{0}$ és $\boldsymbol{A}_1\boldsymbol{x}_1$'$=\boldsymbol{A}_1\boldsymbol{x}_1+\boldsymbol{A}_1\Delta\boldsymbol{x}=\boldsymbol{b}_1+\mathbf{0}=\boldsymbol{b}_1$, az $\boldsymbol{x}_1$' pont tehát S_1-hez tartozik. Ugyanakkor $||\boldsymbol{x}_1-\boldsymbol{x}_2||=||\boldsymbol{x}_1$' $-\ \boldsymbol{x}_2$'$||=d$.

Ebből következik, hogy az összekötő szakaszok hosszának infimuma mindkét pontra ugyanaz. □

3.3.4. Tétel. $\mathbf{R}^m$ két síkja, ha nincs közös pontjuk és legalább egyikük $m-1$ dimenziós, akkor csak párhuzamos lehet.

Bizonyítás. Jellemezzük a két síkot az $\boldsymbol{A}_i\boldsymbol{x}=\boldsymbol{b}_i$ egyenlettel. Ha az első sík az $(m-1)$-dimenziós, akkor $\boldsymbol{A}_1$ $1\ \times m$-es, jelölje egyetlen sorvektorát $\boldsymbol{a}$, $\boldsymbol{A}_2$ (feltételezhetően lineárisan független) sorvektorait pedig $\boldsymbol{b}_1,\boldsymbol{b}_2,...,\boldsymbol{b}_{r_2}$. Ha $\boldsymbol{a}$, $\boldsymbol{b}_1,\boldsymbol{b}_2,...,\boldsymbol{b}_{r_2}$ lineárisan függetlenek, akkor $r=r_1+r_2$, igy a 3.3.1. Tétel miatt a síkoknak van közös pontja, ami ellentmond a feltételnek. Ha nem lineárisan függetlenek, akkor $\boldsymbol{a}$ kifejezhető a $\boldsymbol{b}_1,\boldsymbol{b}_2,...,\boldsymbol{b}_{r_2}$ vektorokkal, ekkor $\boldsymbol{A}_2\boldsymbol{x} = \mathbf{0}$-ból következik, hogy $\boldsymbol{A}_1\boldsymbol{x} = \langle \boldsymbol{a}, \boldsymbol{x}\rangle = \mathbf{0}$, vagyis az első sík altere tartalmazza a másodikét. □

Lényegében minden más dimenzópár esetén léteznek kitérő síkok. Nem jó az a szemlélet, hogy $\mathbf{R}^4$-ben két kétdimenziós sík azért lehet kitérő, mert a dimenziójukat tekintve „elférnek”, hiszen $2 + 2 \leq 4$. $\mathbf{R}^5$-ben ugyanis két háromdimenziós sík is lehet kitérő!

Vizsgáljuk meg a kitérő térelemek távolságát. Itt a kitérő térelemek közé felvehetjük a párhuzamosakat is, az elmondottak nem különböznek.

3.3.5. Tétel. Ha S_1 és S_2 két közös pont nélküli sík $\mathbf{R}^m$-ben, akkor létezik olyan szakasz, mely a két síkot összeköti és mindkettőre ortogonális. A szakaszt *normál transzverzálisnak* nevezzük. A normál transzverzális hossza a két sík távolsága.

Bizonyítás. Az egyik síkról, mondjuk S_1-ről feltehető, hogy altér, ugyanis a koordináta-rendszer origóját áthelyezhetjük S_1 egy tetszőlegesen kiválasztott pontjába. Ekkor a két sík egyenlete legyen S_1: $\boldsymbol{A}_1\boldsymbol{x} = \mathbf{0}$ és S_2: $\boldsymbol{A}_2\boldsymbol{x} = \boldsymbol{b}$. Jelöljük az $\{\boldsymbol{x}\colon\ \boldsymbol{A}_2\boldsymbol{x} = \mathbf{0}\}$ alteret L_2-vel, és képezzük az S_1 és L_2 által generált alteret, legyen ez L. A konstrukcióból következik, hogy S_2 párhuzamos L-lel, ezért a 3.3.3. Tétel miatt minden pontja ugyanolyan távol van L-től. Vetítsük rá L-re ortogonálisan S_2 minden egyes pontját, ez S_2-nek egy eltoltja, ezt a síkot jelöljük S_0-lal. Ha felveszünk egy tetszőleges $\boldsymbol{y} \in S_2$ pontot, és képezzük a $\boldsymbol{v} = \boldsymbol{y}_0 - \boldsymbol{y}$ vektort, ahol $\boldsymbol{y}_0$ az $\boldsymbol{y}$ ortogonális vetülete L-re, akkor a $\boldsymbol{v}$ vektor az az eltolási vektor, ami S_2-t S_0-ba viszi. S_0 egyenlete tehát $\boldsymbol{A}_2(\boldsymbol{x} - \boldsymbol{v}) = \boldsymbol{A}_2\boldsymbol{x} - \boldsymbol{A}_2\boldsymbol{v} = \boldsymbol{A}_2\boldsymbol{x} - \boldsymbol{A}_2\boldsymbol{y}_0 + \boldsymbol{A}_2\boldsymbol{y} = \\ = \boldsymbol{A}_2(\boldsymbol{x} - \boldsymbol{y}_0) + \boldsymbol{b} = \boldsymbol{b}$, vagyis $\boldsymbol{A}_2(\boldsymbol{x} - \boldsymbol{y}_0) = \mathbf{0}$.

Mivel $\boldsymbol{y}_0 \in L$, felírható, hogy $\boldsymbol{y}_0 = \boldsymbol{y}_1 + \boldsymbol{y}_2$, ahol $\boldsymbol{y}_1 \in S_1$ és $\boldsymbol{y}_2 \in L_2$. Az így előállított $\boldsymbol{y}_1$ az S_1 és az S_0 közös pontja lesz, ugyanis kielégíti az S_0 sík egyenletét:

$$\boldsymbol{A}_2(\boldsymbol{y}_1 - \boldsymbol{y}_0) = -\boldsymbol{A}_2\boldsymbol{y}_2 = \mathbf{0}.$$

A normál transzverzális egyik végpontja $\boldsymbol{y}_1$, a másik $\boldsymbol{y}_1 - \boldsymbol{v}$, az ortogonalitás a konstrukcióból adódik, az S_1 és az S_2 síkok távolsága $||\boldsymbol{v}||$. □

Ellentétben a háromdimenziós térrel a normál transzverzális háromnál magasabb dimenzióban nem biztos, hogy egyértelmű, ugyanis általában az $\boldsymbol{y}_0 = \boldsymbol{y}_1 + \boldsymbol{y}_2$ előállítás egyértelműsége nem teljesül.

5. Kérdés. Milyen helyzetű a következő két sík?

$$\begin{aligned} S_1: \quad & 2x + y + w = 4 \\ & -x + 2y + z = 2 \\ S_2: \quad & x + y + 2z + w = 5 \\ & x + 3y + z + w = 4 \end{aligned}$$

Két sík ortogonalitása is analóg módon definiálható.

3.3.5. Definíció. Két sík ortogonális, ha egyik alterének ortogonális kiegészítő altere tartalmazza a másikhoz tartozó alteret.

3.4. Vetítés

A vetítés, vagy pontosabban adott altérre történő vetítés, az $\mathbf{R}^n$ olyan lineáris transzformációja melynek képtere az adott altér és az altér pontjait helybenhagyja.

3.4.1. Definíció. Az $\boldsymbol{A}$: $\mathbf{R}^n \to \mathbf{R}^n$ lineáris transzformációt *vetítésnek* nevezzük, ha $\boldsymbol{A}^2 = \boldsymbol{A}$.

Az $\boldsymbol{A}$ képtere, R_A tehát egyben az az altér, amelyre a vetítés történik. Vegyünk egy $\boldsymbol{y} \in R_A$ pontot, akkor azon pontok halmaza, amelyeknek vetülete erre a pontra képződik, vagyis az $\{\boldsymbol{x}\colon\ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}\}$ halmaz a *vetítő sík*. Az $\boldsymbol{y} = \mathbf{0}$-hoz tartozó vetítő sík, az ún. *vetítő altér*, az M_A magtér.

3.4.2. Tétel. R_A és M_A kiegészítő alterek.

Bizonyítás. Vegyünk egy $\boldsymbol{y}$ közös pontot, akkor van olyan $\boldsymbol{x}$, hogy $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, és ugyanakkor $\boldsymbol{A}\boldsymbol{y} = \mathbf{0}$. De $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}$, akkor viszont $\boldsymbol{y} = \mathbf{0}$. Az altereknek tehát egyetlen közös pontjuk van a $\mathbf{0}$.

Másrészt tetszőleges $\boldsymbol{x} \in \mathbf{R}^n$-re $\boldsymbol{x} = (\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x}) + \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, ahol $\boldsymbol{x}_2 = \boldsymbol{A}\boldsymbol{x} \in R_A$ és $\boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x} = \mathbf{0}$, tehát $\boldsymbol{x}_1 \in M_A$. □

Adott vetítősík R_A-t természetesen egyetlen pontban metszi, ez a vetület.

A különböző $\boldsymbol{y}$-hoz tartozó vetítő síkok egymással párhuzamosak. Közös pontjuk nem lehet, mert egy $\boldsymbol{x}$-hez nem tartozhat két különböző vetület, másrészt a síkokhoz tartozó altér, M_A vetítő altér közös.

3.4.3. Tétel. Ha az $\boldsymbol{A}$ transzformáció vetítés, akkor az $\boldsymbol{I} - \boldsymbol{A}$ transzformáció is vetítés, mely az $\boldsymbol{A}$ vetítő alterére vetít, és vetítő altere R_A.

Bizonyítás. Mivel $(\boldsymbol{I} - \boldsymbol{A})^2 = \boldsymbol{I} - 2\boldsymbol{A} + \boldsymbol{A}^2 = \boldsymbol{I} - \boldsymbol{A}$, az $\boldsymbol{I} - \boldsymbol{A}$ transzformáció vetítés.

Ha $\boldsymbol{x} \in R_A$, akkor $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x} = \mathbf{0}$, vagyis $\boldsymbol{x}$ hozzátartozik $\boldsymbol{I} - \boldsymbol{A}$ magteréhez, és a képlet fordítva is olvasható: ha $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \mathbf{0}$, akkor $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}$, vagyis $\boldsymbol{x} \in R_A$.

Ha $\boldsymbol{x} \in M_A$, akkor $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}$, vagyis $\boldsymbol{x}$ hozzátartozik $\boldsymbol{I} - \boldsymbol{A}$ képteréhez, és a képletsor fordítva is olvasható. □

Ortogonális a vetítés, ha a vetítősík ortogonális a vetületek alterére.

3.4.4. Definíció. Az $\boldsymbol{A}$ vetítést *ortogonális vetítésnek* nevezzük, ha R_A és M_A ortogonális kiegészítő alterek.

3.4.5. Tétel. Az $\boldsymbol{A}$ vetítés akkor és csak akkor ortogonális, ha $\boldsymbol{A} = \boldsymbol{A}^*$.

Bizonyítás. $\Rightarrow$ Legyen $\boldsymbol{x}$ és $\boldsymbol{y} \in \mathbf{R}^n$ tetszőleges, mindkettő felbontható $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, $\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{y}_2$ alakba, ahol $\boldsymbol{x}_1$ és $\boldsymbol{y}_1 \in M_A$, és $\boldsymbol{x}_2$, $\boldsymbol{y}_2 \in R_A$. Ekkor

$$\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{y} \rangle = \langle \boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y}_2 \rangle = \langle \boldsymbol{x}_2, \boldsymbol{y}_2 \rangle,$$
$$\langle \boldsymbol{x}, \boldsymbol{A}^*\boldsymbol{y} \rangle = \langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}_2, \boldsymbol{y}_1 + \boldsymbol{y}_2 \rangle = \langle \boldsymbol{x}_2, \boldsymbol{y}_2 \rangle.$$

A két egyenlőség kivonásával minden $\boldsymbol{x}$-re érvényesen kapjuk, hogy $\langle \boldsymbol{x}, (\boldsymbol{A} - \boldsymbol{A}^*)\boldsymbol{y} \rangle = \boldsymbol{0}$, válasszuk $\boldsymbol{x}$-nek $\boldsymbol{x} = (\boldsymbol{A} - \boldsymbol{A}^*)\boldsymbol{y}$-t, akkor $||(\boldsymbol{A} - \boldsymbol{A}^*)\boldsymbol{y}||^2 = 0$, tehát $(\boldsymbol{A} - \boldsymbol{A}^*)\boldsymbol{y} = \boldsymbol{0}$, azaz minden $\boldsymbol{y}$-ra $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{A}^*\boldsymbol{y}$, amit bizonyítani akartunk.

$\Leftarrow$ A 2.5.6. Következmény alapján R_{A^*} és M_A ortogonális kiegészítő alterek, de jelen esetben $R_{A^*} = R_A$. $\square$

6. Kérdés. Jelöljük az $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ minta átlagát $\bar{x}$-sal, és képezzük az $\boldsymbol{y} = (x_1 - \bar{x}, x_2 - \bar{x}, \ldots, x_n - \bar{x})$ vektort (ld. 2.fejezet 2. Kérdés). Mátrixalgebra felhasználása nélkül mutassuk meg, hogy a lineáris transzformáció, mely az $\boldsymbol{x}$-et az $\boldsymbol{y}$-ba képezi le, ortogonális vetítés! Mennyi annak az altérnek a dimenziója, amelyre a vetítés történik?

7. Kérdés. Igaz-e, hogy az

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 1 \\ -4 & -3 & -4 \\ 2 & 2 & 3 \end{pmatrix}$$

mátrixszal képezett lineáris transzformáció vetítés? Ortogonális vetítés-e? Milyen altérre vetít? Adjuk meg az altér magteres leírását!

Az adott $\boldsymbol{x}_0$ pont (vektor) ortogonális vetületét meghatározni egy adott altérre vonatkozóan eléggé munkaigényes feladat. A munka lépéseit vázoljuk.

Tegyük fel, hogy az altér megadása magteres alakban történt. Legyen $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_r \end{pmatrix}$, tegyük fel, hogy $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r$ lineárisan függetlenek, és az $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ M_A megoldáshalmaza az adott altér. A 2.5.6. Következmény miatt $\boldsymbol{A}^*$ képtere $M_A^\perp$, az M_A ortogonális kiegészítő altere. A 3.1. fejezet módszerét alkalmazva készítsük el $M_A^\perp$ magteres leírását!

1. Először adjuk meg a $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_{m-r}$ lineárisan független vektorokat úgy, hogy ortogonálisak legyenek az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r$ vektorokra. Itt használhatjuk az 1.4.4. Tételben megadott módszert.

2. Képezzük a $\boldsymbol{B} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_{n-r} \end{pmatrix}$ mátrixot, ennek a magtere lesz az $M_A^{\perp}$ altér, vagyis a $\boldsymbol{Bx} = \boldsymbol{0}$ megoldáshalmaza $M_A^{\perp}$.

3. Állítsuk elő a vetítősík egyenletét. A vetítősík páthuzamos $M_A^{\perp}$-sel, és átmegy az $\boldsymbol{x}_0$ ponton, tehát egyenlete $\boldsymbol{Bx} = \boldsymbol{Bx}_0$.

4. Az $\boldsymbol{x}_0$ vetülete a vetítősík és az M_A metszete, tehát meg kell oldani az

$$\begin{aligned} \boldsymbol{Ax} &= \boldsymbol{0} \\ \boldsymbol{Bx} &= \boldsymbol{Bx}_0. \end{aligned}$$

egyenletrendszert. A megoldás mindig egyértelmű (ld. 3.3.1. vagy 3.4.7. Tétel).

8. Kérdés. Mutassuk meg, hogy az

$$\boldsymbol{A} = \frac{1}{21}\begin{pmatrix} 11 & -1 & -3 & 10 \\ -1 & 2 & 6 & 1 \\ -3 & 6 & 18 & 3 \\ 10 & 1 & 3 & 11 \end{pmatrix}$$

mátrixszal megadott lineáris transzformáció vetítés! Ortogonális vetítés-e? Hány dimenziós altérre vetít?

Vetítésnél, ha az $\boldsymbol{x}$ pontot vetítjük, akkor $\boldsymbol{Ax}$ a *vetület*, és $\boldsymbol{x} - \boldsymbol{Ax}$ a *vetítő sugár*. Háromdimenziós geometriában, ortogonális vetítés esetén $\boldsymbol{Ax}$ és $\boldsymbol{x} - \boldsymbol{Ax}$ merőlegesek, és a derékszögű háromszögre a Pythagorasz-tétel érvényes. Ez magasabb dimenzióban is igaz.

3.4.6. Tétel. Ha $\boldsymbol{Ax}$ ortogonális vetítés, akkor

$$||\boldsymbol{x}||^2 = ||\boldsymbol{Ax}||^2 + ||\boldsymbol{x} - \boldsymbol{Ax}||^2.$$

Bizonyítás.

$$\begin{gathered} ||\boldsymbol{x}||^2 = ||(\boldsymbol{x} - \boldsymbol{Ax}) + \boldsymbol{Ax}||^2 = \langle(\boldsymbol{x} - \boldsymbol{Ax}) + \boldsymbol{Ax}\rangle\langle(\boldsymbol{x} - \boldsymbol{Ax}) + \boldsymbol{Ax}\rangle = \\ = ||\boldsymbol{x} - \boldsymbol{Ax}||^2 + 2\langle\boldsymbol{x} - \boldsymbol{Ax}, \boldsymbol{Ax}\rangle + ||\boldsymbol{Ax}||^2, \end{gathered}$$

de

$$\langle\boldsymbol{x} - \boldsymbol{Ax}, \boldsymbol{Ax}\rangle = \langle\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{Ax}), \boldsymbol{x}\rangle = \langle\boldsymbol{Ax} - \boldsymbol{Ax}, \boldsymbol{x}\rangle = \langle\boldsymbol{0}, \boldsymbol{x}\rangle = 0. \quad \square$$

Alkalmazzuk az egyenlőséget $\boldsymbol{x}$ helyett $\boldsymbol{x}-\boldsymbol{y}$-ra, ahol $\boldsymbol{y} \in R_A$:

$$||\boldsymbol{x}-\boldsymbol{y}||^2 = ||\boldsymbol{A}\boldsymbol{x}-\boldsymbol{y}||^2 + ||\boldsymbol{x}-\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}+\boldsymbol{A}\boldsymbol{y}||^2 =$$
$$= ||\boldsymbol{x}-\boldsymbol{A}\boldsymbol{x}||^2 + ||\boldsymbol{A}\boldsymbol{x}-\boldsymbol{y}||^2 \geq ||\boldsymbol{x}-\boldsymbol{A}\boldsymbol{x}||^2.$$

Szavakkal az egyenlőtlenség úgy interpretálható, hogy $\boldsymbol{x}$-hez az altér pontjai közül az $\boldsymbol{A}\boldsymbol{x}$ van a legközelebb, vagy másképpen, ha $\boldsymbol{x}$-et az R_A-ban lévő pontokkal a lehető legjobban akarjuk megközelíteni, akkor az ortogonális vetületet kell választani.

Az ortogonális vetület egyértelmű, ezt mondja ki a következő tétel.

3.4.7. Tétel. Az adott L altérre ortogonálisan vetítő lineáris transzformáció egyértelműen meghatározott.

Bizonyítás. Az $\boldsymbol{x}$ és az L távolsága legyen d, és tegyük fel, hogy $||\boldsymbol{x}-\boldsymbol{A}_1\boldsymbol{x}|| = d$ és $||\boldsymbol{x}-\boldsymbol{A}_2\boldsymbol{x}|| = d$. Az $\boldsymbol{x}-\boldsymbol{A}_1\boldsymbol{x}$ vektorra és az $\boldsymbol{A}_2$ vetítő transzformációra alkalmazzuk a 3.4.6. Pythagorasz-tételt:

$$||\boldsymbol{x}-\boldsymbol{A}_1\boldsymbol{x}||^2 = ||\boldsymbol{A}_2(\boldsymbol{x}-\boldsymbol{A}_1\boldsymbol{x})||^2 + ||\boldsymbol{x}-\boldsymbol{A}_1\boldsymbol{x}-\boldsymbol{A}_2(\boldsymbol{x}-\boldsymbol{A}_1\boldsymbol{x})||^2,$$

$$d^2 = ||\boldsymbol{A}_2\boldsymbol{x}-\boldsymbol{A}_1\boldsymbol{x}||^2 + ||\boldsymbol{x}-\boldsymbol{A}_2\boldsymbol{x}||^2 = ||\boldsymbol{A}_2\boldsymbol{x}-\boldsymbol{A}_1\boldsymbol{x}||^2 + d^2,$$

tehát $\boldsymbol{A}_1\boldsymbol{x} = \boldsymbol{A}_2\boldsymbol{x}$ minden $\boldsymbol{x}$-re, azaz $\boldsymbol{A}_1 = \boldsymbol{A}_2$. $\square$

A 3.4.6. Pythagorasz-tétel módot nyújt egy másik eljárásra, mellyel a vetület meghatározható. A módszert a *legkisebb négyzetek elvének* nevezik: Rögzített $\boldsymbol{x}_0$ esetén határozzuk meg az $||\boldsymbol{x}_0-\boldsymbol{y}||^2$ kifejezés minimumát $\boldsymbol{y} \in L$ esetén. Ez többváltozós szélsőérték feladat, nézzük meg kissé részletesebben.

Két esetet tárgyalunk, az altér megadási módjától függően. Első esetben az r-dimenziós L képteres alakban legyen megadva: legyen $\boldsymbol{B}\boldsymbol{z}$: $\mathbf{R}^m \to \mathbf{R}^n$ lineáris leképezés, melyre $\mathrm{rang}(\boldsymbol{B}) = r$, és $\boldsymbol{B}\boldsymbol{z}$ képtere L. (Az $\boldsymbol{A}$ vetítő transzformációt nem használhatjuk fel, hiszen nem ismerjük.) Adott még az $\boldsymbol{x}_0$ pont. Meg kell határozni az $||\boldsymbol{x}_0-\boldsymbol{B}\boldsymbol{z}||^2$ kifejezés minimumát. A parciális deriváltak nullává tétele a következő egyenletrendszert adja: jelöljük $\boldsymbol{B}$ oszlopvektorait $\boldsymbol{b}_1$, $\boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$-mel, akkor

$$-2\langle \boldsymbol{x}_0-\boldsymbol{B}\boldsymbol{z}, \boldsymbol{b}_i\rangle = 0,$$

vagyis

$$\langle \boldsymbol{B}\boldsymbol{z}, \boldsymbol{b}_i\rangle = \langle \boldsymbol{x}_0, \boldsymbol{b}_i\rangle,$$
$$\langle \boldsymbol{z}, \boldsymbol{B}^*\boldsymbol{b}_i\rangle = \langle \boldsymbol{x}_0, \boldsymbol{b}_i\rangle \quad (i = 1, 2, \ldots, m).$$

Összefoglalva, $\boldsymbol{z}$-re megoldandó az m ismeretlenes

$$\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{z} = \boldsymbol{B}^*\boldsymbol{x}_0$$

egyenletrendszer, majd kiszámítandó a $\boldsymbol{Bz}$ vetület.

Ha az altér magteres formában van megadva, akkor feltételes szélsőértéket kell számolni Lagrange-multiplikátorral. Legyen L most a $\boldsymbol{Cx} = \boldsymbol{0}$ megoldáshalmaza (magtere), ahol $\boldsymbol{Cx}$: $\mathbf{R}^n \to \mathbf{R}^s$, $s = n - r$ és rang$(\boldsymbol{C}) = s$. Keressük most $||\boldsymbol{x}_0 - \boldsymbol{x}||^2$ minimumát abban az esetben, ha $\boldsymbol{Cx} = \boldsymbol{0}$.

Képezzük a Lagrange-kifejezést,

$$||\boldsymbol{x}_0 - \boldsymbol{x}||^2 - 2\langle\lambda\ , \boldsymbol{Cx}\rangle,$$

ahol $\lambda = (\lambda_1,\ \lambda_2,\ \ldots,\lambda_s)$ ismeretlen paraméterekből álló vektor. Ennek a kifejezésnek a parciális deriváltjait tegyük egyenlővé nullával, az előző levezetéshez hasonlóan

$$\boldsymbol{x} - \boldsymbol{x}_0 + \boldsymbol{C}^*\lambda = \boldsymbol{0},$$
$$\boldsymbol{x} = \boldsymbol{x}_0 - \boldsymbol{C}^*\lambda\ .$$

Teljesülnie kell azonban a $\boldsymbol{Cx} = \boldsymbol{0}$ egyenletnek is, ebből

$$\boldsymbol{CC}^*\lambda = \boldsymbol{Cx}_0$$

adódik. Ezt az egyenletrendszert λ-ra megoldva az előző egyenletből $\boldsymbol{x}$ kiszámítható.

A vetítés témaköréhez kapcsolódik a túlhatározott egyenletrendszerek vizsgálata is. Gondoljuk el, például, hogy az $ax + by + cz = d$ egyenlet a, b, c és d együtthatóit különböző körülmények között méréssel határozzuk meg, közben az x, y és z ismeretlenek értéke változatlan marad. Kapunk a három ismeretlenre mondjuk tíz egyenletet. Határozzuk meg x, y és z értékét! A tíz egyenlet biztos, hogy ellentmondó lesz, megoldása nincs, de a legjobban közelítő megoldást még megkereshetjük. A legjobb közelítést itt úgy fogalmazhatjuk meg, hogy a $\boldsymbol{d}$ tízdimenziós vektort nem tudjuk pontosan megkapni, keresünk olyan megoldást, melyre a kapott $\boldsymbol{d}$**'** vektor ehhez $\mathbf{R}^{10}$-beli normában a legközelebb esik.

Írjuk az egyenletrendszert $\boldsymbol{Ax} = \boldsymbol{b}$ alakba. Arról van szó, hogy $\boldsymbol{b} \notin R_A$, de az $\boldsymbol{x} = \boldsymbol{x}_0$ értéket úgy akarjuk meghatározni, hogy $||\boldsymbol{Ax}_0 - \boldsymbol{b}||$ eltérés minimális legyen, vagyis, hogy $\boldsymbol{Ax}_0$ a $\boldsymbol{b}$ ortogonális vetülete legyen R_A-ra. Adott az altér képteres leírása, tehát az első módszer működik.

9. Kérdés. A 2. Kérdés egyenletrendszere, mint megállapítottuk, nem oldható meg. Oldjuk meg, mint túlhatározott egyenletrendszert!

A legkisebb négyzetek elvének legtipikusabb alkalmazása a következő. Adott az $m + 1$ dimenziós, n pontból álló pontfelhő: $(\boldsymbol{x}_i,\ b_i)$, ahol $\boldsymbol{x}_i =$

$= (x_{1i}, x_{2i}, \ldots, x_{mi})$ és $i = 1, 2, \ldots, n$. Más szóval, az $\boldsymbol{x}_i$ mérési paraméterekhez a b_i mért érték tartozik. Feltételezzük, hogy $n > m$. Keressük azt a $z = \langle \boldsymbol{a}, \boldsymbol{x} \rangle + c$ síkot, mely legjobban közelíti a mért pontokat, vagyis az

$$(\langle \boldsymbol{a}, \boldsymbol{x}_1 \rangle + c, \langle \boldsymbol{a}, \boldsymbol{x}_2 \rangle + c, \ldots, \langle \boldsymbol{a}, \boldsymbol{x}_n \rangle + c) \text{ és a } \boldsymbol{b} = (b_1, b_2, \ldots, b_n)$$

vektorok távolsága normában mérve a minimális.

Vezessünk be eggyel magasabb dimenziós $\tilde{\boldsymbol{x}}_i$ és $\tilde{\boldsymbol{a}}$ vektorokat oly módon, hogy $\tilde{\boldsymbol{x}}_i$ utolsó koordinátája, amivel kibővítettük, legyen 1, $\tilde{\boldsymbol{a}}$ utolsó koordinátája pedig c. Jelöljük továbbá $\boldsymbol{X}$-szel az $\boldsymbol{X} = (\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2, \ldots, \tilde{\boldsymbol{x}}_n)$ mátrixot, akkor ezekkel a jelölésekkel

$$\begin{gathered}(\langle \boldsymbol{a}, \boldsymbol{x}_1 \rangle + c, \langle \boldsymbol{a}, \boldsymbol{x}_2 \rangle + c, \ldots, \langle \boldsymbol{a}, \boldsymbol{x}_n \rangle + c) = \\ = (\langle \tilde{\boldsymbol{a}}, \tilde{\boldsymbol{x}}_1 \rangle, \langle \tilde{\boldsymbol{a}}, \tilde{\boldsymbol{x}}_2 \rangle, \ldots, \langle \tilde{\boldsymbol{a}}, \tilde{\boldsymbol{x}}_n \rangle) = \boldsymbol{X}^* \tilde{\boldsymbol{a}},\end{gathered}$$

a minimalizálandó kifejezés pedig a $||\boldsymbol{b} - \boldsymbol{X}^* \tilde{\boldsymbol{a}}||$ alakot ölti. Ezt a feladatot fent már megoldottuk: $\tilde{\boldsymbol{a}}$-ra megoldandó az

$$\boldsymbol{X}\boldsymbol{X}^* \tilde{\boldsymbol{a}} = \boldsymbol{X}\boldsymbol{b}$$

lineáris egyenletrendszer.

A megoldás háttere lényegében a következő eljárás: $\boldsymbol{b}$-t vetítjük $\boldsymbol{X}^*$ képterére, legyen ez $\boldsymbol{b}_0$, és megoldjuk az $\boldsymbol{X}^* \boldsymbol{a} = \boldsymbol{b}_0$ egyenletrendszert.

3.5. Az általánosított inverz és a relatív inverz

I. *Modell és definíció*

Legyen $\boldsymbol{A}$ az $\mathbf{R}^n$ lineáris transzformációja, és tegyük fel, hogy $\operatorname{rang}(\boldsymbol{A}) = r < n$. Ez esetben az $\boldsymbol{A}$: $\mathbf{R}^n \to \mathbf{R}^n$ transzformáció nem bijektív, inverze nem létezik. Kérdés lehet-e a szokásos inverz tulajdonságainak szűkítésével az inverz transzformációt legalább részben helyettesítő általánosított inverzet definiálni.

Jelöljük $\boldsymbol{A}$ képterét R_A-val, magterét M_A-val, ezek valódi, r illetve $n - r$ dimenziós alterei $\mathbf{R}^n$-nek. Az inverz legfőbb tulajdonsága a helyreállítás, azaz $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ ismeretében előállítja $\boldsymbol{x}$-et. Ha $\boldsymbol{y} \notin R_A$ akkor erre nincs remény, mert olyan $\boldsymbol{x}$, amelyre $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ nem létezik, ha viszont $\boldsymbol{y} \in R_A$, akkor végtelen sok ilyen $\boldsymbol{x}$ van. Keressük az $\mathbf{R}^n$ olyan lineáris transzformációját, mely minden $\boldsymbol{y} \in R_A$-hoz hozzárendel egy lehetséges $\boldsymbol{x}$-et, azaz olyan $\boldsymbol{x}$ vektort, melyre $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$. Kérdés azonban, hogy ilyen transzformáció létezik-e mindig?

Legyen L az M_A valamely kiegészítő altere. Megmutatjuk, hogy az $\boldsymbol{A}$: $L \to R_A$ transzformáció bijektív. Ha $\boldsymbol{x}_1 \in L$ és $\boldsymbol{x}_2 \in L$, akkor nyilván $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in L$, másrészt ha $\boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{A}\boldsymbol{x}_2$ lenne, akkor $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in M_A$, de L-nek és M_A -nak közös pontja csak a $\mathbf{0}$, tehát ekkor $\boldsymbol{x}_1 - \boldsymbol{x}_2 = \mathbf{0}$, vagyis $\boldsymbol{x}_1 = \boldsymbol{x}_2$. A transzformáció tehát injektív. Mivel L és R_A dimenziója megegyezik, a transzformáció egyúttal bijektív is (2.5.4. Következmény; a szürjektivitás közvetlenül is könnyen bizonyítható).

Az $\boldsymbol{A}$: $L \to R_A$ transzformációnak tehát létezik az inverze, legyen ez $\boldsymbol{X}$, melyre minden $\boldsymbol{y} \in R_A$ -ra $\boldsymbol{A}\boldsymbol{X}\boldsymbol{y} = \boldsymbol{y}$. Mivel R_A minden $\boldsymbol{y}$ eleme $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ alakú, minden $\boldsymbol{x}$-re $\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}$, azaz $\boldsymbol{A}\boldsymbol{X}\boldsymbol{A} = \boldsymbol{A}$. Az $\boldsymbol{X}$: $R_A \to L$ transzformáció értelmezési tartománya kiterjeszthető $\mathbf{R}^n$ –re is. Legyen $\boldsymbol{x} \in \mathbf{R}^n$, és legyen M az R_A egyik kiegészítő altere, akkor $\boldsymbol{x}$ felírható $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, $\boldsymbol{x}_1 \in R_A$ és $\boldsymbol{x}_2 \in M$ alakban, és $\boldsymbol{X}\boldsymbol{x}$ definiálható oly módon, hogy $\boldsymbol{X}\boldsymbol{x} = \boldsymbol{X}\boldsymbol{x}_1 + \boldsymbol{B}\boldsymbol{x}_2$, ahol $\boldsymbol{X}\boldsymbol{x}_1$ a korábban definiált R_A-n értelmezett transzformáció és $\boldsymbol{B}$ tetszőleges lineáris transzformáció $\mathbf{R}^n$-en. A kapott transzformáció nyilván lineáris, és kiterjesztése a korábbi $\boldsymbol{X}$-nek, mert $\boldsymbol{x}_1 \in R_A$ esetén $\boldsymbol{x}_2 = \boldsymbol{B}\boldsymbol{x}_2 = \mathbf{0}$.

3.5.1. Definíció. Az $\boldsymbol{X}$ transzformációt (mátrixot) az $\boldsymbol{A}$ *általánosított inverzének* nevezzük, ha

$$\boldsymbol{A}\boldsymbol{X}\boldsymbol{A} = \boldsymbol{A}.$$

Az elmondott modell igazolja, hogy az általánosított inverz mindig létezik, és az is látszik, hogy nem egyértelmű, egy adott $\boldsymbol{A}$-hoz több általánosított inverz is megadható.

Bizonyítható, hogy az $\boldsymbol{A}\boldsymbol{X}\boldsymbol{A} = \boldsymbol{A}$ egyenletnek eleget tevő transzformáció mindig a fenti modellel származtatható. Legyen $L = R_{XA}$, vagyis az $\boldsymbol{X}\boldsymbol{A}$ transzformáció képtere. Ekkor L dimenziója r, ugyanis

$$r = \text{rang}(\boldsymbol{A}) \geq \text{rang}(\boldsymbol{X}\boldsymbol{A}) \geq \text{rang}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}) = \text{rang}(\boldsymbol{A}) = r.$$

Továbbá ha $\boldsymbol{z}$ közös pontja L-nek és M_A-nak, akkor $\boldsymbol{A}\boldsymbol{z} = \mathbf{0}$ és $\boldsymbol{z} = \boldsymbol{X}\boldsymbol{A}\boldsymbol{u}$, amiből

$$\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\boldsymbol{X}\boldsymbol{A}\boldsymbol{u} = \boldsymbol{A}\boldsymbol{u} = \mathbf{0},$$

tehát $\boldsymbol{z} = \mathbf{0}$. E két tulajdonság bizonyítja azt, hogy L és M_A kiegészítő alterek.

Az elmondott modell alapján nyilvánvaló a következő tétel.

3.5.2. Tétel. Ha $\boldsymbol{X}$ az $\boldsymbol{A}$ általánosított inverze, akkor az $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ lineáris egyenletrendszernek egy partikuláris megoldása $\boldsymbol{x} = \boldsymbol{X}\boldsymbol{b}$, feltéve, hogy az egyenletrendszer megoldható.

Bizonyítás. Ha az egyenletrendszer megoldható, akkor $\boldsymbol{b} \in R_A$, vagyis alkalmas $\boldsymbol{z}$-re $\boldsymbol{b} = \boldsymbol{Az}$. Helyettesítsük be a feltételezett megoldást:

$$\boldsymbol{Ax} = \boldsymbol{AXb} = \boldsymbol{AXAz} = \boldsymbol{Az} = \boldsymbol{b}. \quad \square$$

Ha $\boldsymbol{X}$ általánosított inverze $\boldsymbol{A}$-nak, és egyben $\boldsymbol{A}$ is általánosított inverze $\boldsymbol{X}$-nek, akkor reflexív általánosított inverzről beszélünk. A reflexivitás nem feltétlenül teljesül.

3.5.3. Tétel. Legyen $\boldsymbol{X}$ az $\boldsymbol{A}$ általánosított inverze. Az $\boldsymbol{A}$ akkor és csak akkor általánosított inverze $\boldsymbol{X}$-nek (reflexivitás), ha rangjuk megegyezik.

Bizonyítás. $\Rightarrow$ Ha $\boldsymbol{AXA} = \boldsymbol{A}$, akkor $\operatorname{rang}(\boldsymbol{X}) \geq \operatorname{rang}(\boldsymbol{A})$. Ha $\boldsymbol{XAX} = \boldsymbol{X}$, akkor $\operatorname{rang}(\boldsymbol{A}) \geq \operatorname{rang}(\boldsymbol{X})$, vagyis $\operatorname{rang}(\boldsymbol{X}) = \operatorname{rang}(\boldsymbol{A})$.

$\Leftarrow$ Mint az előbb láttuk, $L = R_{XA}$ r-dimenziós altér. $R_{XA} \subset R_X$, de mivel R_X is r-dimenziós, $R_X = R_{XA}$. Tetszőleges $\boldsymbol{z} \in \mathbf{R}^n$-re $\boldsymbol{AXAz} = \boldsymbol{Az}$, ezért $\boldsymbol{XAX}(\boldsymbol{Az}) = \boldsymbol{X}(\boldsymbol{Az})$, vagyis R_A minden $\boldsymbol{y}$ elemére $\boldsymbol{XAXy} = \boldsymbol{Xy}$. A képterek egyenlősége miatt minden $\boldsymbol{x} \in \mathbf{R}^n$-hez van olyan $\boldsymbol{y} \in R_A$, hogy $\boldsymbol{Xx} = \boldsymbol{Xy}$, ezért minden $\boldsymbol{x}$-re $\boldsymbol{XAXx} = \boldsymbol{Xx}$, ami a reflexivitást jelenti. $\square$

Fontos speciális esetet jelent az az eset, amikor M_A és R_A kiegészítő alterek. Ekkor L választható R_A -nak, és $\boldsymbol{A}$ bijektív R_A -n, tehát az $\boldsymbol{A}$: $R_A \to R_A$ függvénynek létezik az inverze. Ez okból nevezzük az ilyen általánosított inverzet relatív inverznek, ugyanis R_A -ra szorítkozva valódi inverzet képez.

Legyen $\boldsymbol{y} \in R_A$, akkor $\boldsymbol{Ay}$ egyértelműen meghatározza $\boldsymbol{y}$-t, legyen $\boldsymbol{X}$ ez a hozzárendelés, azaz $\boldsymbol{XAy} = \boldsymbol{y}$. Mivel minden $\boldsymbol{x}$-re $\boldsymbol{y} = \boldsymbol{Ax} \in R_A$, kapjuk, hogy $\boldsymbol{XA}^2\boldsymbol{x} = \boldsymbol{Ax}$, azaz az $\boldsymbol{XA}^2 = \boldsymbol{A}$ összefüggés jellemzi a relatív inverzet.

3.5.4. Definíció. Az $\boldsymbol{X}$-et az $\boldsymbol{A}$ *relatív inverzének* nevezzük, ha $\boldsymbol{XA}^2 = \boldsymbol{A}$.

A relatív inverz nem mindig létezik, létezésének szükséges és elégséges feltételeit írja le a következő tétel.

3.5.5. Tétel. Az $\boldsymbol{A}$ lineáris transzformációra vonatkozó alábbi három állítás ekvivalens:

1. Az $\boldsymbol{A}$-nak van relatív inverze.
2. $\operatorname{rang}(\boldsymbol{A}) = \operatorname{rang}(\boldsymbol{A}^2)$.
3. Az $\boldsymbol{A}$: $R_A \to R_A$ transzformáció bijektív.

Bizonyítás. 1. $\Rightarrow$ 2. Ha $\boldsymbol{XA}^2 = \boldsymbol{A}$, akkor $\operatorname{rang}(\boldsymbol{A}) = \operatorname{rang}(\boldsymbol{XA}^2) \leq \leq \operatorname{rang}(\boldsymbol{A}^2)$ (ld. 2.5.7. Tétel). Másrészt mindig teljesül, hogy $\operatorname{rang}(\boldsymbol{A}^2) \leq \leq \operatorname{rang}(\boldsymbol{A})$, tehát $\operatorname{rang}(\boldsymbol{A}^2) = \operatorname{rang}(\boldsymbol{A})$.

2. $\Rightarrow$ 3. Mivel $R_{A^2} \subset R_A$, de dimenzióban egyeznek, ezért $R_{A^2} = R_A$. Ez azt jelenti, hogy az $\boldsymbol{A}$: $R_A \to R_A$ lineáris transzformáció szürjektív, de a 2.5.4. Következmény miatt azonban ekkor bijektív is.

3. $\Rightarrow$ 1. $\boldsymbol{X}$-et először R_A-n adjuk meg. Legyen $\boldsymbol{x} \in R_A$ tetszőleges, akkor a szürjektivitás miatt valamilyen $\boldsymbol{y} \in R_A$-ra $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{y}$. Mivel $\boldsymbol{y} \in R_A$, valamilyen $\boldsymbol{z} \in \mathbf{R}^n$-re $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{z}$, vagyis $\boldsymbol{x} = \boldsymbol{A}^2\boldsymbol{z}$. Ebben a meggondolásban az $\boldsymbol{y}$ a bijektív tulajdonság miatt az $\boldsymbol{x}$ által egyértelműen meg van határozva (a $\boldsymbol{z}$ nem feltétlenül). Az $\boldsymbol{X}\boldsymbol{x}$ lineáris transzformációt $\boldsymbol{x} \in R_A$-ra úgy definiáljuk, hogy $\boldsymbol{X}\boldsymbol{x} = \boldsymbol{y}$ legyen. Ekkor

$$\boldsymbol{X}\boldsymbol{A}^2\boldsymbol{z} = \boldsymbol{A}\boldsymbol{z},$$

és ez minden $\boldsymbol{z} \in \mathbf{R}^n$-re érvényes lesz, vagyis $\boldsymbol{X}\boldsymbol{A}^2 = \boldsymbol{A}$.

$\boldsymbol{X}$ definícióját ki kell terjeszteni $\mathbf{R}^n$-re is. Mivel a definiáló egyenlet az $\boldsymbol{X}$ számára csak R_A-beli elemekre tartalmaz előírást, a kiterjesztés a linearitás megtartásával tetszőleges lehet. L legyen R_A valamely kiegészítő altere, akkor $\boldsymbol{x} \in \mathbf{R}^n$ felírható $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, $\boldsymbol{x}_1 \in R_A$, $\boldsymbol{x}_2 \in L$ alakban. Legyen $\boldsymbol{X}\boldsymbol{x} = \boldsymbol{X}\boldsymbol{x}_1 + \boldsymbol{B}\boldsymbol{x}_2$, ahol $\boldsymbol{B}$ tetszőleges lineáris transzformáció $\mathbf{R}^n$-en. $\boldsymbol{X}$ az $\mathbf{R}^n$-en értelmezett lineáris transzformáció lesz, és egyben a korábban megadott függvény kiterjesztése marad, ugyanis $\boldsymbol{x} \in R_A$ esetén $\boldsymbol{x}_2 = \boldsymbol{0}$, és így $\boldsymbol{B}\boldsymbol{x}_2 = \boldsymbol{0}$. $\square$

A bizonyítás harmadik részéből látható, hogy $\boldsymbol{X}$ nincs egyértelműen meghatározva a teljes téren, de az értelmezési tartományátR_A-ra korlátozva már egyértelművé válik.

A 2 $\Rightarrow$ 3 alapján $R_{A^2} = R_A$, vagyis R_A minden eleme $\boldsymbol{A}^2\boldsymbol{x}$ alakú, ahol $\boldsymbol{x} \in \mathbf{R}^n$, és $\boldsymbol{A}^2\boldsymbol{x}$ mindig eleme R_A -nak. Ha $\boldsymbol{X}\boldsymbol{A}^2 = \boldsymbol{A}$, akkor $\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}^2\boldsymbol{x}$, vagyis $\boldsymbol{A}\boldsymbol{X}\boldsymbol{y} = \boldsymbol{y}$, minden $\boldsymbol{y} \in R_A$ -ra, így $\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\boldsymbol{z}$ minden $\boldsymbol{z} \in \mathbf{R}^n$-re. Beláttuk tehát a következő állítást.

3.5.6. Következmény. Minden relatív inverz egyben általánosított inverz is.

Az $\boldsymbol{X}\boldsymbol{A}^2 = \boldsymbol{A}$ egyenlettel definiált relatív inverz megfelel a modelben leírtaknak. Az $\boldsymbol{X}\boldsymbol{A}$ transzformáció képtere

$$R_{XA} = \{\boldsymbol{X}\boldsymbol{A}\boldsymbol{x} : \boldsymbol{x} \in \mathbf{R}^n\} = \{\boldsymbol{X}\boldsymbol{y} : \ \boldsymbol{y} \in R_A\} = \{\boldsymbol{X}\boldsymbol{A}^2\boldsymbol{z} : \boldsymbol{z} \in \mathbf{R}^n\} = \\ = \{\boldsymbol{A}\boldsymbol{z} : \boldsymbol{z} \in \mathbf{R}^n\} = R_A,$$

tehát L választása itt R_A.

10. Kérdés. Igaz-e, hogy bármely $\boldsymbol{A}$-hoz található olyan $\boldsymbol{P}$ permutáló mátrix, hogy $\boldsymbol{P}\boldsymbol{A}$-nak létezik a relatív inverze? (Lásd 4.4.5. Definíció)

3.5.7. Tétel. Legyen $\boldsymbol{X}$ az $\boldsymbol{A}$ relatív inverze. Az $\boldsymbol{A}$ akkor és csak akkor relatív inverze $\boldsymbol{X}$-nek (reflexivitás), ha rangjuk megegyezik.

Bizonyítás. Ha a rangok különböznek, akkor $\boldsymbol{A}$ még általánosított inverz sem lehet. Ellenkező esetben, $R_X \subset R_A$, de mivel a dimenziójuk egyenlő, $R_X = R_A$. Ez azt jelenti, hogy minden $\boldsymbol{y}$-hoz van olyan $\boldsymbol{x}$, hogy $\boldsymbol{A}\boldsymbol{x} =$

$= \boldsymbol{Xy}$. Ezért
$$\boldsymbol{AX}^2\boldsymbol{y} = \boldsymbol{AXAx} = \boldsymbol{Ax} = \boldsymbol{Xy},$$
ami bizonyítja az állítást. □

II. *Explicit előállítás*

A következőkben többször is felhasználjuk az önmagában is érdekes alábbi segédtételt.

3.5.8. Segédtétel. Legyen az $n \times n$-es $\boldsymbol{A}$ mátrix $\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix}$ szerkezetű, ahol $\boldsymbol{A}_0$ $r \times r$-es teljesrangú mátrix. $\boldsymbol{A}$ rangja akkor és csak akkor r, ha $\boldsymbol{W} = \boldsymbol{V}\boldsymbol{A}_0^{-1}\boldsymbol{U}$.

Bizonyítás. $\Rightarrow$ Képezzük a $\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{V}\boldsymbol{A}_0^{-1} & \boldsymbol{I} \end{pmatrix}$ mátrixot és balról szorozzuk be vele az $\boldsymbol{A}$-t (a hipermátrixokkal a szorzás a számokkal történő mátrixszorzás mintájára végezhető el):
$$\boldsymbol{QA} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{V}\boldsymbol{A}_0^{-1} & \boldsymbol{I} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{0} & \boldsymbol{W} - \boldsymbol{V}\boldsymbol{A}_0^{-1}\boldsymbol{U} \end{pmatrix}.$$
Mivel $\boldsymbol{Q}$ teljesrangú mátrix, $\boldsymbol{QA}$ rangja is r, amiből $\boldsymbol{W} - \boldsymbol{V}\boldsymbol{A}_0^{-1}\boldsymbol{U} = \boldsymbol{0}$ következik.
$\Leftarrow$ Ha $\boldsymbol{W} = \boldsymbol{V}\boldsymbol{A}_0^{-1}\boldsymbol{U}$, akkor $\boldsymbol{QA}$ rangja nyilván r, és így $\boldsymbol{A}$ rangja is r. □

(Az analógia kedvéért jegyezzük meg, hogy $a \neq 0$ esetén az $\begin{pmatrix} a & u \\ v & w \end{pmatrix}$ mátrix rangja akkor és csak akkor 1, ha a determinánsa 0, vagyis, ha $w = \frac{vu}{a}$.)

Tetszőleges $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix}$ mátrixhoz, ahol $\operatorname{rang}(\boldsymbol{A}) = \operatorname{rang}(\boldsymbol{A}_0)$, igen egyszerűen megadható reflexív általánosított inverz, mégpedig az $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{A}_0^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$ alakban. Könnyen ellenőrizhető a 3.5.8. Segédtétel felhasználásával, hogy
$$\boldsymbol{AXA} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_0^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix} =$$
$$= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{V}\boldsymbol{A}_0^{-1} & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{V}\boldsymbol{A}_0^{-1}\boldsymbol{U} \end{pmatrix} = \boldsymbol{A}$$
és
$$\boldsymbol{XAX} = \begin{pmatrix} \boldsymbol{A}_0^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_0^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} =$$

$$= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A}_0^{-1}\boldsymbol{U} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}_0^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} = \boldsymbol{X}.$$

A megadott $\boldsymbol{X}$ általában nem relatív inverz, noha $\boldsymbol{A}$ relatív inverze mindig létezik.

Az általánosított inverz általános alakja is megadható. Itt az alapgondolat a következő: ha $\boldsymbol{X}$ általánosított inverze $\boldsymbol{A}$-nak, akkor $\boldsymbol{Q}_2^{-1}\boldsymbol{X}\boldsymbol{Q}_1$ is általánosított inverze $\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_2$-nek bármely invertálható $\boldsymbol{Q}_1$ és $\boldsymbol{Q}_2$ esetén. Az $\boldsymbol{AXA} = \boldsymbol{A}$ relációból ugyanis következik, hogy

$$\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_2\boldsymbol{Q}_2^{-1}\boldsymbol{X}\boldsymbol{Q}_1\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_2 = \boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_2.$$

1. $\boldsymbol{A}$ legyen tetszőleges r rangú mátrix, akkor oszlopcserékkel és sorcserékkel elérhető, hogy a bal felső sarokban álló $r \times r$-es mátrix rangja r legyen. A sorcserék alkalmas $\boldsymbol{P}_1$ permutáló mátrixszal balról, az oszlopcserék $\boldsymbol{P}_2$ permutáló mátrixszal jobbról történő szorzással végezhetők el. Így feltehető, hogy $\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{P}_2$ már a kívánt alakú.

2. Feltehető tehát, hogy $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix}$ alakú, ahol $\boldsymbol{A}_0$ $r\times$ r-es teljesrangú mátrix. A

$$\boldsymbol{Q}'_1 = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{V}\boldsymbol{A}_0^{-1} & \boldsymbol{I} \end{pmatrix} \text{ és a } \boldsymbol{Q}'_2 = \begin{pmatrix} \boldsymbol{I} & -\boldsymbol{A}_0^{-1}\boldsymbol{U} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix}$$

mátrixokkal balról, ill jobbról szorozva $\boldsymbol{A}$-t, $\boldsymbol{A}$ a kívánt alakot ölti: az $\boldsymbol{A}_0$-tól különböző elemei nullázódnak:

$$\boldsymbol{A}_2 = \boldsymbol{Q}'_1\boldsymbol{A}\boldsymbol{Q}'_2 = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

3. Az $\boldsymbol{A}_2$ általános inverze könnyen kiszámítható, hiszen az

$$\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{D} & \boldsymbol{E} \\ \boldsymbol{F} & \boldsymbol{G} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$$

egyenletnek az általános megoldása $\boldsymbol{D} = \boldsymbol{A}_0^{-1}$, és $\boldsymbol{E}$, $\boldsymbol{F}$, $\boldsymbol{G}$ tetszőleges. $\boldsymbol{A}_2$ általánosított inverze tehát

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{A}_0^{-1} & \boldsymbol{E} \\ \boldsymbol{F} & \boldsymbol{G} \end{pmatrix}.$$

4. Vezessük be a $\boldsymbol{Q}_1^{-1} = \boldsymbol{Q}'_1\boldsymbol{P}_1$ és a $\boldsymbol{Q}_2 = \boldsymbol{P}_2\boldsymbol{Q}'_2$ jelöléseket, akkor $\boldsymbol{A}_2 = \boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_2$, ennek általánosított inverze $\boldsymbol{Y}$, és $\boldsymbol{X} = \boldsymbol{Q}_2\boldsymbol{Y}\boldsymbol{Q}_1^{-1}$.

A relatív inverz explicit megadásához az előzőekhez hasonló észrevétel jelenti az alapgondolatot. Ha $\boldsymbol{X}$ relatív inverze $\boldsymbol{A}$-nak, akkor $\boldsymbol{Q}^{-1}\boldsymbol{X}\boldsymbol{Q}$ relatív inverze $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$-nak bármely invertálható $\boldsymbol{Q}$ esetén. Az állítás igazolásához azt kell belátni, hogy az $\boldsymbol{X}\boldsymbol{A}^2 = \boldsymbol{A}$ összefüggésből következik, hogy

$$\boldsymbol{Q}^{-1}\boldsymbol{X}\boldsymbol{Q}\ \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}\ \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q},$$

de ez magától értetődő. Az alapgondolatot az jelenti, hogy képezve az $\boldsymbol{A}$ mátrix $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ alakú transzformáltját, hozzuk $\boldsymbol{A}$-t minél egyszerűbb alakra. Az erősebb korlátozást az jelenti, hogy mindkét oldalon ugyanaz a mátrix, illetve annak inverze szerepel.

1. Ha $\operatorname{rang}(\boldsymbol{A}) = r$, akkor oszlopcserékkel elérhető, hogy az $\boldsymbol{A}$ első r oszlopa lineárisan független legyen. Az oszlopcserét végezzük el alkalmas $\boldsymbol{P}$ permutáló mátrixszal történő jobb oldali szorzással (ld. 4.5.4. Definíció). Az $\boldsymbol{A}\boldsymbol{P}$ mátrixot azonban a $\boldsymbol{P}^{-1}$ permutáló mátrixszal szorozni kell balról is, ami sorcseréket jelent, azonban a sorcserék az első r oszlopvektor lineáris függetlenségét nem befolyásolják. Feltehető tehát, hogy $\boldsymbol{A}$ első r oszlopvektora lineáris független. (A bal felső $r \times r$-es részmátrix itt nem biztos, hogy teljesrangú.)

2. Ha $\operatorname{rang}(\boldsymbol{A}) = r$, és az első r oszlopvektor lineárisan független, akkor a többi oszlopvektor már ezek lineáris kombinációjaként előállítható, vagyis

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{A}_0\boldsymbol{C} \\ \boldsymbol{V} & \boldsymbol{V}\boldsymbol{C} \end{pmatrix},$$

ahol $\boldsymbol{C}$ megfelelően választott $r \times (n-r)$-es mátrix. Válasszuk meg a $\boldsymbol{Q}_1$ mátrixot

$$\boldsymbol{Q}_1 = \begin{pmatrix} \boldsymbol{I} & -\boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix}$$

alakban, akkor

$$\boldsymbol{A}_1 = \boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1 = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{A}_0\boldsymbol{C} \\ \boldsymbol{V} & \boldsymbol{V}\boldsymbol{C} \end{pmatrix}\begin{pmatrix} \boldsymbol{I} & -\boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} =$$
$$= \begin{pmatrix} \boldsymbol{A}_0 + \boldsymbol{C}\boldsymbol{V} & \boldsymbol{0} \\ \boldsymbol{V} & \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{V} & \boldsymbol{0} \end{pmatrix}.$$

Ha $\boldsymbol{A}$ relatív inverze létezik, akkor $\boldsymbol{A}_1 = \boldsymbol{Q}_1^{-1}\boldsymbol{A}Q_1$ relatív inverze is létezik, tehát a négyzetének a rangja r. Ezt felírva

$$r = \operatorname{rang}(\boldsymbol{A}_1^2) = \operatorname{rang}\left(\begin{pmatrix} \boldsymbol{B}^2 & \boldsymbol{0} \\ \boldsymbol{V}\boldsymbol{B} & \boldsymbol{0} \end{pmatrix}\right) = \operatorname{rang}\left(\begin{pmatrix} \boldsymbol{B}^2 \\ \boldsymbol{V}\boldsymbol{B} \end{pmatrix}\right) =$$

$$= \operatorname{rang}(\begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{V} \end{pmatrix} \boldsymbol{B}) \leq \operatorname{rang}(\boldsymbol{B}) \leq r,$$

amiből kapjuk, hogy $\boldsymbol{B} = \boldsymbol{A}_0 + \boldsymbol{C}\boldsymbol{V}$ invertálható mátrix.

3. Alakítsuk tovább az $\boldsymbol{A}_1$ mátrixot. Válasszuk meg $\boldsymbol{Q}_2$ mátrixot:

$$\boldsymbol{Q}_2 = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{V}\boldsymbol{B}^{-1} & \boldsymbol{I} \end{pmatrix},$$

akkor

$$\boldsymbol{A}_2 = \boldsymbol{Q}_2^{-1}\boldsymbol{A}_1\boldsymbol{Q}_2 = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{V}\boldsymbol{B}^{-1} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{V} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{V}\boldsymbol{B}^{-1} & \boldsymbol{I} \end{pmatrix} =$$
$$= \begin{pmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

4. Jelöljük $\boldsymbol{Y}$-nal $\boldsymbol{A}_2$ relatív inverzét, és keressük $\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{D} & \boldsymbol{E} \\ \boldsymbol{F} & \boldsymbol{G} \end{pmatrix}$ alakban, akkor

$$\boldsymbol{Y}\boldsymbol{A}_2^2 = \begin{pmatrix} \boldsymbol{D} & \boldsymbol{E} \\ \boldsymbol{F} & \boldsymbol{G} \end{pmatrix} \begin{pmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{D}\boldsymbol{B}^2 & \boldsymbol{0} \\ \boldsymbol{F}\boldsymbol{B}^2 & \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

és ez akkor és csak akkor teljesül, ha $\boldsymbol{D} = \boldsymbol{B}^{-1}$ és $\boldsymbol{F} = \boldsymbol{0}$, az $\boldsymbol{E}$ és $\boldsymbol{G}$ részmátrixok tetszőlegesek.

$\boldsymbol{A}_2$ relatív inverze tehát

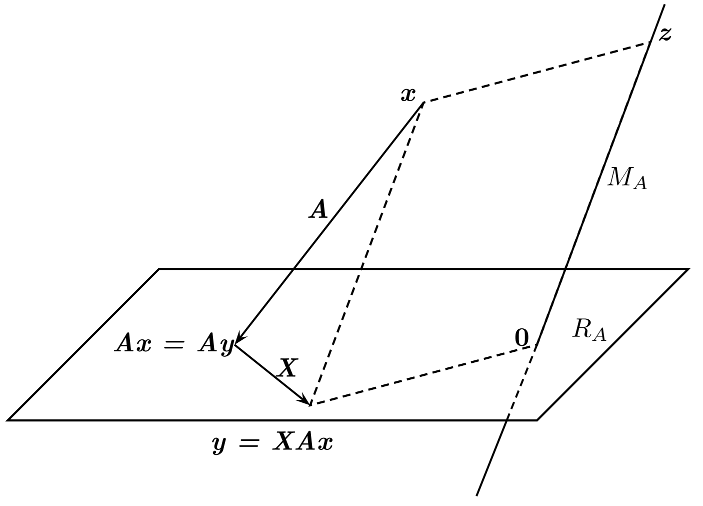
$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{B}^{-1} & \boldsymbol{E} \\ \boldsymbol{0} & \boldsymbol{G} \end{pmatrix}.$$

5. Az $\boldsymbol{A}_2$ relatív inverzéből a $\boldsymbol{Q}_2$, $\boldsymbol{Q}_1$ és a $\boldsymbol{P}$ segítségével már könnyen előállíthatjuk $\boldsymbol{A}$ inverzét.

11. Kérdés. Megfelel-e a relatív inverz előállított alakja az általánosított inverzre kapott alaknak?

III. *Kapcsolat a vetítésekkel*

Legyen az $\boldsymbol{A}$ lineáris transzformáció magtere M_A képtere R_A, és tegyük fel, hogy ezek kiegészítő alterek, akkor az $\boldsymbol{x} \in \mathbf{R}^n$ felbontható $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$ alakba, ahol $\boldsymbol{y} \in R_A$ és $\boldsymbol{x} \in M_A$. Az $\boldsymbol{y}$ az $\boldsymbol{x}$-nek az R_A-ra képezett vetülete az M_A vetítő sík esetén. Az $\boldsymbol{A}\boldsymbol{x}$ az R_A valamely pontja, és mivel $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{0}$, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{y}$. Ha $\boldsymbol{X}$ az $\boldsymbol{A}$ relatív inverze, akkor $\boldsymbol{X}\boldsymbol{A}\boldsymbol{y} = \boldsymbol{y}$, de ekkor $\boldsymbol{X}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ is teljesül, ami a vetítő transzformáció előállítását adja. $\boldsymbol{XA}$ tehát az R_A-ra történő vetítés. Lásd az ábrát!



Ennél lényegesen több is igaz.

3.5.9. Tétel. Ha $\boldsymbol{X}$ az $\boldsymbol{A}$ általánosított inverze, akkor $\boldsymbol{AX}$ és $\boldsymbol{XA}$ egyaránt vetítések. $\boldsymbol{AX}$ képtere R_A, és $\boldsymbol{XA}$ vetítő síkja M_A. Ha $\boldsymbol{X}$ relatív inverz, akkor ráadásul $\boldsymbol{XA}$ képtere is R_A.

Bizonyítás. Ha $\boldsymbol{AXA} = \boldsymbol{A}$, akkor $\boldsymbol{AXAX} = \boldsymbol{AX}$, azaz $(\boldsymbol{AX})^2 = \boldsymbol{AX}$, valamint $\boldsymbol{XAXA} = \boldsymbol{XA}$, azaz $(\boldsymbol{XA})^2 = \boldsymbol{XA}$, vagyis $\boldsymbol{AX}$ és $\boldsymbol{XA}$ egyaránt vetítés.

$R_{AX} \subset R_A$, mert $R_A = \{\boldsymbol{Ax}\colon\ \boldsymbol{x} \in \mathbf{R}^n\} \supset \{\boldsymbol{Ax}\colon\ \boldsymbol{x} \in R_X\} = \{\boldsymbol{AXx}\colon\ \boldsymbol{x} \in \mathbf{R}^n\} = R_{AX}$. Másrészt ugyanilyen okból $R_{AXA} = R_A \subset \subset R_{AX}$, tehát $R_A = R_{AX}$.

Az $\boldsymbol{XA}$ vetítő síkja $V = \{\boldsymbol{x}\colon\ \boldsymbol{XAx} = \boldsymbol{0}\}$. Ha $\boldsymbol{x} \in M_A$, akkor $\boldsymbol{Ax} = \boldsymbol{0}$, és $\boldsymbol{XAx} = \boldsymbol{0}$, vagyis $\boldsymbol{x} \in V$. Ha $\boldsymbol{x} \in V$, akkor $\boldsymbol{XAx} = \boldsymbol{0}$, és $\boldsymbol{AXAx} = \boldsymbol{Ax} = \boldsymbol{0}$, vagyis $\boldsymbol{x} \in M_A$.

Ha $\boldsymbol{X}$ relatív inverz, akkor $\boldsymbol{A} = \boldsymbol{XA}^2$ miatt $R_A = \{\boldsymbol{XA}^2\boldsymbol{x} : \boldsymbol{x} \in \mathbf{R}^n\} = \{\boldsymbol{XAy} : \boldsymbol{y} \in R_A\} \subset R_{XA}$. Másrészt $\operatorname{rang}(\boldsymbol{XA}) \le \operatorname{rang}(\boldsymbol{A})$, így R_{XA} dimenziója nem lehet nagyobb, mint R_A dimenziója, tehát $R_{XA} = R_A$. $\square$

Az átalánosított inverz, valamint az arról elmondottak lehetőséget adnak arra, hogy egy adott altérre történő ortogonális vetítés mátrixát meghatározzuk. 3.4.-ben adott pont vetületét legkisebb négyzetek módszerével már ki tudtuk számolni, de az a módszer kevésbé alkalmas a transzformáció meghatározására. Az itt közölt módszert az előző ábra szemlélteti, de ebben az esetben az ábrán M_A ortogonális R_A-ra. Először módszerként, lépésekre bontva tárgyaljuk, majd explicit képletet adunk az ortogonális vetítő mátrix meghatározására.

1. Az r-dimenziós L altér, melyre a vetítés történik, megadható generáló vektoraival, vagy képteres formában (ld. 3.1.). A generáló vektorok ismeretében a képteres megadás mátrixa közvetlenül felírható: az $n \times r$-es mátrix oszlopvektorai legyenek az alteret generáló lineárisan független vektorok. A mátrixot bővítsük ki $n \times n$-essé nullákat hozzávéve, és tegyük fel, hogy az első r darab sorvektor is lineárisan független. A kapott $\boldsymbol{A}$ mátrix a következő alakú

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_0 & 0 \\ \boldsymbol{V} & 0 \end{pmatrix},$$

ahol $\operatorname{rang}(\boldsymbol{A}_0) = r$ és L az $\boldsymbol{A}$ transzformáció képtere. Az első r sorvektor lineáris függetlensége sorcserékkel mindig elérhető, az eredményként kapott vetítő mátrixban a sorcseréket – oszlopcserékkel párosítva – vissza kell cserélni.

2. Ezt a transzformációt kell úgy átalakítani, hogy képtere, L ne változzon meg, és magtere, $L^{\perp}$ legyen. Vezessük be a $\boldsymbol{C} = \boldsymbol{V}\boldsymbol{A}_0^{-1}$ jelölést, akkor az átalakítás az

$$\boldsymbol{A}_1 = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{A}_0\boldsymbol{C}^* \\ \boldsymbol{V} & \boldsymbol{V}\boldsymbol{C}^* \end{pmatrix}$$

képlettel hajtható végre.

Bizonyítás. $R_A \subset R_{A_1}$, mert az $\begin{pmatrix} \boldsymbol{A}_0 \\ \boldsymbol{V} \end{pmatrix}$ oszlopvektorai közösek a két mátrixban. R_{A_1} dimenziója szintén r a 3.5.8. Segédtétel miatt, tehát $R_{A_1} = R_A = L$.

Ha bebizonyítjuk, hogy $R_{A_1} = R_{A_1^*} = L$, akkor a 2.5.6. Következmény miatt $M_{A_1} = L^{\perp}$. Az $\boldsymbol{A}_1$ mátrix első r sorvektora $R_{A_1^*}$ bázisát képezi, írjuk ezeket oszlopvektorként, akkor az

$$\begin{pmatrix} \boldsymbol{A}_0^* \\ \boldsymbol{C}\boldsymbol{A}_0^* \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_0 \\ \boldsymbol{V} \end{pmatrix} \boldsymbol{A}_0^{-1}\boldsymbol{A}_0^*$$

átalakítás mutatja, hogy $\boldsymbol{A}_1$-ben az első r sorvektor az első r oszlopvektor lineáris kombinációja, tehát $R_{A_1} = R_{A_1^*}$. $\square$

3. Határozzuk meg $\boldsymbol{A}_1$ mátrix $\boldsymbol{X}$ relatív inverzét a levezetett általános alak alapján. A $\boldsymbol{Q}_1 = \begin{pmatrix} \boldsymbol{I} & -\boldsymbol{C}^* \\ 0 & \boldsymbol{I} \end{pmatrix}$ választás mellett az $\boldsymbol{A}_2$ az $\boldsymbol{A}_2 = \boldsymbol{Q}_1^{-1}\boldsymbol{A}_1\boldsymbol{Q}_1 = \begin{pmatrix} \boldsymbol{A}_0 + \boldsymbol{C}^*\boldsymbol{V} & 0 \\ \boldsymbol{V} & 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{B} & 0 \\ \boldsymbol{V} & 0 \end{pmatrix}$ alakot ölti. A $\boldsymbol{Q}_2$ legyen $\boldsymbol{Q}_2 = \begin{pmatrix} \boldsymbol{I} & 0 \\ \boldsymbol{V}\boldsymbol{B}^{-1} & \boldsymbol{I} \end{pmatrix}$, akkor $\boldsymbol{A}_3 = \boldsymbol{Q}_2^{-1}\boldsymbol{A}_2\boldsymbol{Q}_2 = \begin{pmatrix} \boldsymbol{B} & 0 \\ 0 & 0 \end{pmatrix}$. Az $\boldsymbol{A}_3$

relatív inverze tehát $\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{B}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$, a korábbi $\boldsymbol{F}$ és a $\boldsymbol{G}$ nullmátrixnak vehető. Visszatranszformálással az $\boldsymbol{A}_1$ relatív inverzének kiszámítására az $\boldsymbol{X} = \boldsymbol{Q}_1\boldsymbol{Q}_2\boldsymbol{Y}\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1}$ képlet alkalmazható.

4. Az ortogonális vetítés előáll az $\boldsymbol{X}\boldsymbol{A}_1$ szorzat alakjában a 3.5.8. Tétel alapján.

5. Számítsuk ki $\boldsymbol{X}\boldsymbol{A}_1$-et!

$$\boldsymbol{X}\boldsymbol{A}_1 = \boldsymbol{Q}_1\boldsymbol{Q}_2\boldsymbol{Y}\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1}\boldsymbol{Q}_1\boldsymbol{Q}_2\boldsymbol{A}_3\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1} =$$
$$= \boldsymbol{Q}_1\boldsymbol{Q}_2\boldsymbol{Y}\boldsymbol{A}_3\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1} = \boldsymbol{Q}_1\boldsymbol{Q}_2\begin{pmatrix} \boldsymbol{I} & 0 \\ 0 & 0 \end{pmatrix}\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1}.$$

A szorzásokat elvégezve

$$\boldsymbol{X}\boldsymbol{A}_1 = \begin{pmatrix} \boldsymbol{I} - \boldsymbol{C}^*\boldsymbol{V}\boldsymbol{B}^{-1} & (\boldsymbol{I} - \boldsymbol{C}^*\boldsymbol{V}\boldsymbol{B}^{-1})\boldsymbol{C}^* \\ \boldsymbol{V}\boldsymbol{B}^{-1} & \boldsymbol{V}\boldsymbol{B}^{-1}\boldsymbol{C}^* \end{pmatrix}.$$

A formula azonban, mint az alábbi tétel mutatja, egyszerűbb alakra hozható.

A tétel fontosságára való tekintettel közvetlen bizonyítást is adunk rá (2. Bizonyítás).

3.5.10. Tétel. Legyen az $\boldsymbol{A}$ lineáris transzformáció $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_0 & 0 \\ \boldsymbol{V} & 0 \end{pmatrix}$ alakú, ahol $\boldsymbol{A}_0$ $r \times r$-es teljesrangú mátrix. Az $\boldsymbol{A}$ képterére történő vetítés mátrixa

$$\boldsymbol{R} = \begin{pmatrix} \boldsymbol{D} & \boldsymbol{D}\boldsymbol{C}^* \\ \boldsymbol{C}\boldsymbol{D} & \boldsymbol{C}\boldsymbol{D}\boldsymbol{C}^* \end{pmatrix},$$

ahol $\boldsymbol{C} = \boldsymbol{V}\boldsymbol{A}_0^{-1}$ és $\boldsymbol{D} = (\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})^{-1}$.

1. Bizonyítás. Mivel a fentebbi mátrixban $\boldsymbol{B} = \boldsymbol{A}_0 + \boldsymbol{C}^*\boldsymbol{V}$,

$$\boldsymbol{I} - \boldsymbol{C}^*\boldsymbol{V}\boldsymbol{B}^{-1} = (\boldsymbol{B} - \boldsymbol{C}^*\boldsymbol{V})\boldsymbol{B}^{-1} = \boldsymbol{A}_0(\boldsymbol{A}_0 + \boldsymbol{C}^*\boldsymbol{V})^{-1} =$$
$$= (\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{V}\boldsymbol{A}_0^{-1})^{-1} = (\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})^{-1}.$$

Vezessük be a $\boldsymbol{D} = (\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})^{-1}$ jelölést, akkor hasonlóan átalakítható a

$$\boldsymbol{V}\boldsymbol{B}^{-1} = \boldsymbol{V}(\boldsymbol{A}_0 + \boldsymbol{C}^*\boldsymbol{V})^{-1} = \boldsymbol{V}((\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{V}\boldsymbol{A}_0^{-1})\boldsymbol{A}_0)^{-1} = \boldsymbol{V}\boldsymbol{A}_0^{-1}\boldsymbol{D} = \boldsymbol{C}\boldsymbol{D}$$

kifejezés is. Az $(\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})^{-1}$ inverz létezése a $\boldsymbol{B}^{-1}$ létezéséből következik. □

2. Bizonyítás. 1. $\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C}$ invertálható mátrix. Nézzük a magterét, vagyis azon $\boldsymbol{x}$-ek halmazát, melyre $(\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})\boldsymbol{x} = \boldsymbol{0}$, azaz $\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{x} = -\boldsymbol{x}$. Ekkor $||\boldsymbol{x}||^2 = \langle \boldsymbol{x}, -\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{x}\rangle = -\langle \boldsymbol{C}\boldsymbol{x}, \boldsymbol{C}\boldsymbol{x}\rangle = -||\boldsymbol{C}\boldsymbol{x}||^2$, vagyis $\boldsymbol{x} = \boldsymbol{0}$, ami az invertálhatóságot jelenti.

2. $\boldsymbol{R}$ képtere a megadott altér. $\boldsymbol{R}$ első r oszlopvektora mátrix alakban

$$\begin{pmatrix} \boldsymbol{D} \\ \boldsymbol{CD} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} \\ \boldsymbol{VA}_0^{-1} \end{pmatrix} \boldsymbol{D} = \begin{pmatrix} \boldsymbol{A}_0 \\ \boldsymbol{V} \end{pmatrix} \boldsymbol{A}_0^{-1}\boldsymbol{D},$$

amiből látható, hogy $\boldsymbol{R}$ első r oszlopvektora a megadott vektorok lineáris kombinációja, és a rangjuk r. A $\boldsymbol{R}$ további oszlopvektorai szemmel láthatóan már lineáris kombinációi az első r oszlopvektornak, tehát $\boldsymbol{R}$ rangja is r.

3. $\boldsymbol{R}$ vetítés.

$$\boldsymbol{R}^2 = \begin{pmatrix} \boldsymbol{D}^2 + \boldsymbol{DC}^*\boldsymbol{CD} & (\boldsymbol{D}^2 + \boldsymbol{DC}^*\boldsymbol{CD})\boldsymbol{C}^* \\ \boldsymbol{CD}^2 + \boldsymbol{CDC}^*\boldsymbol{CD} & (\boldsymbol{CD}^2 + \boldsymbol{CDC}^*\boldsymbol{CD})\boldsymbol{C}^* \end{pmatrix},$$

ahol $\boldsymbol{D}^2 + \boldsymbol{DC}^*\boldsymbol{CD} = \boldsymbol{D}(\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})\boldsymbol{D} = \boldsymbol{D}$, és $\boldsymbol{CD}^2 + \boldsymbol{CDC}^*\boldsymbol{CD} = \boldsymbol{CD}(\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})\boldsymbol{D} = \boldsymbol{CD}$, tehát $\boldsymbol{R}^2 = \boldsymbol{R}$.

4. $\boldsymbol{R}$ ortogonális vetítés, mert mátrixa szimmetrikus, hiszen $\boldsymbol{D}$ szimmetrikus mátrix. □

A módszer vagy a képlet bonyolultnak tűnik, de az elvégzendő két mátrixinvertálás azonban csak $r \times r$-es mátrixokra vonatkozik. A gyakorlati végrehajtást illetően lásd a következő kérdésre adott választ.

12. Kérdés. Mi annak a lineáris transzformációnak a mátrixa, mely $\mathbf{R}^4$-ben ortogonálisan vetít a (2, 5, 5, 1) és (1, 3, 2, 1) vektorok által meghatározott síkra? (Használjuk az MsExcelt az 5.8.-ban leírtak szerint.)

IV. *Lineáris egyenletrendszer általános megoldása*

Láttuk, hogy a lineáris egyenletrendszer egy partikuláris megoldása, ha létezik, előállítható az általánosított inverzzel. Azt mondhatjuk, hogy ezért találták ki az általánosított inverzet. Az általános megoldás is megadható egyetlen általánosított inverz segítségével.

3.5.11. Tétel. Ha az $\boldsymbol{A}$ négyzetes mátrix valamelyik általánosított inverzét $\boldsymbol{X}$-szel jelöljük, és az $\boldsymbol{Ax} = \boldsymbol{b}$ egyenlet megoldható, akkor $\boldsymbol{x} = \boldsymbol{Xb}$ megoldása az egyenletrendszernek. Az egyenletrendszer általános megoldása

$$\boldsymbol{x} = \boldsymbol{Xb} + (\boldsymbol{I} - \boldsymbol{XA})\boldsymbol{t}$$

alakú, ahol $\boldsymbol{t} \in \mathbf{R}^n$ tetszőlegesen megválasztható vektor.

Bizonyítás. Azt, hogy $\boldsymbol{x} = \boldsymbol{Xb}$ kielégíti az egyenletrendszert, már korábban megmutattuk.

Az $\boldsymbol{Ax} = \boldsymbol{0}$ homogén lineáris egyenletrendszer megoldása az M_A magtér, ami a 3.5.9. Tétel miatt megegyezik $\boldsymbol{XA}$ vetítés magterével, a 3.4.3. Tétel értelmében ez utóbbi pedig azonos $\boldsymbol{I} - \boldsymbol{XA}$ képterével, vagyis $M_A = M_{XA} = R_{I-XA}$. R_{I-XA} minden eleme $(\boldsymbol{I} - \boldsymbol{XA})\boldsymbol{t}$ alakú, tehát $\boldsymbol{Ax} = \boldsymbol{0}$ összes megoldása $\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{XA})\boldsymbol{t}$ alakban írható fel. Könnyen látható, hogy minden ilyen alakú kifejezés megoldás:

$$\boldsymbol{A}(\boldsymbol{I} - \boldsymbol{XA})\boldsymbol{t} = \boldsymbol{At} - \boldsymbol{AXAt} = \boldsymbol{At} - \boldsymbol{At} = \boldsymbol{0}.$$

Az $\boldsymbol{Ax} = \boldsymbol{b}$ általános megoldása a kettő összegéből adódik. $\square$

Egy pillanatra megütközünk azon, hogy n szabad paraméter függvényében állítottuk elő az általános megoldást, holott úgy tudjuk, hogy a szabad változók száma $n - r$, ahol $r = \operatorname{rang}(\boldsymbol{A})$. Voltaképpen ez nem így van. R_{I-XA} dimenziója valóban $n - r$, így $\boldsymbol{I} - \boldsymbol{XA}$ oszlopvektorai közül $n - r$ lineárisan függetlent kiválasztva ezek lineáris kombinációi is előállítják a képteret. Így $\boldsymbol{t}$ azon koordinátái, melyek indexei nem egyeznek meg a kiválasztott lineárisan független oszlopvektorok indexeivel, nyugodtan nullázhatók, és ily módon csupán r szabad paramétere van a megoldásseregnek.

A 3.6.1. Tétel beváltja az általánosított inverzhez fűződő reményeinket, hiszen az egyenletrendszer megoldása hasonlóan történik, mint a létező inverz esetén. Túlzott reményeket azonban ne fűzzünk hozzá, a szokásosaktól eltérő megoldási módszert nem ad. Ha az általánosított inverznek a 3.5.6. Definíciót követő legegyszerűbb alakját választjuk, a megoldási módszer azonos az 5.7. fejezetben leírtakkal.

Ha ismerjük az általánosított inverz egy partikuláris alakját – márpedig ez könnyen előállítható az előzőek szerint – akkor az általános alakot is megadja az alábbi tétel.

3.5.12. Tétel. Ha $\boldsymbol{X}$ az $\boldsymbol{A}$ általánosított inverze, akkor az $\boldsymbol{A}$ általánosított inverzeinek általános alakja

$$\boldsymbol{X} + (\boldsymbol{I} - \boldsymbol{XA})\boldsymbol{B} + \boldsymbol{C}(\boldsymbol{I} - \boldsymbol{AX}),$$

ahol $\boldsymbol{B}$ és $\boldsymbol{C}$ tetszőleges $n \times n$-es mátrixok.

Bizonyítás. $\Leftarrow$ Az $\boldsymbol{X}_0 = \boldsymbol{X} + (\boldsymbol{I} - \boldsymbol{XA})\boldsymbol{B} + \boldsymbol{C}(\boldsymbol{I} - \boldsymbol{AX})$ általánosított inverz, mert

$$\begin{aligned}\boldsymbol{AX}_0\boldsymbol{A} &= \boldsymbol{AXA} + \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{XA})\boldsymbol{BA} + \boldsymbol{AC}(\boldsymbol{I} - \boldsymbol{AX})\boldsymbol{A} = \\ &= \boldsymbol{A} + (\boldsymbol{A} - \boldsymbol{AXA})\boldsymbol{BA} + \boldsymbol{AC}(\boldsymbol{A} - \boldsymbol{AXA}) = \boldsymbol{A} + \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{A}.\end{aligned}$$

$\Rightarrow$ Ha $\boldsymbol{X}$ és $\boldsymbol{X}_0$ két általánosított inverz, akkor $\boldsymbol{Y} = \boldsymbol{X}_0 - \boldsymbol{X}$ olyan transzformáció, melyre $\boldsymbol{AYA} = \boldsymbol{0}$, másképpen fogalmazva, $\boldsymbol{Y}$ az $\boldsymbol{A}$ képterét, R_A-t, az $\boldsymbol{A}$ magterébe, M_A-ba képezi le.

Vezessük be a $\boldsymbol{V} = \boldsymbol{I} - \boldsymbol{XA}$, $\boldsymbol{W} = \boldsymbol{I} - \boldsymbol{AX}$ egyszerűsített jelöléseket, akkor ezek a 3.4.3. Tétel szerint vetítések. $\boldsymbol{W}$ magtere R_A (a 3.4.3. és 3.5.9. Tételek szerint), képterét jelöljük L-lel. Mivel $\boldsymbol{W}$ vetítés, R_A és L

kiegészítő alterek (3.4.2. Tétel). Az $\boldsymbol{x} \in \mathbf{R}^n$ felírható $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ alakban, ahol $\boldsymbol{x}_1 \in R_A$ és $\boldsymbol{x}_2 \in L$, így $\boldsymbol{Yx} = \boldsymbol{Yx}_1 + \boldsymbol{Yx}_2 = \boldsymbol{Dx} + \boldsymbol{Cx}$, vagyis $\boldsymbol{Y} = \boldsymbol{D} + \boldsymbol{C}$, ahol $\boldsymbol{Dx} = \boldsymbol{Yx}_1$, és $\boldsymbol{Cx} = \boldsymbol{Yx}_2$.

Ha $\boldsymbol{x} \in R_A$, akkor $\boldsymbol{x}_2 = \boldsymbol{0}$, tehát $\boldsymbol{Cx} = \boldsymbol{0}$. Ha $\boldsymbol{x} \in \mathbf{R}^n$, és $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ a fentiek szerint, akkor $\boldsymbol{Cx} = \boldsymbol{Cx}_2 = \boldsymbol{CW}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \boldsymbol{CWx}$, mert $\boldsymbol{W}$ magtere R_A és képtere L. Ezzel az előállítás második tagját megkaptuk.

Ha $\boldsymbol{x} \in \mathbf{R}^n$, akkor $\boldsymbol{Dx} = \boldsymbol{Yx}_1 \in M_A$, tehát $\boldsymbol{D}\colon \mathbf{R}^n \to M_A$ leképezés. Legyen $\boldsymbol{B} = \boldsymbol{I} + \boldsymbol{D} - \boldsymbol{V}$, akkor $\boldsymbol{VBx} = \boldsymbol{V}(\boldsymbol{I} + \boldsymbol{D} - \boldsymbol{V})\boldsymbol{x} = \boldsymbol{Vx} + \boldsymbol{VDx} - \boldsymbol{Vx} = \boldsymbol{Dx}$, mert $\boldsymbol{Dx} \in M_A$ és $\boldsymbol{V}$ az M_A-ra vetít. Ezzel a $\boldsymbol{DX}$-et is előállítottuk a kívánt alakban. □

Alkalmazzuk a tétel eredményét az $\boldsymbol{Ax} = \boldsymbol{b}$ egyenletrendszer megoldására. Ha feltételezzük, hogy az egyenletrendszer megoldható és $\boldsymbol{b} \neq \boldsymbol{0}$, akkor a 3.5.11. Tétel eredményét kapjuk meg, hiszen $\boldsymbol{b} \in R_A$ esetén $\boldsymbol{C}(\boldsymbol{I} - \boldsymbol{AX})\boldsymbol{b} = \boldsymbol{0}$ és $\boldsymbol{t} = \boldsymbol{Cb}$ a $\boldsymbol{C}$ alkalmas választásával mindig elérhető. A homogén egyenletekre is érvényes 3.5.11. Tétel a 3.5.12.-ből közvetlenül nem következik.

3.6. Válaszok a kérdésekre

1. $\mathbf{R}^4$-ben legyen az L altér bázisa $\boldsymbol{a} = (1, 2, 3, 0)$ és $\boldsymbol{b} = (1, 3, 2, 1)$. Adjuk meg azt a lineáris leképezést, amelynek L a magtere. Alkalmazzuk az 1.4.4. Tétel konstrukciós eljárását!

Mivel $\boldsymbol{e}_4$ ortogonális $\boldsymbol{a}$-ra, az eljárás előírása szerint $\boldsymbol{b}_1 = (0, 0, 0, 1)$. $\boldsymbol{b}_2$-t $\boldsymbol{b}_2 = (1, \alpha, 0, 0)$ alakban keressük. Ahhoz, hogy ez ortogonális legyen $\boldsymbol{a}$-ra $\alpha = -\frac{1}{2}$-nek választandó, tehát $\boldsymbol{b}_2 = (1, -\frac{1}{2}, 0, 0)$. $\boldsymbol{b}_3$-at $\boldsymbol{b}_3 = (0, 1, \alpha, 0)$ alakban keresve $\boldsymbol{a}$-ra ortogonális lesz, ha $\alpha = -\frac{2}{3}$; tehát $\boldsymbol{b}_3 = (0, 1, -\frac{2}{3}, 0)$. $\boldsymbol{c}_1 = \boldsymbol{b}_1 + \alpha \boldsymbol{b}_2$ alapján a $\boldsymbol{b}$-re vonatkozó ortogonalitás $\alpha = 2$ esetén teljesül, így $\boldsymbol{c}_1 = (2, -1, 0, 1)$. $\boldsymbol{c}_2 = \boldsymbol{b}_2 + \alpha \boldsymbol{b}_3$-ból $\alpha = \frac{3}{10}$, ezért $\boldsymbol{c}_2 = (1, -\frac{1}{5}, -\frac{1}{5}, 0)$. $\boldsymbol{c}_2$ helyett vehetünk $\boldsymbol{c}_2 = (5, -1, -1, 0)$-t.

A kapott $\boldsymbol{c}_1$ és $\boldsymbol{c}_2$ vektorok ortogonálisak $\boldsymbol{a}$-ra és $\boldsymbol{b}$-re, így az L minden (x, y, z, w) elemére is, ezt felírva

$$\begin{aligned} 2x - y + w &= 0 \\ 5x - y - z &= 0, \end{aligned}$$

és ezzel megkaptuk a lineáris leképezést, melynek magtere L.

2. Megoldható-e az

$$x - y + 2z + t = 6$$

$$\begin{aligned} 3x-2y+z-t&=0\\ x-3z-3t&=1 \end{aligned}$$

egyenletrendszer?

Két mátrix rangját kellene kiszámítani, de ez egyszerre elvégezhető. Írjuk fel a kibővített mátrixot, és számoljunk rangot, de fogadjuk meg, hogy az utolsó oszlopot (a konstansok oszlopát) aktív oszlopként nem használjuk! Jelzésként szaggatott vonallal válasszuk el.

$$\left(\begin{array}{cccc:c} 1 & -1 & 2 & 1 & 6\\ 3 & -2 & 1 & -1 & 0\\ 1 & 0 & -3 & -3 & 1 \end{array}\right) \sim \left(\begin{array}{cccc:c} 1 & -1 & 5 & 4 & 5\\ 3 & -2 & 10 & 8 & -3\\ 1 & 0 & 0 & 0 & 0 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{cccc:c} 0 & -1 & 5 & 4 & 5\\ 0 & -2 & 10 & 8 & -3\\ 1 & 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{cccc:c} 0 & -1 & 5 & 4 & 5\\ 0 & 0 & 0 & 0 & -13\\ 1 & 0 & 0 & 0 & 0 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{cccc:c} 0 & -1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & -13\\ 1 & 0 & 0 & 0 & 0 \end{array}\right),$$

ebből látható, hogy az egyenletrendszer mátrixának a rangja 2, míg a kibővített mátrix rangja 3, tehát az egyenletrendszer nem oldható meg.

3. Az

$$\begin{aligned} 5x+y-z&=5\\ x-y+z&=1\\ x+y-z&=1 \end{aligned}$$

egyenletrendszernek $x=1$, $y=3$, $z=3$ megoldása. Van-e más megoldás?

A megadott megoldás helyes, tehát az egyenletrendszer megoldható. A mátrixának a rangja $r=2$, tehát $m>r$, a megoldások száma végtelen, így van más megoldás is.

4. Az

$$\begin{aligned} x+y-z+2w&=0\\ x+z-w&=0 \end{aligned}$$

megoldásai hány dimenziós alteret alkotnak? Adjuk meg az általános megoldást és a megoldáshalmaz egy bázisát!

A homogén egyenletrendszer mindig megoldható. Az egyenletrendszer mátrixának a rangja $r = 2$, tehát $s = m - r = 2$, vagyis a megoldások halmaza (a magtér) 2-dimenziós. Az általános megoldás:

$$\begin{aligned} x &= -z + w \\ y &= 2z - 3w \end{aligned}$$

z és w tetszőleges.

Az általános megoldásból láthatóan az (x, y, z, w) lineáris kombinációja a (-1, 2, 1, 0) és az (1, -3, 0, 1) vektoroknak.

5. Milyen helyzetű a következő két sík?

$$\begin{aligned} S_1 : \quad & 2x + y + w = 4 \\ & -x + 2y + z = 2 \\ S_2 : \quad & x + y + 2z + w = 5 \\ & x + 3y + z + w = 4 \end{aligned}$$

Rangszámolással eldönthető, hogy a négy egyenletből álló egyenletrendszer nem oldható meg, tehát a síkoknak nincs közös pontjuk. Mivel mindkét sík kétdimenziós, párhuzamosak csak úgy lehetnének, ha az ortogonális kiegészítő altereik megegyeznek. A második sík (1, 1, 2, 1) vektora azonban nincs benne a (2, 1, 0, 1) és a $(-1, 2, 1, 0)$ által generált altérben, hiszen a három vektor lineárisan független. A síkok tehát kitérők.

6. Jelöljük az $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ minta átlagát $\bar{x}$-sal, és képezzük az $\boldsymbol{y} = (x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x})$ vektort (ld. 2. fejezet 2. Kérdés). Mátrixalgebra felhasználása nélkül mutassuk meg, hogy a lineáris transzformáció, mely az $\boldsymbol{x}$-et az $\boldsymbol{y}$-ba képezi le, ortogonális vetítés! Mennyi annak az altérnek a dimenziója, amelyre a vetítés történik?

Mivel $\boldsymbol{y}$ koordinátáinak az átlaga nulla, az eljárást $\boldsymbol{y}$-ra alkalmazva újra $\boldsymbol{y}$-t kapunk, tehát vetítésről van szó. A $\mathbf{0}$-ra képződő $\boldsymbol{x}$ vektorokra $x_i = \bar{x}$ minden i-re, tehát minden koordinátájuk egyenlő, azaz $c{\cdot}\mathbf{1}$ alakúak. $\langle \mathbf{1}, \boldsymbol{y} \rangle$ azonban 0, vagyis a vetítősugár merőleges minden vetületre. Ezért a vetítés ortogonális vetítés. (Ez a mátrixának szimmetriájából is látszik.) Mivel az ortogonális kiegészítő altér egydimenziós, a vetítés $(n-1)$-dimenziós altérre történik.

7. Igaz-e, hogy az

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

mátrixszal képezett lineáris transzformáció vetítés? Ortogonális vetítés-e? Milyen altérre vetít? Adjuk meg az altér magteres leírását!

Az $\boldsymbol{A}^2 = \boldsymbol{A}$ összefüggés könnyen igazolható. Mivel $\boldsymbol{A}$ nem szimmetrikus mátrix (azaz $\boldsymbol{A}^* \neq \boldsymbol{A}$), $\boldsymbol{A}$ nem lehet ortogonális vetítés.

rang$(\boldsymbol{A}) = 2$, tehát R_A kétdimenziós és bázisa pl. az első két oszlopvektor (az első helyett egyszerűbb a felét venni): $\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ és $\boldsymbol{a}_2 = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$. Ezekre ortogonális vektor $\mathbf{1} = (1,\ 1,\ 1)$, könnyen kapható 1.1.4. Tételben leírt ortogonalizálással, vagy közvetlenül, egyenletrendszerrel. A vektorra ortogonális altér egyenlete $\langle \mathbf{1}, \boldsymbol{x} = 0\rangle$, vagy másképpen

$$x_1 + x_2 + x_3 = 0.$$

8. Mutassuk meg, hogy az

$$A = \frac{1}{21}\begin{pmatrix} 11 & -1 & -3 & 10 \\ -1 & 2 & 6 & 1 \\ -3 & 6 & 18 & 3 \\ 10 & 1 & 3 & 11 \end{pmatrix}$$

mátrixszal megadott lineáris transzformáció vetítés! Ortogonális vetítés-e? Hány dimenziós altérre vetít?

Az $\boldsymbol{A}^2 = \boldsymbol{A}$ összefüggés könnyen igazolható. Mivel $\boldsymbol{A}$ szimmetrikus mátrix (azaz $\boldsymbol{A}^* = \boldsymbol{A}$), $\boldsymbol{A}$ ortogonális vetítés. rang$(\boldsymbol{A}) = 2$, tehát kétdimenziós altérre történik a vetítés.

9. A 2. Kérdés egyenletrendszere,

$$\begin{aligned} x - y + 2z + t &= 6 \\ 3x - 2y + z - t &= 0 \\ x - 3z - 3t &= 1, \end{aligned}$$

mint megállapítottuk, nem oldható meg. Oldjuk meg, mint túlhatározott egyenletrendszert!

Ha az egyenletrendszer $\boldsymbol{Bx} = \boldsymbol{b}$ alakú, akkor a levezetett képlet alapján $\boldsymbol{x}$-et a $\boldsymbol{B}^*\boldsymbol{Bx} = \boldsymbol{B}^*\boldsymbol{b}$ egyenletrendszerből kell meghatározni.

$$\boldsymbol{B}^*\boldsymbol{B} = \begin{pmatrix} 11 & -7 & 2 & -5 \\ -7 & 5 & -4 & 1 \\ 2 & -4 & 14 & 10 \\ -5 & 1 & 10 & 11 \end{pmatrix} \text{ és } \boldsymbol{B}^*\boldsymbol{b} = \begin{pmatrix} 7 \\ -6 \\ 9 \\ 3 \end{pmatrix},$$

$\operatorname{rang}(\boldsymbol{B}^*\boldsymbol{B}) = 2$, ezért az egyenletrendszernek végtelen sok megoldása van. Ilyen például $x = \frac{23}{24}, y = z = 0,\ t = \frac{17}{24}$. A vetületi pont, $\boldsymbol{Bx}$ azonban egyértelmű.

A végtelen sok megoldás túlhatározott egyenletrendszernél nem tipikus. Itt voltaképpen kevés egyenlet áll a rendelkezésünkre, más kérdés, hogy azok között is ellentmondás van.

10. Kérdés. Igaz-e, hogy bármely $\boldsymbol{A}$-hoz található olyan $\boldsymbol{P}$ permutáló mátrix, hogy $\boldsymbol{PA}$-nak létezik a relatív inverze?

Legyen $\boldsymbol{A}$ rangja r. Oszlopcserékkel elérhető, hogy az első r oszlopvektor lineárisan független legyen, vagyis $\boldsymbol{AP}_1$ már ilyen tulajdonságú mátrix, és itt $\boldsymbol{P}_1$ valamilyen permutáló mátrix. További sorcserékkel elérhető, hogy a bal felső $r \times r$-es részmátrix rangja r legyen, tehát a $\boldsymbol{B} = \boldsymbol{P}_2\boldsymbol{AP}_1$ már ilyen tulajdonságú mátrix. Mivel $\operatorname{rang}(\boldsymbol{B}^2) = r = \operatorname{rang}(\boldsymbol{B})$, $\boldsymbol{B}$-nek létezik a relatív inverze, de ekkor $\boldsymbol{P}_1\boldsymbol{BP}^{-1} = \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{A}$-nak is létezik, tehát $\boldsymbol{P} = \boldsymbol{P}_1\boldsymbol{P}_2$ választható.

11. Megfelel-e a relatív inverz előállított alakja az általánosított inverzre kapott alaknak?

A kérdés természetesen arra vonatkozik, hogy a relatív inverz a 3.5.4. Következmény értelmében egyúttal általánosított inverz is, tehát a megadott előállítását az általánosított inverz alakjában is fel kell tudni írni. A $\boldsymbol{B} = \boldsymbol{A}_0^2 + \boldsymbol{U}\mathbf{V}$ rövidített jelöléssel legyen

$$\boldsymbol{X} = \boldsymbol{Q}_1^{-1} \begin{pmatrix} \boldsymbol{A}_0\boldsymbol{B}^{-1} & \boldsymbol{D} \\ \mathbf{0} & \boldsymbol{E} \end{pmatrix} \boldsymbol{Q}_1 = \boldsymbol{Q}_2\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1} \begin{pmatrix} \boldsymbol{A}_0\boldsymbol{B}^{-1} & \boldsymbol{D} \\ \mathbf{0} & \boldsymbol{E} \end{pmatrix} \boldsymbol{Q}_1,$$

akkor az alábbi mátrixot kell kiszámítani:

$$\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1} \begin{pmatrix} \boldsymbol{A}_0\boldsymbol{B}^{-1} & \boldsymbol{D} \\ \mathbf{0} & \boldsymbol{E} \end{pmatrix} = \boldsymbol{Q}_2^{-1} \begin{pmatrix} \boldsymbol{I} & \mathbf{0} \\ \boldsymbol{VA}_0^{-1} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}_0\boldsymbol{B}^{-1} & \boldsymbol{D} \\ \mathbf{0} & \boldsymbol{E} \end{pmatrix} =$$

$$= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A}_0^{-1}\boldsymbol{U} \\ \mathbf{0} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}_0\boldsymbol{B}^{-1} & \boldsymbol{D} \\ \boldsymbol{VB}^{-1} & \boldsymbol{VA}_0^{-1}\boldsymbol{D} + \boldsymbol{E} \end{pmatrix} =$$

$$= \begin{pmatrix} \boldsymbol{A}_0\boldsymbol{B}^{-1} + \boldsymbol{A}_0^{-1}\boldsymbol{U}\boldsymbol{V}\boldsymbol{B}^{-1} & \cdots \\ \cdots & \cdots \end{pmatrix}.$$

Ebben a kifejezésben $\boldsymbol{A}_0\boldsymbol{B}^{-1}+\boldsymbol{A}_0^{-1}\boldsymbol{U}\boldsymbol{V}\boldsymbol{B}^{-1} = \boldsymbol{A}_0^{-1}(\boldsymbol{A}_0^2+\boldsymbol{U}\boldsymbol{V})\boldsymbol{B}^{-1} = \boldsymbol{A}_0^{-1}$, és ezt kellett megmutatnunk.

12. Mi annak a lineáris transzformációnak a mátrixa, mely $\mathbf{R}^4$-ben ortogonálisan vetít a (2, 5, 5, 1) és (1, 3, 2, 1) vektorok által meghatározott síkra? (Használjuk az MsExcelt az 5.8.-ban leírtak szerint.)

Alkalmazzuk a 3.5.10. Tétel formuláját! A megadottak alapján $\boldsymbol{A}_0 = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$, $\boldsymbol{V} = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$, ebből $\boldsymbol{A}_0^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$, és $\boldsymbol{C} = \boldsymbol{V}\boldsymbol{A}^{-1} = \begin{pmatrix} 5 & -1 \\ -2 & 1 \end{pmatrix}$. Újabb invertálással számítsuk ki $\boldsymbol{D}$-t:

$$\boldsymbol{D} = (\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})^{-1} = \frac{1}{41}\begin{pmatrix} 3 & 7 \\ 7 & 30 \end{pmatrix},$$

és töltsük ki a vetítő mátrixot:

$$\begin{pmatrix} \boldsymbol{D} & \boldsymbol{D}\boldsymbol{C}^* \\ \boldsymbol{C}\boldsymbol{D} & \boldsymbol{C}\boldsymbol{D}\boldsymbol{C}^* \end{pmatrix} = \frac{1}{41}\begin{pmatrix} 3 & 7 & 8 & 1 \\ 7 & 30 & 5 & 16 \\ 8 & 5 & 35 & -11 \\ 1 & 16 & -11 & 14 \end{pmatrix}.$$

Az eredmény könnyen ellenőrizhető: Egyrészt szimmetrikus mátrix (hiszen ortogonális vetítés), másrészt négyzete önmaga (hiszen vetítés), harmadsorban pedig a mátrix valamennyi oszlopvektora lineárisan függ a megadott vektoroktól.

4. fejezet

Bázistranszformációk

4.1. Vektor és mátrix transzformálása

A tér bizonyos jelenségei, geometriai sajátosságai, alaki tulajdonságai változatlanok maradnak, ha a tér bázisát megváltoztatjuk. A távolság- és szögviszonyok addig nem változnak meg amíg az új bázist „ferdeszögű koordinátarendszerként" kezeljük. De ha ezen bázis mellett a normát és a skalárszorzatot a régi képlettel számoljuk, akkor a ferdeszögű koordináta-rendszert „kiegyenesítjük" és vektorait egységvektoroknak tekintjük, ekkor az így kapott érték megváltozik. Mindezek ellenére a bázis alkalmas megválasztása számos feladat megoldását segíti elő, mind a lineáris algebra, mind a lineáris programozás témakörében.

A fejezetben az $\mathbf{R}^n$ átalakításáról lesz szó, tehát $\mathbf{R}^n \to \mathbf{R}^n$ lineáris transzformációkról, és a hozzá tartozó mátrixok is mindig $n \times n$-es, négyzetes mátrixok. Az $\mathbf{R}^n$ eredeti, derékszögűnek vélt bázisát $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$-nel fogjuk jelölni. Ha áttérünk az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ bázisra, akkor természetesen megköveteljük az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ vektorok lineáris függetlenségét. Az áttérés egy lineáris transzformációt jelent, mely az $\boldsymbol{e}_i$ vektorokat átviszi az $\boldsymbol{a}_i$ vektorokba, és ezáltal az egész térnek is egy transzformációját definiálja. Ennek a transzformációnak a mátrixa $\boldsymbol{C} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$. Eddig is már többször használtuk és bizonyítottuk a következő egyszerű állítást.

4.1.1. Tétel. Az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ vektorrendszer akkor és csak akkor lineárisan független (és egyben akkor és csak akkor bázisa $\mathbf{R}^n$-nek), ha $\operatorname{rang}(\boldsymbol{C}) = n$.

Bizonyítás. A rang definíciójából következik. □

A $\boldsymbol{C}$ bázistranszformációról mindig feltételezzük, hogy teljesrangú, ezért ha $\boldsymbol{C}$ bázistranszformáció, akkor mindig invertálható, azaz $\boldsymbol{C}^{-1}$ mindig létezik.

A $\boldsymbol{Cx}$ a transzformált $\boldsymbol{x}$ vektor régi bázisra vonatkozó koordinátáit adja meg. Többnyire azonban nem erre van szükség, hanem az eredeti $\boldsymbol{x}$ vektor új bázisra vonatkozó koordinátái érdekelnek.

1. Kérdés. Az 1. fejezet 11. Kérdése során láttuk, hogy az origóba helyezett egységkocka origóból kiinduló $(n-1)$-dimenziós lapátló vektorai $\mathbf{R}^n$ bázisát alkotják. Adjuk meg a testátló vektor koordinátáit ebben a bázisban!

4.1.2. Tétel. Az $\boldsymbol{x}$ vektor új bázisra vonatkozó koordinátái megegyeznek a $\boldsymbol{C}^{-1}\boldsymbol{x}$ vektor régi bázisra vonatkozó koordinátáival.

Bizonyítás. Mivel $\boldsymbol{a}_i = \boldsymbol{Ce}_i$, az új bázis $\boldsymbol{Ce}_1$, $\boldsymbol{Ce}_2, \ldots, \boldsymbol{Ce}_n$ alakban is írható. Jelöljük $\boldsymbol{x}$ koordinátáit az új bázisban $x'_1, x'_2, ..., x'_n$-nel, akkor $\boldsymbol{x} = \sum\limits_{k=1}^{n} x'_k \boldsymbol{C}\mathbf{e}_k$. Ebből $\boldsymbol{C}^{-1}\boldsymbol{x} = \sum\limits_{k=1}^{n} x'_k \boldsymbol{e}_k$, ami bizonyítja az állítást. □

A lineáris programozás egyik fő módszere a lépésenként végrehajtott bázistranszformáció, ami azt jelenti, hogy a bázisvektorokat egyenként cserélik le. Lépésenként haladva egyrészt jobban lehet alkalmazkodni a feladat jellegéhez, másrészt $\boldsymbol{C}^{-1}$ meghatározása nem mindig megy könnyen.

4.1.3. Definíció. Azt a bázistranszformációt, mely a bázisban az $\boldsymbol{e}_i$ vektort adott $\boldsymbol{a}$-re cseréli (feltételezve, hogy $\langle \boldsymbol{a}, \boldsymbol{e}_i \rangle \neq 0$), de a többi vektort meghagyja, *elemi bázistranszformációnak* nevezzük. A transzformáció mátrixa: $\boldsymbol{C} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_{i-1}, \boldsymbol{a}, \boldsymbol{e}_{i+1}, \ldots, \boldsymbol{e}_n)$.

Legyen $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$, akkor alapvető feltétel, hogy $\langle \boldsymbol{a}, \boldsymbol{e}_i \rangle = = a_i \neq 0$, ugyanis ebben az esetben $\boldsymbol{C}$ i-edik sora nulla lévén nem teljesrangú mátrix, így nem hoz létre újabb bázist. Ha $a_i \neq 0$ teljesül, akkor könnyen ellenőrizhetően

$$\boldsymbol{C}^{-1} = (\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_{i-1}, \boldsymbol{c}, \boldsymbol{e}_{i+1}, ..., \boldsymbol{e}_n),$$

ahol

$$\boldsymbol{c} = \frac{1}{a_i}(-a_1, -a_2, ..., -a_{i-1}, 1\ , -a_{i+1}, ..., -a_n).$$

A $\boldsymbol{C}^{-1}\boldsymbol{x}$ számszerű meghatározására viszonylag egyszerű, táblázatos eljárást dolgoztak ki (ld. 4.2.).

A másik fontos kérdés az, hogy hogyan alakul át egy lineáris transzformáció mátrixa, ha a bázist megváltoztatjuk.

4.1.4. Tétel. Tekintsük az $\boldsymbol{y} = \boldsymbol{Ax}$ lineáris transzformációt. Ha a bázison a $\boldsymbol{C}$ lineáris transzformációt hajtjuk végre, akkor az új bázis mellett a korábbi lineáris transzformáció mátrixa $\boldsymbol{C}^{-1}\boldsymbol{AC}$ lesz.

Bizonyítás. Az $\boldsymbol{A}$ transzformáció az $\boldsymbol{x}$-hez az $\boldsymbol{Ax}$-et rendeli hozzá, ami az új bázisra vonatkoztatva a $\boldsymbol{C}^{-1}\boldsymbol{x}$-hez történő $\boldsymbol{C}^{-1}\boldsymbol{Ax}$ hozzárendelést jelenti (4.1.2. Tétel). Ha a $\boldsymbol{C}^{-1}\boldsymbol{x}$-re alkalmazzuk a $\boldsymbol{C}^{-1}\boldsymbol{AC}$ transzformációt, akkor $(\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C})(\boldsymbol{C}^{-1}\boldsymbol{x}) = \boldsymbol{C}^{-1}\boldsymbol{A}(\boldsymbol{C}\boldsymbol{C}^{-1})\boldsymbol{x} = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{x}$, tehát éppen ezt a hozzárendelést kapjuk. $\square$

$\boldsymbol{AB}$ szorzat transzformálása tényezőnként számítható:

$$\boldsymbol{C}^{-1}\boldsymbol{ABC} = \boldsymbol{C}^{-1}\boldsymbol{A}(\boldsymbol{C}\boldsymbol{C}^{-1})\boldsymbol{BC} = (\boldsymbol{C}^{-1}\boldsymbol{AC})(\boldsymbol{C}^{-1}\boldsymbol{BC}).$$

Ennek következménye, hogy mátrixok hatványának számítása sok esetben megkönnyíthető, ha a mátrix az új bázisban egyszerűbb alakú, ugyanis

$$\boldsymbol{C}^{-1}\boldsymbol{A}^n\boldsymbol{C} = (\boldsymbol{C}^{-1}\boldsymbol{AC})^n.$$

4.1.5. Definíció. Két mátrixot, $\boldsymbol{A}$-t és $\boldsymbol{B}$-t *hasonlónak* nevezünk, ha van olyan $\boldsymbol{C}$ invertálható mátrix, hogy $\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{AC}$.

A hasonlóság az $n \times n$-es mátrixok világában ekvivalencia reláció. Ez három tulajdonságot követel meg:

a) $\boldsymbol{A}$ hasonló $\boldsymbol{A}$-hoz, ugyanis $\boldsymbol{A} = \boldsymbol{I}^{-1}\boldsymbol{AI}$. (Reflexív tulajdonság.)

b) Ha $\boldsymbol{A}$ hasonló $\boldsymbol{B}$-hez, akkor $\boldsymbol{B}$ is hasonló $\boldsymbol{A}$-hoz. (A definíció eleve szimmetrikusnak kezelte a relációt: $\boldsymbol{A}$-t és $\boldsymbol{B}$-t hasonlónak nevezte.) Ha ugyanis $\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{AC}$, akkor $\boldsymbol{C}$-vel balról, $\boldsymbol{C}^{-1}$-gyel jobbról beszorozva $\boldsymbol{CBC}^{-1} = \boldsymbol{A}$, ami a $\boldsymbol{C}^{-1}$ mátrixszal képezett hasonlósági reláció.

c) Ha $\boldsymbol{A}$ hasonló $\boldsymbol{A}_1$-hez és $\boldsymbol{A}_1$ hasonló $\boldsymbol{B}$-hez, akkor $\boldsymbol{A}$ is hasonló $\boldsymbol{B}$-hez (tranzitív tulajdonság). Ha ugyanis $\boldsymbol{A}_1 = \boldsymbol{C}^{-1}\boldsymbol{AC}$ és $\boldsymbol{B} = \boldsymbol{D}^{-1}\boldsymbol{A}_1\boldsymbol{D}$, akkor $\boldsymbol{B} = \boldsymbol{D}^{-1}(\boldsymbol{C}^{-1}\boldsymbol{AC})\boldsymbol{D} = (\boldsymbol{CD})^{-1}\boldsymbol{A}(\boldsymbol{CD})$.

2. Kérdés. Hasonlók-e az

$$\begin{pmatrix} 1 & 3 & 1 \\ 2 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix} \text{ és az } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

mátrixok?

4.2. Elemi bázistranszformáció

Az elemi, lépésenként végrehajtott bázistranszformációról az előző pontban már volt szó, legalábbis a definícióját kimondtuk (4.1.3. Definíció). Itt a gyakorlatban történő numerikus végrehajtásáról, annak táblázatos formájáról lesz szó. Ennek hibamentes elsajátítása a lineáris programozás alappillére.

Készítsünk táblázatot a szereplő vektorokkal az alábbi formában. Tüntessük fel a bázist, a bázisba bevonandó $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ vektort és a nyomon követendő $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ vektort (ld. a következő oldal táblázatát).

	$\boldsymbol{a}$	$\boldsymbol{x}$
$\boldsymbol{e}_1$	$\boxed{a_1}$	x_1
$\boldsymbol{e}_2$	a_2	x_2
$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{e}_n$	a_n	x_n

Ha $\boldsymbol{e}_1$ helyére $\boldsymbol{a}$-t hozunk be a bázisba, akkor a 4.1.3-at követő képlet szerint $\boldsymbol{x}$ új koordinátái $\boldsymbol{C}^{-1}\boldsymbol{x} = (\frac{x_1}{a_1},\ x_2 - \frac{x_1}{a_1}a_2,\ x_3 - \frac{x_1}{a_1}a_3, ..., x_n - \frac{x_1}{a_1}a_n)$ lesznek. Vezessük be a $\delta = \frac{x_1}{a_1}$ jelölést, akkor $\boldsymbol{C}^{-1}\boldsymbol{x} = (\delta,\ x_2 - \delta a_2,$ $x_3 - \delta a_3, ..., x_n - \delta a_n)$ kissé egyszerűbb képlet adódik. A táblázatban az a_1 pozíciójával tudjuk jelölni, hogy az $\boldsymbol{e}_1$ helyére az $\boldsymbol{a}$-t hozzuk be, az a_1-et az elemi bázistranszformáció *generáló elemének* nevezzük, és ezt a táblázatban az a_1 bekeretezése jelöli. Generáló elemnek nulla nem választható, ebben az esetben a létrejövő új bázis nem teljesítené a lineáris függetlenség kritériumát. Ha $a_1 = 0$, akkor $\boldsymbol{a} = a_2\boldsymbol{e}_2 + \ldots + a_n\boldsymbol{e}_n$, ami mutatja a lineáris összefüggést. Az új táblázat:

	$\boldsymbol{x}$
$\boldsymbol{a}$	δ
$\boldsymbol{e}_2$	$x_2 - \delta a_2$
$\vdots$	$\vdots$
$\boldsymbol{e}_n$	$x_n - \delta a_n$

Szavakkal elmondva:

1. Kiszámítjuk a δ értékét. (Más generáló elemre más és más értéket kapunk! Gyakran az első táblázat utolsó sora után be is írjuk ezt az értéket.)

2. A generáló érték sorába beírjuk a δ-t. (A többi sor más módon számítódik!)

3. A többi sorban az $\boldsymbol{x}$ régi koordinátájából levonjuk az $\boldsymbol{a}$ megfelelő koordinátájának δ-szorosát.

Természetesen több vektor sorsa is követhető az elemi bázistranszformáció elvégzésekor. Ekkor a táblázat több oszlopból áll, és minden oszlopra el kell végezni a kijelölt transzformációt. A generáló elem értelemszerűen bármely sorból választható.

Ha a_1 a generáló elem, akkor a táblázatból $\boldsymbol{e}_1$ eltűnik. Ha szükség van rá a továbbiakban, akkor megőrzendő vektorként fel kell venni a korábbi táblázat egy további oszlopába (koordinátái $(1, 0, 0, \ldots, 0)$), és erre vonatkozóan is meg kell állapítani az új koordinátákat.

Az elemi bázistranszformáció az új helyzetre vonatkozóan további báziselem lecserélésével folytatható.

3. Kérdés. Egy elemi bázistranszformációt alkalmazva az $\boldsymbol{x}$ vektor (3, 0, 2, 1) koordinátái (1, 2, -1, 0)-ra változtak. Megállapítható-e, hogy mely bázisvektor maradt biztosan a régi? Megállapítható-e, hogy melyiket cseréltük le? Mik a bázisba bevont vektor eredeti koordinátái?

4.3. Az elemi bázistranszformáció közvetlen felhasználásai

I. *Lineáris függőség vizsgálata*

Egy vektor akkor és csak akkor vihető be a bázisba, ha a kapott vektorrendszer lineárisan független marad. Ez akkor és csak akkor valósul meg, ha a generáló elem nem nulla. Ezt használjuk fel a függőség vizsgálatakor.

Legyen adott a $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ vektorrendszer. Tetszőleges sorrendben kezdjük bevonni a vektorokat a bázisba elemi bázistranszformációval, az egyes lépéseknél a bevont vektorokat a továbbiakban is tartsuk bent. Az eljárást addig folytassuk, amíg lehet. A befejezés után több végállás lehetséges.

Ha valamennyi vektort be lehetett vinni a bázisba, akkor a báziselemek lineáris függetlensége miatt ennek egy részrendszere, a $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ vektorrendszer is lineárisan független.

Ha több hely nincs, a régi bázis valamennyi vektorát már felváltottuk, és további vektorok várakoznak még a bevonásra, akkor a bevitt vektorok lineárisan függetlenek, a továbbiak függősége pedig leolvasható a táblázatból. Az alábbi numerikus példán nézzük ezt meg.

	$\boldsymbol{v}_3$	$\boldsymbol{v}_5$
$\boldsymbol{v}_2$	1	5
$\boldsymbol{v}_4$	4	-2
$\boldsymbol{v}_6$	0	2
$\boldsymbol{v}_1$	-2	1

A táblázat feltünteti $\boldsymbol{v}_3$ koordinátát a $\boldsymbol{v}_2$, $\boldsymbol{v}_4$, $\boldsymbol{v}_6$, $\boldsymbol{v}_1$ bázisra vonatkozóan, ez azt jelenti, hogy $\boldsymbol{v}_3 = \boldsymbol{v}_2 + 4\boldsymbol{v}_4 - 2\boldsymbol{v}_1$, vagyis a vektorok között ez a lineáris összefüggés áll fenn. Hasonlóan $\boldsymbol{v}_5 = 5\boldsymbol{v}_2 - 2\boldsymbol{v}_4 + 2\boldsymbol{v}_6 + \boldsymbol{v}_1$. Az eredmény természetes, hiszen 4-dimenziós térben 6 vektor nem lehet lineárisan független.

Ugyancsak nem tudjuk folytatni az eljárást, ha már csak 0 generáló elemet lehet választani (de ez tiltott!). Akkor a végállás kicsit más:

	$\boldsymbol{v}_3$	$\boldsymbol{v}_5$
$\boldsymbol{v}_2$	1	5
$\boldsymbol{e}_2$	0	0
$\boldsymbol{v}_4$	4	-2
$\boldsymbol{e}_4$	0	0
$\boldsymbol{v}_6$	0	2
$\boldsymbol{v}_1$	-2	1

Ebből a táblázatból is leolvasható, hogy $\boldsymbol{v}_3 = \boldsymbol{v}_2 + 4\boldsymbol{v}_4 - 2\boldsymbol{v}_1$ és $\boldsymbol{v}_5 = 5\boldsymbol{v}_2 - 2\boldsymbol{v}_4 + 2\boldsymbol{v}_6 + \boldsymbol{v}_1$, tehát a bázisba bevont vektorok lineárisan függetlenek, de több lineárisan független nincs köztük.

A módszer a rangszámításnál valamivel nehézkesebb, de annyival nyújt többet, hogy megadja a vektorok közötti lineáris függőséget is.

II. *Lineáris egyenletrendszerek megoldása*

Írjuk $\boldsymbol{A}$-t $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m)$ alakba, akkor az $\boldsymbol{Ax} = \boldsymbol{b}$ egyenletrendszer a $\sum\limits_{k=1}^{m} \boldsymbol{a}_k x_k = \boldsymbol{b}$ alakot ölti. Innen látható, hogy az egyenletrendszer megoldása az $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m$ és a $\boldsymbol{b}$ vektorok közötti lineáris összefüggés felderítését jelenti. Ezért pontosan ugyanazt csináljuk, mint a lineáris függőség vizsgálatánál: az $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m$ vektorokat megkíséreljük bevonni a bázisba, a $\boldsymbol{b}$ vektor változását csak nyomon követjük. Az elemi bázistranszformációk egymás utáni elvégzését addig folytatjuk, amíg csak lehet (a bevont együttható-vektorokat már nem cseréljük le). A végállás kiértékelése azonban kissé eltér az előző esettől.

Ha az összes $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m$ vektort sikerült bevonni a bázisba, és utolsó utáni lépésként a $\boldsymbol{b}$ vektort is be tudnánk vinni, akkor az $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m, \boldsymbol{b}$ vektorrendszer lineárisan független lenne, ami azt jelenti, hogy nincs megoldása az egyenletrendszernek. Ha a $\boldsymbol{b}$-t már nem tudjuk bevonni, akkor a lineáris függőség leolvasható:

	$\boldsymbol{b}$
$\boldsymbol{a}_2$	2
$\boldsymbol{a}_4$	-5
$\boldsymbol{a}_3$	1
$\boldsymbol{e}_4$	0
$\boldsymbol{a}_1$	3

$\boldsymbol{b} = 3\boldsymbol{a}_1 + 2\boldsymbol{a}_2 + \boldsymbol{a}_3 - 5\boldsymbol{a}_4$, ami azt jelenti, hogy $x_1 = 3$, $x_2 = 2$, $x_3 = 1$, $x_4 = -5$. Ez egyértelmű megoldást jelent. (Lényeges, hogy az eredeti bázisvektorok sorában 0 álljon, ezért nem lehetne a $\boldsymbol{b}$-t a bázisba bevonni, ezért lineárisan összefüggő!)

Ha nem sikerül az összes $\boldsymbol{a}_i$ vektort bevonni a bázisba, akkor továbbra is az a döntő, hogy utolsó utáni lépésként a $\boldsymbol{b}$ bevonható lenne-e. Mindenesetre a kívül maradt vektorokra a lineáris összefüggés felírható (ld. a példát). Ha $\boldsymbol{b}$ bevihető, akkor a bevont vektoroktól lineárisan független, velük $\boldsymbol{b}$ nem fejezhető ki, és ezen a kívül maradt vektorok sem segíthetnek, vagyis az egyenletrendszer nem oldható meg.

Ha a végállásban a bázisban maradt összes régi vektor sora nullákból áll, akkor az egyenletrendszer egy lehetséges megoldását ugyanúgy leolvashatjuk, mint az előző példában, a kívül maradt vektorokhoz tartozó ismeretleneknek 0 értéket választva.

	$\boldsymbol{a}_3$	$\boldsymbol{a}_5$	$\boldsymbol{b}$
$\boldsymbol{a}_2$	1	5	7
$\boldsymbol{e}_2$	0	0	0
$\boldsymbol{a}_4$	4	-2	-3
$\boldsymbol{a}_1$	-2	0	4

A kapott megoldás tehát $x_1 = 4$, $x_2 = 7$, $x_3 = 0$, $x_4 = -3$, $x_5 = 0$.

Vegyük fel azonban tetszőlegesen az x_3 és x_5 értékét, és képezzük a $\boldsymbol{b} - \boldsymbol{a}_3x_3 - \boldsymbol{a}_5x_5$ vektort:

	$\boldsymbol{a}_3$	$\boldsymbol{a}_5$	$\boldsymbol{b} - \boldsymbol{a}_3x_3 - \boldsymbol{a}_5x_5$
$\boldsymbol{a}_2$	1	5	$7 - x_3 - 5x_5$
$\boldsymbol{e}_2$	0	0	0
$\boldsymbol{a}_4$	4	-2	$-3 - 4x_3 + 2x_5$
$\boldsymbol{a}_1$	-2	0	$4 + 2x_3$

Az utolsó oszlop értékei így megoldását adják az $\boldsymbol{a}_2x_2 + \boldsymbol{a}_4x_4 + \boldsymbol{a}_1x_1 =$

$= \boldsymbol{b} - \boldsymbol{a}_3 x_3 - \boldsymbol{a}_5 x_5$ egyenletrendszernek, ami azonos az eredeti egyenletrendszerrel. Az egyenletrendszer általános megoldása tehát:

$$\begin{aligned} x_1 &= 4 + 2x_3, \\ x_2 &= 7 - x_3 - 5x_5, \\ x_4 &= -3 - 4x_3 + 2x_5, \end{aligned}$$

x_3 és x_5 tetszőleges.

III. *Mátrix inverzének meghatározása*

Legyen $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ teljesrangú négyzetes mátrix. Ennek az a tulajdonsága, hogy az $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ bázist az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ bázisba transzformálja:

$$(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) \overset{\boldsymbol{A}}{\mapsto} (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n).$$

Az $\boldsymbol{A}^{-1}$ mátrix a fordítottját cselekszi:

$$(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n) \overset{\boldsymbol{A}^{-1}}{\mapsto} (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n),$$

tehát azt a mátrixot keressük, amely elvégzi ezt a feladatot.

Készítsük el a bázistranszformációs táblázatot:

	$\boldsymbol{a}_1$ $\ldots$ $\boldsymbol{a}_n$
$\boldsymbol{e}_1$	
$\vdots$	$\boldsymbol{A}$
$\boldsymbol{e}_n$	

Elemi bázistranszformációk sorozatával teljesen cseréljük ki a bázist, de az $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ vektorokat őrizzük meg. Sorok cseréjével (ha kell az oszlopok cseréjével is) hozzuk létre a vektorok természetes sorrendjét, akkor a

	$\boldsymbol{e}_1$ $\ldots$ $\boldsymbol{e}_n$
$\boldsymbol{a}_1$	
$\vdots$	$\boldsymbol{A}^{-1}$
$\boldsymbol{a}_n$	

táblázatot kapjuk, ami tartalmazza az inverz mátrixot.

4.4. Ortogonális bázistranszformációk

Fontos szerepet játszanak a bázistranszformációk között azok, amelyek a derékszögű koordináta-rendszert és a koordináta egységvektorok hosszát megtartják. Várhatóan ezek és csak ezek rendelkeznek távolságtartó (normatartó) tulajdonsággal.

Ha $\boldsymbol{Ax}$ az $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ koordináta egységvektorokat az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ bázisba transzformálja, és az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ vektorokról feltételezzük szintén, hogy egységvektorok, valamint páronként egymásra ortogonálisak, akkor ún. *ortogonális transzformációt* végzünk. Az ortogonális transzformáció mátrixát ortogonális mátrixnak nevezzük.

4.4.1. Definíció. Az $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ négyzetes mátrix *ortogonális*, ha $||\boldsymbol{a}_i|| = 1$ minden i-re, és $\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = 0$ minden $i \neq j$-re.

Bár a nevében nincs benne, de ortogonális mátrixnál mindig a normáltságot, vagyis az oszlopvektorok egyenlő hosszát is feltételezzük. Ha az oszlopvektorok hossza egyenlő, akkor már lényegében nem jelent lényeges megszorítást, ha egységnyinek vesszük.

4.4.2. Tétel. Az $\boldsymbol{A}$ akkor és csak akkor ortogonális mátrix, ha $\boldsymbol{A}^{-1} = \boldsymbol{A}^*$.

Bizonyítás. Képezzük az $\boldsymbol{A}^*\boldsymbol{A}$ szorzatot. $\boldsymbol{A}^*$ sorai voltaképpen $\boldsymbol{A}$ oszlopai, ezért az $\boldsymbol{A}^*\boldsymbol{A}$ elemeit $\boldsymbol{A}$ két oszlopának a skalárszorzata adja meg, ez 1, ha a két oszlop azonos és 0, ha különböző, vagyis $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I}$. A 2.7.1. Tétel értelmében tehát $\boldsymbol{A}^* = \boldsymbol{A}^{-1}$. A bizonyítás fordított sorrendben is elmondható.□

4.4.3. Következmény. Ha $\boldsymbol{A}$ ortogonális mátrix, akkor $\boldsymbol{A}^*$ is az.

Bizonyítás. $\boldsymbol{A}^{-1}$ szintén derékszögű, egységvektorokból álló bázist hoz létre, tehát mátrixa ortogonális. □

$\boldsymbol{A}^*$ ortogonalitása azt jelenti, hogy ha a mátrix oszlopai ortogonális egységvektorok, akkor ugyanez érvényes a soraira is. Ha az oszlopvektorok egyenlő hosszát nem tételezzük fel, csak az ortogonalitásukat, akkor ez az állítás nem lesz igaz.

4.4.4. Tétel. Az $\boldsymbol{Ax}$ lineáris transzformáció akkor és csak akkor távolságtartó, ha $\boldsymbol{A}$ ortogonális. A távolságtartó leképezés egyben szögtartó és skalárszorzat-tartó is.

Bizonyítás. $\Leftarrow$ Ha $\boldsymbol{A}$ ortogonális, akkor

$$||\boldsymbol{Ax}||^2 = ||\sum_i \boldsymbol{a}_i x_i||^2 = \sum_i ||\boldsymbol{a}_i x_i||^2 = \sum_i x_i^2 ||\boldsymbol{a}_i||^2 = \sum_i x_i^2 = ||\boldsymbol{x}||^2.$$

$\Rightarrow$ Ha $\boldsymbol{Ax}$ távolságtartó, akkor skalárszorzat-tartó is, hiszen

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \frac{1}{2}(||\boldsymbol{a} + \boldsymbol{b}||^2 - ||\boldsymbol{a}||^2 - ||\boldsymbol{b}||^2).$$

Ha skalárszorzat-tartó is, akkor szögtartó is, hiszen $\cos\alpha = \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{||\boldsymbol{a}|| \cdot ||\boldsymbol{b}||}$. Ha szögtartó, akkor az egységvektorok hosszán kívül azok szögét is megtartja, azaz $\boldsymbol{A}$ ortogonális. □

Ortogonális mátrixok szorzata is ortogonális, hiszen a transzformációk elvégzésével a normatartás öröklődik.

1. **Kérdés.** Hogy lehet könnyen meghatározni az

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

mátrix inverzét?

A legegyszerűbb ortogonális transzformáció a permutáló transzformáció, mely a koordináta egységvektorokat koordináta egységvektorokba képezi le, de más sorrendben. A mátrixának oszlopvektorai különböző sorrendben a koordináta egységvektorok.

4.4.5. Definíció. *Permutáló mátrixnak* nevezzük egy négyzetes mátrixot, ha minden sorában és minden oszlopában pontosan egy elem 1-es, a többi nulla. A permutáló mátrix által létrehozott transzformáció a *permutáló transzformáció.*

Ha egy permutáló mátrixszal balról megszorzunk egy mátrixot, akkor a sorai, ha jobbról, akkor az oszlopai felcserélődnek. Nézzünk egy példát:

$$\boldsymbol{P}\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 11 & 21 & 31 \\ 2 & 12 & 22 & 32 \\ 3 & 13 & 23 & 33 \\ 4 & 14 & 24 & 34 \end{pmatrix} = \begin{pmatrix} 2 & 12 & 22 & 32 \\ 4 & 14 & 24 & 34 \\ 1 & 11 & 21 & 31 \\ 3 & 13 & 23 & 33 \end{pmatrix}$$

Általánosan megfogalmazva, ha $a_{1i_1}, a_{2i_2}, ..., a_{ni_n}$ jelöli a $\boldsymbol{P}$ permutáló mátrix nem nulla elemeit, és az $\boldsymbol{A}$ mátrix sorvektorai rendre $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$, akkor a $\boldsymbol{PA}$ mátrix sorvektorai rendre $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, ..., \boldsymbol{a}_{i_n}$. Röviden $\boldsymbol{P}$ nem nulla elemeinek oszlopindexei mutatják a $\boldsymbol{PA}$ mátrixban a sorok sorrendjét. Jobbról történő szorzásnál hasonló szabály érvényes a sor és az oszlop szóhasználat felcserélésével.

Az $i_1, i_2, ..., i_n$ számsorozat az $1, 2, \ldots, n$ számok egy permutációja. Definiáljuk a permutációban az inverziók számát.

4.4.6. Definíció. Az 1, 2, …, n számok $i_1, i_2, ..., i_n$ permutációjában az *inverziók száma* azon $(i,\ j)$ $(1 \leq i < j \leq n)$ számpárok száma, ahol a permutációban a j szerepel előbb (sorrendjük eltérő a természetes sorrendtől).

Az inverziók száma a permutáció egy fontos adata. Jelentőségét közelebb hozza az alábbi tétel.

4.4.7. Tétel. Ha az 1, 2, …, n számok valamely permutációjában az inverziók száma k, akkor az adott permutáció szomszédos elemek cseréjével történő előállításához legalább k lépés szükséges, és k lépés mindig elégséges is.

Bizonyítás. Szomszédos elemek cseréjénél az inverziók száma mindig eggyel változik, tehát legalább k lépés mindig kell.

Az elégségesség bizonyításához vegyük észre, hogy egy permutációban, ha van inverzió, akkor mindig van szomszédos inverzió is. Két inverzióban álló elem között, ugyanis, mindig van szomszédos inverzió is, mert, ha nem lenne, akkor a köztes elemek monoton növekedőek lennének, de akkor a két elem nem alkotna inverziót. Az inverzióban álló szomszédos párok cseréjével az inverziók száma mindig eggyel csökkenthető, így a permutáció helyreállítható. A permutáció előállítása ennek a sorrendnek a megfordításával érhető el. $\square$

Az a szám természetesen nem adható meg egyértelműen, hogy egy adott permutáció hány lépésben állítható elő szomszédos elemek cseréjével, hiszen „ügyetlenkedni” korlátlan lehetőségünk van. Mivel minden cserénél az inverziószám eggyel változik, páratlan lépésszámmal mindig páratlan, és páros lépésszámmal mindig páros inverziószámot érhetünk el. Az adott permutációnak ezért lényeges adata az inverziószám paritása, és ennek alapján beszélhetünk *páros, ill. páratlan permutációról.*

4.5. Tükrözések

A vetítéshez hasonló, ahhoz köthető transzformációtípus. Itt is, bár a neve alapján kevésbé gondolunk rá, megengedett a nem ortogonális tükrözés is. A tükrözés lényege, hogy megismételve a transzformációt visszaáll az eredeti állapot.

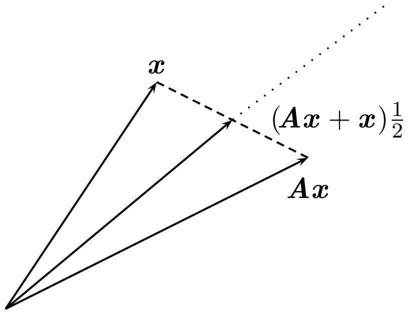
4.5.1. Definíció. Az identikustól különböző $\boldsymbol{Ax}$ lineáris transzformációt *tükrözésnek* nevezzük, ha $\boldsymbol{A}^2 = \boldsymbol{I}$.

$\boldsymbol{A}^2 = \boldsymbol{I}$ tuladonsággal rendelkező, tükröző mátrixot involutórius mátrixnak is nevezik.

Az $\boldsymbol{A}^2 = \boldsymbol{I}$ összefüggés nyilván ugyanazt jelenti, mint $\boldsymbol{A} = \boldsymbol{A}^{-1}$. Az

$\boldsymbol{A}^2 = \boldsymbol{I}$-ből következően az inverz a 2.7.1. Tétel alapján biztosan létezik.

A vetítéssel való kapcsolatát a $\boldsymbol{B} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{I})$ jelenti, ahol $\boldsymbol{A}$ a tükrözés és $\boldsymbol{B}$ a vetítés. Valóban $\boldsymbol{BB} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{I})\frac{1}{2}(\boldsymbol{A} + \boldsymbol{I}) = \frac{1}{4}(\boldsymbol{A}^2 + 2\boldsymbol{A} + \boldsymbol{I}) = = \frac{1}{2}(\boldsymbol{A}+\boldsymbol{I}) = \boldsymbol{B}$. Az összefüggés természetesen geometriailag szemléltethető:



Az altér, amire a tükrözés történik, a „tengely”, melynek pontjai helyben maradnak az $\{\boldsymbol{x} + \boldsymbol{Ax}\colon \boldsymbol{x} \in \mathbf{R}^n\}$ halmaz, vagyis az $(\boldsymbol{A} + \boldsymbol{I})\boldsymbol{x}$ transzformáció képtere. Valóban $\boldsymbol{A}(\boldsymbol{x} + \boldsymbol{Ax}) = \boldsymbol{Ax} + \boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{Ax} + \boldsymbol{x}$. A vetítő síkok az $L = \{\boldsymbol{x}\colon \boldsymbol{x} + \boldsymbol{Ax} = \boldsymbol{0}\}$ altérrel párhuzamos síkok. A tengelyként szolgáló altér és a vetítősík mindig egy pontban metszi egymást, ugyanis a tengely és az L kiegészítő alterek (ld. 3.4.2 Tétel).

5. Kérdés. Tükrözés-e a $\boldsymbol{Ax}$ transzformáció, ha

$$\boldsymbol{A} = \begin{pmatrix} 3 & -8 & 4 \\ 2 & -7 & 4 \\ 2 & -8 & 5 \end{pmatrix}?$$

Melyik altérre tükröz?

4.5.2. Definíció. A tükrözés *ortogonális*, ha a tengelyként szolgáló altere és az $L = \{\boldsymbol{x}\colon (\boldsymbol{A} + \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}\}$ ortogonális kiegészítő alterek.

A vetítés és a tükrözés közötti kapcsolatból közvetlenül következik az alábbi tétel.

4.5.3. Tétel. Az $\boldsymbol{A}$ tükrözés akkor és csak akkor ortogonális, ha $\boldsymbol{A}$ szimmetrikus, azaz ha $\boldsymbol{A} = \boldsymbol{A}^*$.

Az ortogonális tükrözés – a szemléletünk szerint – normatartó (ellentétben a nem ortogonális tükrözéssel), de ezt bizonyítani kell. Ennél kissé többet is állíthatunk.

4.5.4. Tétel. Az $\boldsymbol{A}$, $\boldsymbol{A} \neq \boldsymbol{I}$ mátrix által létrehozott lineáris transzformáció akkor és csak akkor ortogonális tükrözés, ha $\boldsymbol{A}$ ortogonális és szimmetrikus.

Bizonyítás. $\Rightarrow$ Ha $\boldsymbol{A}$ tükrözés, akkor $\boldsymbol{A} = 2\boldsymbol{B} - \boldsymbol{I}$, ahol $\boldsymbol{B}$ vetítés.

Alkalmazzuk a Pythagorasz-tételt (3.4.5. Tétel), akkor $||\boldsymbol{x}||^2 = ||\boldsymbol{B}\mathbf{x}||^2 + ||\boldsymbol{x} - \boldsymbol{B}\mathbf{x}||^2$. $\boldsymbol{Ax}$-re is alkalmazzuk a Pythagorasz-tételt, felhasználva, hogy $\boldsymbol{BAx} = \boldsymbol{B}(2\boldsymbol{B} - \boldsymbol{I})\boldsymbol{x} = \boldsymbol{Bx}$, és $\boldsymbol{Ax} - \boldsymbol{BAx} = (2\boldsymbol{B} - \boldsymbol{I})\boldsymbol{x} - \boldsymbol{Bx} = \boldsymbol{Bx} - \boldsymbol{x}$, $||\boldsymbol{Ax}||^2 = ||\boldsymbol{BAx}||^2 + ||\boldsymbol{Ax} - \boldsymbol{BAx}||^2 = ||\boldsymbol{Bx}||^2 + ||\boldsymbol{Bx} - \boldsymbol{x}||^2 = ||\boldsymbol{x}||^2$, ami a normatartást jelenti. $\boldsymbol{A}$ tehát ortogonális, és mivel $\boldsymbol{B}$ ortogonális vetítés, szimmetrikus, így $\boldsymbol{A}$ is az.

$\Leftarrow$ $\boldsymbol{A}$ ortogonális, és szimmetrikus voltából következik, hogy $\boldsymbol{A}^2 = \boldsymbol{AA}^* = \boldsymbol{AA}^{-1} = \boldsymbol{I}$. Továbbá $\boldsymbol{A}$ szimmetrikus, tehát ortogonális tükrözést hoz létre. $\square$

6. Kérdés. A 2. fejezet 8. kérdésében megmutattuk, hogy a

$$\boldsymbol{B} = \frac{1}{21}\begin{pmatrix} 11 & -1 & -3 & 10 \\ -1 & 2 & 6 & 1 \\ -3 & 6 & 18 & 3 \\ 10 & 1 & 3 & 11 \end{pmatrix}$$

mátrixú lineáris transzformáció ortogonális vetítés. Adjuk meg a hozzátartozó ortogonális tükrözés mátrixát!

Minden vetítés, és így minden tükrözés is kanonikus (diagonális) alakra hozható. Válasszunk egy-egy bázist a vetítés képterében és a magterében, legyen ez $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r$, ill. $\boldsymbol{a}_{r+1}, \ldots, \boldsymbol{a}_n$. Az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ bázisra vonatkozóan a vetítés minden $\boldsymbol{x}$ vektorra vonatkozóan az első r koordinátát megőrzi, a továbbiakat lenullázza. A transzformáció mátrixa ebből leolvasható:

$$\boldsymbol{B}_0 = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Minden vetítő mátrix tehát $\boldsymbol{B}_0$-hoz hasonló mátrix, ennek más koordinátarendszerben felírt alakja. Ebből megalkotható a tükröző mátrix kanonikus alakja is, hiszen $\boldsymbol{A}_0 = 2\boldsymbol{B}_0 - \boldsymbol{I}$:

$$\boldsymbol{A}_0 = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \text{-}\boldsymbol{I} \end{pmatrix}.$$

Minden tükrözés mátrixa tehát $\boldsymbol{A}_0$-hoz hasonló. (A hasonlóság fogalmát ld. 4.1.5. Definíció.)

7. Kérdés. A 3. fejezet 7. Kérdésében szereplő

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 1 \\ -4 & -3 & -4 \\ 2 & 2 & 3 \end{pmatrix}$$

mátrixszal megadott vetítés milyen $\boldsymbol{C}$ mátrixszal transzformálható kanonikus alakra. Ellenőrizzük az eredményünket!

4.6. Forgatások

Tréfásnak tűnik a kérdés: Lehet-e kétdimenziós sík körül forgatni, mondjuk négy dimenzióban? A forgatások tanulmányozásához először az elérhető dimenziókban nézzük meg a kérdést.

Két dimenzióban pont körüli forgatás létezik. Három dimenzióban egyenes körül lehet forgatni, a pont körül „gömbcsuklós forgatás” létezik, aminek két szabadsági foka van, ami két paraméterrel írható le. Most csak az egy paraméterrel jellemezhető forgatásokkal fogunk foglalkozni.

Mi a jellemzője az egy szabadságfokú forgatásnak? Távolságtartó transzformáció, ami a „tengely” pontjait helyben hagyja, és minden erre merőleges vektort adott szöggel elfordít.

$\mathbf{R}^2$-ben az origó körüli φ szöggel történő elforgatás az $(x,\ y)$ pontot, melynek polárkoordinátái $(r,\ \alpha)$, az $(r,\ \alpha+\varphi)$ pontba viszi át, ehhez az

$$(r\cos(\alpha+\varphi), r\sin(\alpha+\varphi)) =$$

$$= (r\cos\alpha\cos\varphi - r\sin\alpha\sin\varphi,\ r\cos\alpha\sin\varphi + r\sin\alpha\cos\varphi) =$$

$$= (x\cos\varphi - y\sin\varphi, x\sin\varphi + y\cos\varphi)$$

derékszögű koordináták tartoznak. Ha az $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ és az

$$\boldsymbol{A}_0 = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

jelöléseket használjuk akkor az $\boldsymbol{x}$ vektor az elforgatás során az $\boldsymbol{A}_0\boldsymbol{x}$-be megy át. Hasonlóan $\mathbf{R}^3$-ban a z-tengely körüli φ szöggel történő elforgatás az $\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ vektort az $\boldsymbol{Ax}$-be viszi át, ahol most

$$\boldsymbol{A} = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ez általánosítható magasabb dimenzióra is. A forgatás történjen egy $(n-2)$-dimenziós altér körül, ennek pontjai helyben maradnak, az erre ortogonális kiegészítő altér vektorai pedig φ szöggel elfordulnak. Vegyük fel a derékszögű koordináta-rendszert úgy, hogy az $n-2$ koordinátatengely a tengelyként szolgáló altérben legyen. A másik két koordinátatengely, mondjuk az első kettő, már ezekre merőleges, tehát kötelezően az ortogonális kiegészítő altérben van. Képezzük az

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_0 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi & \cdots & 0 \\ \sin\varphi & \cos\varphi & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

mátrixot. Az $\boldsymbol{A}$ mátrix ortogonális, tehát $\boldsymbol{Ax}$ normatartó transzformáció. $\boldsymbol{Ax}$ minden olyan vektort, mely a tengelyként szolgáló altérben van, vagyis az első két koordinátája nulla, önmagába visz át. Továbbá az ortogonális kiegészítő altérben lévő $\boldsymbol{x} = (x,\, y,\, 0,\, \ldots,\, 0)$ és $\boldsymbol{Ax}$ α szögére:

$$\cos\alpha = \frac{\langle \boldsymbol{x}, \boldsymbol{A}\mathrm{x}\rangle}{||\boldsymbol{x}|| \cdot ||\boldsymbol{A}\mathrm{x}||} = \frac{x^2\cos\varphi - xy\sin\varphi + xy\sin\varphi + y^2\cos\varphi}{||\boldsymbol{x}||^2} =$$

$$= \frac{x^2+y^2}{||\boldsymbol{x}||^2}\cos\varphi = \cos\varphi,$$

vagyis az elfordulás szöge minden ilyen vektorra φ. (Az elfordulás szögét előjelezhetjük is, bár ez problémás. A szögtartás miatt az elfordulás szöge azonban mindig azonos előjelű.) Ezzel beláttuk, hogy ténylegesen egyetlen szögadattal jellemezhető forgatásról van szó.

A fenti kérdésre a válasz: négydimenziós térben létezik bármely kétdimenziós altér (vagy sík) körüli forgatás. Általában megadható $\mathbf{R}^n$-ben is bármely $(n-2)$-dimenziós altér körüli egyparaméteres forgatás.

4.6.1. Definíció. $\mathbf{R}^n$-ben adott $(n-2)$-dimenziós altér körüli *forgatás*, mely az altér pontjait helybenhagyja, és az ortogonális altér vektorait φ szöggel elforgatja, alkalmas derékszögű koordináta-rendszerben az

$$\boldsymbol{A} = \begin{pmatrix} \cos\varphi & -\sin\varphi & \cdots & 0 \\ \sin\varphi & \cos\varphi & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

mátrixszal adható meg.

Algebrailag is könnyen ellenőrizhető, de a transzformáció természetéből is következik, hogy $\boldsymbol{A}$ inverze a -φ szöggel történő elforgatás.

8. Kérdés. Az L alteret adjuk meg két generáló elemével, legyenek ezek $\boldsymbol{a} = (1,\, 1,\, 0,\, 0)$ és $\boldsymbol{b} = (0,\, 0,\, 1,\, 1)$. Forgassuk el a $\boldsymbol{v} = (2,\, 1,\, 1,\, 2)$ vektort L körül 45^o-kal! $\boldsymbol{v}$ és elforgatása tényleg 45^o-os szöget fog bezárni?

4.6.2. Tétel. Bármely ortogonális transzformáció elvégezhető legfeljebb $n - 1$ forgatás és esetleg egy ortogonális tükrözés egymás utáni végrehajtásával.

Bizonyítás. Az ortogonális transzformáció az $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\ldots$, $\boldsymbol{e}_n$ vektorokat vigye át az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ vektorokba, ahol az utóbbi vektorok is egységvektorok és páronként ortogonálisak.

Vegyünk fel egy kétdimenziós L alteret, mely tartalmazza az $\boldsymbol{e}_1$ és az $\boldsymbol{a}_1$ vektorokat, az első forgatást L ortogonális kiegészítő altere körül akarjuk elvégezni. Legyen $\boldsymbol{e}_1$ és $\boldsymbol{a}_1$ szöge φ, és forgassunk $L^\perp$ körül φ vagy -φ szöggel, akkor egyik esetben $\boldsymbol{e}_1$ az $\boldsymbol{a}_1$-re képződik le.

Tegyük fel most, hogy $\boldsymbol{e}_i = \boldsymbol{a}_i$ ($i = 1, 2, \ldots, k - 1$; $k < n$). Ha $\boldsymbol{e}_k$ és $\boldsymbol{a}_k$ lineárisan függetlenek, akkor legyen L az általuk generált kétdimenziós altér, ha nem, akkor $\boldsymbol{e}_k$ és az $\boldsymbol{e}_{k+1}$ által generált altér. Mindkét esetben az $\boldsymbol{e}_i$ vektorok az $L^\perp$ altér elemei ($i = 1, 2, \ldots, k - 1$), tehát az $L^\perp$ körüli forgatás nem mozdítja el ezeket. A forgatás szöge legyen φ vagy -φ, ahol φ az $\boldsymbol{e}_k$ és $\boldsymbol{a}_k$ szöge. Valamelyik forgatás az $\boldsymbol{e}_k$-t az $\boldsymbol{a}_k$-ba képezi le.

Folytatva az eljárást $n - 1$ forgatással elérhető, hogy $\boldsymbol{e}_1 = \boldsymbol{a}_1$, $\boldsymbol{e}_2 = \boldsymbol{a}_2$, $\ldots$, $\boldsymbol{e}_{n-1} = \boldsymbol{a}_{n-1.}$ Az $\boldsymbol{e}_n$ és az $\boldsymbol{a}_n$ vektor egyaránt ortogonális az $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\ldots$, $\boldsymbol{e}_{n-1}$ vektorokra, tehát hozzátartoznak az általuk generált altér egydimenziós ortogonális kiegészítő alteréhez. Ez azt jelenti, hogy $\boldsymbol{e}_n = \boldsymbol{a}_n$ vagy $\boldsymbol{e}_n =$ -$\boldsymbol{a}_n$. Az utóbbi esetben egy ortogonális tükrözésre van szükség, mely $\boldsymbol{e}_n$-et -$\boldsymbol{a}_n$-be visz át $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\ldots$, $\boldsymbol{e}_{n-1}$ helybenhagyása mellett. □

9. Kérdés. $\mathbf{R}^4$-ben forgatás-e az a lineáris transzformáció, amelynek mátrixa

$$\begin{pmatrix} \cos\varphi & \sin\varphi & 0 & 0 \\ \sin\varphi & -\cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ vagy } \begin{pmatrix} \cos\varphi & \sin\varphi & 0 & 0 \\ -\sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}?$$

4.7. Válaszok a kérdésekre

1. Az 1. fejezet 11. Kérdése során láttuk, hogy az origóba helyezett egységkocka origóból kiinduló $(n - 1)$-dimenziós lapátló vektorai $\mathbf{R}^n$ bázisát alkotják. Adjuk meg a testátló vektor koordinátáit ebben a bázisban!

A testátló vektor koordinátái $\mathbf{1} = (1, 1, \ldots, 1)$, a lapátló vektorok koordinátái az 1. fejezet 11. kérdésére adott válaszban megadott $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ vektorok. Mivel $\boldsymbol{a}_1 + \boldsymbol{a}_2 + \ldots + \boldsymbol{a}_n = (n - 1)\mathbf{1}$**,** a testátló vektor koordinátái az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ bázisban: $\left(\frac{1}{n-1}, \frac{1}{n-1}, \ldots, \frac{1}{n-1}\right)$.

2. Hasonlók-e az

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix} \text{ és a } B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

mátrixok?

A második transzformációnál van olyan nem nulla vektor a térben, melyre $\boldsymbol{Bx} = \boldsymbol{x}$, nevezetesen az $\boldsymbol{x} = (a, 0, 0)$. Ha hasonlók, akkor más koordinátarendszerben ugyanazt a transzformációt létesíti, tehát $\boldsymbol{A}$ esetében is kell lennie olyan nullától különböző $\boldsymbol{x}$ vektornak, melyre $\boldsymbol{Ax} = \boldsymbol{x}$. Ez utóbbi az $\boldsymbol{x} = (x, y, z)$ jelölés mellett az

$$\begin{aligned} x + 3y + z &= x \\ 2x + 2y &= y \\ x - y + 2z &= z \end{aligned}$$

egyenletrendszer teljesülését kívánja meg. Az egyenletrendszer rangja azonban 3, ezért csak az $\boldsymbol{x} = \mathbf{0}$ megoldása van. A mátrixok tehát nem hasonlók. (Lásd még a sajátértékekről szóló 6.1. fejezetet.)

3. Egy elemi bázistranszformációt alkalmazva az $\boldsymbol{x}$ vektor (3, 0, 2, 1) koordinátái (1, 2, -1, 0)-ra változtak. Megállapítható-e, hogy mely bázisvektor maradt biztosan a régi? Megállapítható-e, hogy melyiket cseréltük le? Mik a bázisba bevont vektor eredeti koordinátái?

A második és a negyedik bázisvektor biztosan változatlan maradt, mert a generáló elem sorában vagy mindkét helyen nulla áll, vagy egyik sem nulla. Az első és a harmadik bázisvektor bármelyikét lecserélhettük, mégpedig az első esetén a bevont vektor $(3, -2, 3, 1)$, a harmadik esetén $(-2, 2, -2, -1)$.

4. Hogy lehet könnyen meghatározni az

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

mátrix inverzét?

Az $\frac{1}{3}\boldsymbol{A}$ mátrix ortogonális mátrix, inverze a transzponáltja. $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$, tehát

$$\boldsymbol{A}^{-1} = \frac{1}{3}(\frac{1}{3}\boldsymbol{A})^{-1} = \frac{1}{3}(\frac{1}{3}\boldsymbol{A})^{*} = \frac{1}{9}\boldsymbol{A}^{*}.$$

5. Tükrözés-e az $\boldsymbol{A}$ transzformáció, ha

$$\boldsymbol{A} = \begin{pmatrix} 3 & -8 & 4 \\ 2 & -7 & 4 \\ 2 & -8 & 5 \end{pmatrix}?$$

Melyik altérre tükröz?

Könnyű ellenőrizni, hogy $\boldsymbol{A}^2 = \boldsymbol{I}$, tehát $\boldsymbol{A}$ tükrözés. Tudjuk, hogy

$$\boldsymbol{B} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{I}) = \begin{pmatrix} 2 & -4 & 2 \\ 1 & -3 & 2 \\ 1 & -4 & 3 \end{pmatrix}$$

vetítés. Mivel $\mathrm{rang}(\boldsymbol{B}) = 2$, és a $\boldsymbol{B}$ képtere megegyezik a vetítés tengelyével, a tengely a $(2, 1, 1)$ és a $(2, 2, 3)$ vektorok által generált kétdimenziós altér.

6. A II. fejezet 8. kérdésében megmutattuk, hogy a

$$\boldsymbol{B} = \frac{1}{21}\begin{pmatrix} 11 & -1 & -3 & 10 \\ -1 & 2 & 6 & 1 \\ -3 & 6 & 18 & 3 \\ 10 & 1 & 3 & 11 \end{pmatrix}$$

mátrixú lineáris transzformáció ortogonális vetítés. Adjuk meg a hozzátartozó ortogonális tükrözés mátrixát!

Ha $\boldsymbol{A}$-val jelöljük a hozzá tartozó tükrözés mátrixát, akkor

$$\boldsymbol{A} = 2\boldsymbol{B} - \boldsymbol{I} = \frac{1}{21}\begin{pmatrix} 1 & -2 & -6 & 20 \\ -2 & -17 & 12 & 2 \\ -6 & 12 & 15 & 6 \\ 20 & 2 & 6 & 1 \end{pmatrix}.$$

Ellenőrizhetően $\boldsymbol{A}$ szimmetrikus, ortogonális mátrix (és ezért $\boldsymbol{A}^2 = \boldsymbol{I}$).

7. A 3. fejezet 7. Kérdésében szereplő

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 1 \\ -4 & -3 & -4 \\ 2 & 2 & 3 \end{pmatrix}$$

mátrixszal megadott vetítés milyen $\boldsymbol{C}$ mátrixszal transzformálható kanonikus alakra. Ellenőrizzük az eredményünket!

$\text{Rang}(\boldsymbol{A}) = 2$, R_A tehát kétdimenziós, két lineárisan független generáló eleme: $(2, -4, 2)$ és $(1, -3, 2)$. Az első vektor helyett vehetünk $(1, -2, 1)$-t. Az M_A egydimenziós, generáló eleme a

$$\begin{aligned} 2x + y + z &= 0 \\ -4x - 3y - 4z &= 0 \\ 2x + 2y + 3z &= 0 \end{aligned}$$

egyenletrendszer egyik megoldásaként adódik. Az általános megoldás pl. a $(\frac{z}{2},\ -2z,\ z)$ vektor, ahol z tetszőleges szám. Legyen mondjuk $z = 2$, akkor az $(1, -4, 2)$ vektort kapjuk. Alkalmazni kell a $\boldsymbol{C}$-vel képzett bázistranszformációt, ahol $\boldsymbol{C}$ ezekből a vektorokból épül fel:

$$\boldsymbol{C} = \begin{pmatrix} 1 & 1 & 1 \\ -2 & -3 & -4 \\ 1 & 2 & 2 \end{pmatrix}.$$

A kanonikus alak előállításához $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$-t kell kiszámolni, ehhez $\boldsymbol{C}^{-1}$-re is szükség van, ami például a 4.3. pont módszerével elérhető.

8. Az L alteret adjuk meg két generáló elemével: $\boldsymbol{a} = (1, 1, 0, 0)$-val és $\boldsymbol{b} = (0, 0, 1, 1)$-vel. Forgassuk el a $\boldsymbol{v} = (2, 1, 1, 2)$ vektort L körül 45^o-kal! $\boldsymbol{v}$ és elforgatása tényleg 45^o-os szöget fog bezárni?

Ahhoz, hogy a forgatás kanonikus alakját használni tudjuk, koordinátarendszer transzformációra van szükség. Vegyünk fel egy koordináta-rendszert a megadott és az arra ortogonális altérben. Az 1.4.4. tétel módszerével (ortogonalizálás) ez megtehető, de itt ez felesleges, mert a megadott vektorok már ortogonálisak, csak normálni kell ezeket, a további, ezekre ortogonális irányok pedig könnyen kitalálhatók. Így az új koordinátavektorok:

$$\boldsymbol{a}_1 = \frac{1}{\sqrt{2}}(1,\ 1,\ 0,\ 0), \quad \boldsymbol{a}_2 = \frac{1}{\sqrt{2}}(0,\ 0,\ 1,\ 1),$$

$$\boldsymbol{a}_3 = \frac{1}{\sqrt{2}}(1,\ -1,\ 0,\ 0), \quad \boldsymbol{a}_4 = \frac{1}{\sqrt{2}}(0,\ 0,\ 1,\ -1).$$

A $\boldsymbol{v}$ vektor új koordinátái (pl. elemi bázistranszformációkkal meghatározható módon) ebben a koordináta-rendszerben: $\frac{1}{\sqrt{2}}(3,\ 3,\ 1,\ -1)$. Alkalmazzuk a forgatás kanonikus alakját:

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3 \\ 3 \\ 1 \\ -1 \end{pmatrix} \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 \\ 3 \\ \sqrt{2} \\ 0 \end{pmatrix},$$

ez az elforgatott vektor koordinátás alakja az új rendszerben. A vissza-transzformálás:

$$\boldsymbol{v}_1 = \frac{3}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{3}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 3+\sqrt{2} \\ 3-\sqrt{2} \\ 3 \\ 3 \end{pmatrix}.$$

$\boldsymbol{v}_1$ nem tartozik hozzá az ortogonális altérhez, tehát nem várható, hogy $\boldsymbol{v}$-vel bezárt szöge 45^o legyen (nem is annyi, kb 15^o).

9. $\mathbf{R}^4$-ben forgatás-e az a lineáris transzformáció, amelynek mátrixa

$$\begin{pmatrix} \cos\varphi & \sin\varphi & 0 & 0 \\ \sin\varphi & -\cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ vagy } \begin{pmatrix} \cos\varphi & \sin\varphi & 0 & 0 \\ -\sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}?$$

A második mátrix $-\varphi$ szöggel történő forgatás mátrixa. Az első nem forgatás, hanem egy forgatás és egy tükrözés összetételéből adódó transzformáció:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\varphi & \sin\varphi & 0 & 0 \\ \sin\varphi & -\cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 & 0 \\ -\sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

5. fejezet

Determinánsok

5.1. A determináns definíciója

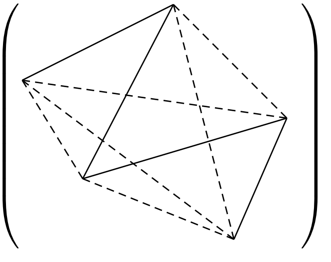
A determináns a négyzetes mátrixokhoz rendelt számérték. A hozzárendelés módja eléggé összetett, számos definiálási mód létezik, itt viszonylag közérthető utat választunk, de látszólag eltérünk a fő irányvonaltól, az n-dimenziós terek vizsgálatától. Fogadjuk el egyelőre, hogy egy fontos segédeszközzel ismerkedünk meg, később meglátjuk, hogy ez az eszköz szorosan kapcsolódik a lineáris (sőt nem csak a lineáris) transzformációkhoz.

Ebben a fejezetben végig feltételezzük, hogy a szereplő mátrixok négyzetes mátrixok. A hozzárendelés módját nehéz, és nem célszerű egy rövid definícióban megadni, hosszasabban fogunk beszélni róla.

Az $n \times n$-es $\boldsymbol{A}$ mátrix esetében *elemi szorzatnak* nevezzük a különböző sorokból és különböző oszlopokból kiválasztott n elem szorzatát. Összesen $n!$ darab elemi szorzatot képezhetünk, ugyanis az elemeknek ennyi megfelelő kiválasztása van. Az első sor elemét n-féleképpen választhatjuk ki, a második sor elemét már csak $(n-1)$-féleképpen, mert az első sorban kiválasztott elem egy oszlopot letilt. A harmadik sorbéli elemet $(n-2)$-féleképpen választhatjuk, mert két oszlop van letiltva, és így tovább, az utolsó sor elemének választása már egyértelmű.

Egy adott elemi szorzat $a_{1k_1}a_{2k_2}\cdots a_{nk_n}$ alakban írható fel, ahol $k_1, k_2, \ldots, k_n$ az 1, 2, …, n számok valamely permutációja. Ebből is látszik, hogy az elemi szorzatok száma a permutációk számával egyezik, azaz $n!$. Egy adott permutációban – és így egy adott elemi szorzatban is – a 4.4.6. Definícióban definiáltuk az inverziók számát. *Inverziónak* nevezzük azt a jelenséget, amikor egy nagyobb szám megelőzi a kisebbet, ezen előfordulások száma az inverziók száma. A 4.4.7. Tétel megadja e szám jelentését is: azon szomszédos sorcserék (vagy oszlopcserék) minimális száma, amellyel az elemi szorzat tényezői a főátlóba vihetők át.

Nézzünk egy példát: 5×5-ös mátrixban vegyük az $a_{13}a_{21}a_{35}a_{42}a_{54}$ elemi szorzatot, ennek az első indexek rendezése után a második indexek 3, 1, 5, 2, 4 permutációja felel meg. Itt inverziót jelent a (3, 1), (3, 2), (5, 2) és az (5, 4) számpár, az inverziók száma tehát négy. Az inverziók száma grafikusan is – talán kisebb hibalehetőséggel – meghatározható a következőképpen. Rajzoljuk fel az „üres” mátrixot, azaz ne tüntessük fel a szereplő elemeket, majd jelöljük be az elemi szorzatba beválasztott elemek helyét. Ezután kössük össze minden bejelölést minden másikkal:



A jobbra-felfelé haladó összeköttetések (az ábrán folytonos vonallal jelölve) mutatják az inverziókat, az inverziók száma a jobbra-felfelé haladó összekötések száma. *Párosnak* nevezzük a permutációt, ha az inverziók száma páros, egyébként *páratlannak* hívjuk.

5.1.1. Definíció. Az $\boldsymbol{A}$ $n \times n$-es mátrix *determinánsának* nevezzük az összes elemi szorzat megfelelő előjellel képezett összegét. Az előjeladás szabálya: az elemi szorzat előjele megmarad, ha az elemi szorzathoz tartozó permutáció páros, megváltozik, ha páratlan. Az $\boldsymbol{A}$ determinánsát $|\boldsymbol{A}|$ jellel jelöljük.

A permutáció páros, ha a fenti ábrában a jobbra-felfelé haladó összekötések száma páros.

2×2-es mátrix determinánsa: $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$, vagyis a főátlóbeli elemek szorzatából az ún. mellékátlóbeli elemek szorzatát le kell vonni.

1. Kérdés. Eddigi ismereteinkkel határozzuk meg a

$$\begin{pmatrix} 0 & 0 & 5 & 0 \\ 2 & 0 & 0 & -1 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 0 & 12 \end{pmatrix}$$

mátrix determinánsát!

5.2. A determinánsok elemi tulajdonságai

A definíció alapján – némi meggondolással – a determinánsokra vonatkozó néhány tulajdonság közvetlenül megállapítható.

5.2.1. Tulajdonság. $|\boldsymbol{A}| = |\boldsymbol{A}^*|$.

Bizonyítás. A transzponálás a főátlóra vonatkozó tükrözés. Adott elemi szorzat elemeit tükrözve a transzponált mátrix elemi szorzatának tényezőit kapjuk meg. A két determináns tehát ugyanazokból az elemi szorzatokból tevődik össze. Egy adott elemi szorzathoz és tükörképéhez ugyanolyan előjel tartozik, mert a permutációk inverzióinak a száma egyenlő: az a_{ij} és az a_{kl} ($i \neq k$ és $j \neq l$) elemek pontosan akkor alkotnak inverziót, ha $k - i$ és $l - j$ előjele különböző, ez a szabály a sorok és oszlopok cseréje után is változatlan marad. □

Az 5.2.1. Tulajdonság alapján a továbbiakban a determináns sorai és oszlopai szimmetrikus szerepet játszanak, minden sorokra vonatkozó tulajdonság oszlopokra is érvényes és fordítva.

5.2.2. Tulajdonság. Ha az $\boldsymbol{A}$ mátrixnak egy sora csupa nullából áll, akkor $|\boldsymbol{A}| = 0$.

Bizonyítás. Minden elemi szorzat tartalmaz egy elemet ebből a csupa nulla sorból, így minden elemi szorzat nulla. □

5.2.3. Tulajdonság. Ha egy determináns egyik sorának minden elemét c-vel megszorozzuk, a determináns értéke c-szeres lesz.

Bizonyítás. Minden elemi szorzat pontosan egy elemet tartalmaz ebből a c-szeres sorból, így minden elemi szorzat c-szeres lesz. □

5.2.4. Tulajdonság. Ha a determináns két sorát felcseréljük, a determináns előjelet vált (abszolút értéke változatlan).

Bizonyítás. A determináns mindkét esetben ugyanazokból az elemi szorzatokból tevődik össze, csupán az előjeleiket kell megvizsgálni. Tegyük fel, hogy az i-edik és a j-edik sort cseréljük fel ($i < j$). Tekintsünk egy elemi szorzatot, és tényezőit rajzoljuk be az üres mátrixba. Azok az összekötések, melyek végpontjai nem változnak, megtartják haladásuk irányát (jobbra-fel vagy jobbra-le). Azok az összekötések is megtartják haladásuk irányát, amelyek az i-edik előtti, vagy a j-edik utáni sorokból indulnak egyik változó pont felé. Az i-edik és a j-edik sor közötti pontokból induló és a változó pontokban végződő összekötések iránya a sorcserével mind biztosan megváltozik, de ilyen összekötés páros sok van, tehát ez az előjel szempontjából közömbös. Marad a két változó pont összekötése, ennek iránya is megváltozik, és ez minden elemi szorzat előjelét megváltoztatja. □

5.2.5. Tulajdonság. A determináns értéke nulla, ha egyik sor a másik c-szerese. Speciálisan a determináns nulla, ha két sora egyenlő.

Bizonyítás. A determinánsból a c konstans kiemelhető (5.2.3. Tulajdonság) és két egyenlő sor áll elő, vagyis elég a második állítást bizonyítani. Cseréljük fel a két egyenlő sort. A mátrix ezzel nem változik, így a determinánsa sem változhat. Másrészt sorcserével a determináns előjelet vált (5.2.4. Tulajdonság), ezért a determináns értéke egyenlő a mínusz egyszeresével, de akkor csak nulla lehet. □

5.2.6. Tulajdonság. Két determináns összeadható, ha valamelyik soruk kivételével a többi megegyezik. A két determináns összege a két különböző sorvektor összeadásával és a további sorvektorok változatlan leírásával kapott mátrix determinánsa.

Bizonyítás. Válasszunk ki az első mátrixból egy tetszőleges elemi szorzatot, és adjuk hozzá az ugyanilyen pozícióban lévő elemi szorzatot a második mátrixból. Az egyenlő elemek kiemelhetők, a különböző sorok elemei pedig összeadódnak, és így pontosan az eredmény mátrix ugyanilyen pozíciójú elemi szorzatát kapjuk. Az előjelek is megegyeznek, mivel a pozíciók azonosak. □

5.2.7. Tulajdonság. A determináns értéke nem változik meg, ha egy sorához a másik sor többszörösét hozzáadjuk.

Bizonyítás. Tegyük fel, hogy az i-edik sorhoz adjuk hozzá a j-edik sor c-szeresét. A műveletet végezzük el úgy, hogy az eredeti determinánshoz adjuk hozzá azt a determinánst, melynek minden sora azonos az eredetivel, csak az i-edik sora megegyezik a j-edik sor c-szeresével. Ezzel a hozzáadással a determináns értéke nem változik meg, hiszen nullát adunk hozzá (5.2.5. Tulajdonság). Az 5.2.6. Tulajdonság miatt a determinánsok összeadhatók, és az i-edik sorhoz a j-edik c-szeresét adjuk hozzá. □

Vessük össze a 2.6.2. Tétellel. A determinánsszámolás és a rangszámítás ugyanazzal a módszerrel történhet: ugyanúgy addig nullázunk, amíg egyetlen nem nulla elemi szorzat marad, akkor ezt kiszámítjuk és előjelezzük a definícióban elmondott előjelszabály szerint. Később (ld. 5.3.) a determináns kiszámításának ez a módja még finomítható.

5.2.8. Tulajdonság. $\mathrm{rang}(\boldsymbol{A}) = n$ akkor és csak akkor, ha $|\boldsymbol{A}| \neq 0$.

Bizonyítás. $\Rightarrow$ Ha $\mathrm{rang}(\boldsymbol{A}) = n$, akkor a rangszámítással a végállásban egyetlen nem nulla elemi szorzat marad, így a determináns értéke sem nulla.

$\Leftarrow$ Ha $\mathrm{rang}(\boldsymbol{A}) < n$, akkor a rangszámításnál a végállásban lesz csupa nullából álló sor, így ennek a determinánsa is nulla. □

5.2.9 Következmény. Ha $|\boldsymbol{A}| \neq 0$, akkor pusztán oszlopműveletekkel (vagy pusztán sorműveletekkel) is elérhető az a végállás, melyben minden nem nulla elem teljesen magányos.

Bizonyítás. Válasszuk ki az első sor egyik nem nulla elemét, legyen ez $a_{1i_1} \neq 0$. Ilyen van, mert $|\boldsymbol{A}| \neq 0$. Ezzel az elemmel az első sor nullázható, tehát elérhető, hogy a_{1i_1}az első sorban magányos legyen.

Tegyük fel most, hogy az $a_{1i_1}, a_{2i_2}, ..., a_{k-1i_{k-1}}$ elemek magányosak az első $k-1$ sorban, és az $i_1, i_2, ..., i_{k-1}$ számok különbözők. Van olyan $a_{ki_k} \neq \neq 0$ elem, melyre i_k különbözik az $i_1, i_2, ..., i_{k-1}$ számoktól, hiszen $|\boldsymbol{A}| \neq 0$. Ezzel az elemmel a k-adik sor nullázható és elérhető, hogy a_{ki_k} is sorában magányos legyen. Az eljárás folytatható és az n-edik lépés után a kívánt végállás alakul ki. □

2. Kérdés. $\boldsymbol{A}$ nem négyzetes mátrix. $\boldsymbol{A}\boldsymbol{A}^*$ és $\boldsymbol{A}^*\boldsymbol{A}$ közül melyiknek a determinánsa biztosan nulla?

5.3. Kifejtési tétel

Természetesnek látszik az aldetermináns fogalma: a táblázat egy négyzetes részének a determinánsa. Gyakran azonban csupán egyetlen sort és oszlopot hagyunk el, ez jellemezhető a sor és az oszlop metszéspontjában álló elemmel. Amikor azt mondjuk, hogy az a_{ij} elemhez tartozó aldetermináns, akkor az i-edik sor és a j-edik oszlop elhagyásával kapott mátrix determinánsáról van szó. Az aldeterminánshoz rendszerint hozzátartozik egy előjel is, ezt bele szoktuk olvasztani a fogalomba, és ekkor előjeles aldeterminánsról beszélünk.

5.3.1. Definíció. Az a_{ij} elemhez tartozó *előjeles aldetermináns* az i-edik sor és a j-edik oszlop elhagyásával kapott mátrix determinánsa szorozva $(-1)^{i+j}$-nel. Szokásos jelölése: A_{ij}.

Az aldeterminánsok előjelezésére kaptunk egy másik előjelszabályt, az előjel csupán a_{ij} pozíciójától, i-től és j-től függ. Bejelölve a mátrixba az egyes pozíciókhoz tartozó előjelet, az alábbi ábrát kapjuk:

$$\begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix},$$

ennek alapján ezt az előjelszabályt *sakktábla-szabálynak* is nevezik.

5.3.2. Tétel (kifejtési tétel). Tetszőlegesen választott i mellett

$$|\boldsymbol{A}| = \sum_{j=1}^{n} a_{ij} A_{ij}.$$

Természetesen a j-edik oszlopot választva is el lehet végezni a kifejtést: $|\boldsymbol{A}| = \sum\limits_{i=1}^{n} a_{ij} A_{ij}$.

Szokásos a kifejtési tétellel is definiálni a determináns értékét, hiszen visszavezeti az $n \times n$-es determináns kiszámítását n darab $(n-1) \times (n-1)$-es determináns kiszámítására, így rekurzív definíciót nyújt. A determináns kiszámítására közvetlenül használni általában nem gazdaságos, túl sok munkát igényel.

Bizonyítás. Elég azt belátni, hogy a_{ij} szorzója tényleg A_{ij}. Gyűjtsük össze azokat az elemi szorzatokat, amelyekben a_{ij} szerepel, és emeljük ki belőlük a_{ij}-t. A zárójelbe kerülő tagok az a_{ij}-hez tartozó aldetermináns elemi szorzatai lesznek, csupán ezek előjeleit kell tisztázni.

Vegyünk egy elemi szorzatot. a_{ij} a mátrixot négy részre bontja:

$$\begin{pmatrix} \beta & \vdots & \alpha \\ --- & a_{ij} & --- \\ \gamma & \vdots & \end{pmatrix}$$

Az elemi szorzat tényezői a négy negyedbe esnek számukat jelölje α, β, γ és $n-1-\alpha-\beta-\gamma$. Ezek egymás közötti összekötései az aldeterminánsnak megfelelő előjelet adnak. Az $\boldsymbol{A}$ determinánsában azonban ezeket az a_{ij}-vel is össze kell kötni. Ezek közül a jobbra-felfelé haladók száma $\alpha+\gamma = (\alpha+\beta)+(\gamma+\beta)-2\beta = (i-1)+(j-1)-2\beta = i+j-2(\beta-1)$, ami előjelben $2(\beta-1)$ páros volta miatt $(-1)^{i+j}$-nek felel meg. Ez megfelel az aldetermináns előjelezési szabályának. $\square$

5.3.3. Következmény. Ha az i-edik sorhoz tartozó előjeles aldeterminánsokat egy másik sor elemeivel szorozva összegezzük, nullát kapunk. Képletben:

$$\sum_{j=1}^{n} a_{kj}A_{ij} = 0, \text{ ha } k \neq i.$$

Bizonyítás. Írjuk be az $\boldsymbol{A}$ mátrix i-edik sorvektora helyére a k-adik sorvektorát, és ezt a mátrixot fejtsük ki az i-edik sor szerint. A kifejtési tétel alapján a fenti összeggel a megváltoztatott mátrix determinánsát kapjuk meg, ami nulla, hiszen két sora megegyezik. $\square$

3. Kérdés. Ki lehet-e számolni a

$$\begin{vmatrix} 1 & 1 & 2 & -1 \\ 2 & & 4 & 0 \\ 2 & 4 & -1 & 3 \\ 3 & 6 & 1 & 2 \end{vmatrix}$$

determináns számszerű értékét, ha a hiányzó számot nem tudjuk elolvasni?

4. Kérdés. Egy számokkal megadott mátrixban egy elemet x-re cserélünk, és a determinánsát $f(x)$-szel jelöljük. Hogyan lehet kiszámolni $f'(x)$-et? Hogyan lehet ugyanezt kiszámítani, ha két elemet cserélünk ki x-re?

5.4. Az inverz mátrix számítása

Újabb számítási eljárást adhatunk az inverz mátrix meghatározására a determinánsok ismeretében. A számolás mindenképpen hosszadalmas, műveletigényes eljárás, ezt az eljárást is nehézkes elvégezni, de legalább explicit képletet kapunk.

Képezzünk mátrixot az adott mátrix előjeles aldeterminánsaiból:

$$\boldsymbol{S} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}.$$

Szorozzuk össze $\boldsymbol{A}$-t és $\boldsymbol{S}^*$-ot. A sor-oszlopszorzás miatt $\boldsymbol{A}$ sorait kell skalárisan megszorozni $\boldsymbol{S}$ soraival. 5.3.3. és 5.3.2. alapján különböző sorszámú sorok szorzata 0, míg az azonosaké $|\boldsymbol{A}|$ lesz, ez mátrixalakban azt jelenti, hogy $\boldsymbol{A}\boldsymbol{S}^* = |\boldsymbol{A}|\cdot\boldsymbol{I}$. Átrendezve $\boldsymbol{A}\frac{\boldsymbol{S}^*}{|\boldsymbol{A}|} = \boldsymbol{I}$, vagyis $\boldsymbol{A}^{-1} = \frac{\boldsymbol{S}^*}{|\boldsymbol{A}|}$, ha $|\boldsymbol{A}| \neq 0$. Ezzel lényegében bebizonyítottuk az alábbi tételt:

5.4.1 Tétel. $\boldsymbol{A}^{-1}$ akkor és csak akkor létezik, ha $|\boldsymbol{A}| \neq 0$. Ebben az esetben $\boldsymbol{A}^{-1} = \frac{\boldsymbol{S}^*}{|\boldsymbol{A}|}$.

Bizonyítás. A tétel előtt elmondottakat az $\boldsymbol{A}^{-1}$ létezésével kell kiegészíteni. Ha $|\boldsymbol{A}| = 0$, akkor 5.2.8. szerint $\text{rang}(\boldsymbol{A}) < n$, így R_A dimenziója kisebb, mint $\mathbf{R}^n$ dimenziója, tehát $\boldsymbol{A}\boldsymbol{x}$ nem szürjektív, ezért az inverze sem létezhet. Ha $|\boldsymbol{A}| \neq 0$, akkor rang $(\boldsymbol{A}) = n$, az R_A dimenziója n, így az $\boldsymbol{A}$ szürjektív. Ekkor $\boldsymbol{A}$ a 2.5.4. Következmény miatt bijektív, tehát $\boldsymbol{A}^{-1}$ létezik. $\square$

5. Kérdés. Igaz-e, hogy az $\boldsymbol{S}$ előjeles aldeterminánsaiból képezett mátrix elemei arányosak az $\boldsymbol{A}$ azonos pozícióban lévő elemeivel?

5.5. Köbtartalom

$\mathbf{R}^n$-ben derékszögű, egységvektorokból álló koordináta-rendszert feltételezünk, és az n-dimenziós köbtartalom egysége az egységkocka térfogata.

Adott n vektor, $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, …, $\boldsymbol{a}_n$, és az általuk kifeszített parallelepipedon térfogatát akarjuk kiszámítani. A P parallelepipedon ugyanúgy definiálható, mint az egységkocka:

$$P = \{\boldsymbol{x}\colon\ \boldsymbol{x} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \ldots + c_n\boldsymbol{a}_n,\ 0 \leq c_i \leq 1,\ \text{bármely } i\text{-re}\}.$$

(Ha az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, …, $\boldsymbol{a}_n$ vektorok lineárisan függetlenek, és bázisnak tekintjük ezeket, akkor ebben a bázisban az egységkockáról van szó.) Fogadjuk el a térfogatszámítás egyik alapképletét: parallelepipedon térfogata alapterület szorozva magasság. Itt az „alapterület” az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, …, $\boldsymbol{a}_n$ vektorok közül kiválasztott $n-1$ vektor által kifeszített parallelepipedon köbtartalma, míg a magasság a kihagyott vektor és a többiek által meghatározott altér távolsága. Ezzel a képlettel az n-dimenziós térfogat visszavezethető eggyel kisebb dimenziós parallelepipedon térfogatára, és az eljárás folytatható.

Ebből az alapképletből következik, hogy a parallelepipedon térfogata nem változik meg, ha bármely vektor végpontját a többi vektor által meghatározott $(n-1)$-dimenziós alaplappal párhuzamosan eltoljuk.

5.5.1. Tétel. Az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, …, $\boldsymbol{a}_n$ vektorok által kifeszített parallelepipedon térfogata az $|\boldsymbol{A}|$ abszolút értéke, ahol $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$.

Bizonyítás. Feltehető, hogy $|\boldsymbol{A}| \neq 0$. Ha ugyanis ez nem teljesül, akkor a megadott vektorok egyike lineárisan kifejezhető a többivel, vagyis ennek a vektornak a távolsága a többi által generált altértől nulla, ekkor a térfogat is nulla, tehát az állítás érvényes.

Az 5.2.7. tulajdonság miatt $|\boldsymbol{A}|$ értéke nem változik meg, ha egy oszlophoz hozzáadom egy másik oszlop többszörösét. A parallelepipedon térfogata sem változik meg ezzel, hiszen az alaplap síkjában fekvő vektort adok hozzá a megváltoztatandó vektorhoz (ld. 3.3.3. Tétel).

Az 5. 2. 9. Következmény szerint ilyen oszlopműveletekkel elérhető, hogy a mátrix minden nem nulla eleme teljesen magányos legyen, közben sem a determináns értéke, sem a térfogat nem változik. Az így kapott mátrixnak megfelelő parallelepipedon téglatest, melynek térfogata a magányos elemek szorzatának az abszolút értéke, de ez egyben a determináns abszolút értéke is. □

Mivel az $\boldsymbol{Ax}$ lineáris transzformáció a koordináta egységvektorok által meghatározott kockát az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, …, $\boldsymbol{a}_n$ vektorok által kifeszített parallelepipedonba viszi át, kimondhatjuk a tétel alapján, hogy ennek köbtartalmát d-szeresére változtatja, ahol d az $|\boldsymbol{A}|$ abszolút értéke. A térfogatváltozás ezen szabálya bármely (Jordan értelemben) létező térfogatú halmazra érvényes, hiszen kockákkal ezen térfogatok megközelíthetők. Azt mondhatjuk, hogy $|\boldsymbol{A}|$ abszolút értéke az $\boldsymbol{Ax}$ lineáris transzformáció „térfogati torzítását” adja meg.

Ha $\boldsymbol{f}(\boldsymbol{x})$: $\mathbf{R}^n \to \mathbf{R}^n$ függvény, akkor az n-dimenziós integrálok $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ transzformációjánál a $d\boldsymbol{y} = J(\boldsymbol{x})d\boldsymbol{x}$ helyettesítés lép fel, ahol $J(\boldsymbol{x})$ az ún. Jacoby-determináns. A Jacoby-determináns nem más, mint a transzformációt az $\boldsymbol{x}$ hely környezetében lineárisnak véve a determinánssal kiszámított „lokális térfogati torzulás”.

5.5.2. Következmény. Ha $\boldsymbol{A}$ ortogonális mátrix, akkor $|\boldsymbol{A}| = 1$ vagy $|\boldsymbol{A}| =$ -1.

Bizonyítás. Távolságtartás miatt az egységkocka képe egységoldalú kocka lesz, aminek a köbtartalma 1, tehát $|\boldsymbol{A}|$ abszolút értéke 1. □

Az elmondottakból következik, hogy $|\boldsymbol{BA}|$ abszolút értéke megegyezik $|\boldsymbol{B}|$ és $|\boldsymbol{A}|$ abszolút értékeinek szorzatával. Az $\boldsymbol{Ax}$ transzformáció ugyanis a térfogatot $||\boldsymbol{A}||$-szorosára változtatja ($||\boldsymbol{A}||$ az $\boldsymbol{A}$ determinánsának az abszolút értékét jelöli). A $\boldsymbol{Bx}$ transzformáció minden térfogatot, így az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ vektorok által kifeszített parallelepipedon térfogatát is, $||\boldsymbol{B}||$-szorosára változtatja meg, így az egységkocka köbtartalma a $\boldsymbol{BA}$ transzformáció során $||\boldsymbol{B}|| \cdot ||\boldsymbol{A}||$-szorosára változik, azaz $||\boldsymbol{B}\mathbf{A}|| = ||\boldsymbol{B}|| \cdot ||\boldsymbol{A}||$. Ennél azonban több is igaz.

5.5.3. Tétel. Ha $\boldsymbol{A}$ és $\boldsymbol{B}$ $n \times n$-es mátrixok, akkor $|\boldsymbol{BA}| = |\boldsymbol{B}| \cdot |\boldsymbol{A}|$.

Bizonyítás. Feltehetjük, hogy $|\boldsymbol{B}| \neq 0$, ugyanis, ha $|\boldsymbol{B}| = 0$, akkor $\boldsymbol{BA}$ sem teljes rangú mátrix, így $|\boldsymbol{BA}| = 0$, és a bizonyítandó állítás teljesül.

Számítsuk ki a $|\boldsymbol{BA}|$ determinánst. A kiszámolás alapgondolata az, hogy $\boldsymbol{BA}$-t olyan mátrixszal szorozzuk be, ami a determináns értékét nem változtatja meg. A $\boldsymbol{C}$ mátrix legyen olyan mátrix, melynek egy adott, főátlótól különböző pozíciójú eleme, mondjuk $c_{ij} = c$, a többi 0. Ha a $\boldsymbol{Q} = \boldsymbol{I} + \boldsymbol{C}$ mátrixszal beszorzok egy mátrixot balról, akkor a mátrix j-edik sorának a c-szerese hozzáadódik az i-edik sorához, miáltal a determinánsa nem változik meg. Ugyanez a helyzet jobbról történő szorzásnál, csak oszlopok vonatkozásában. A $\boldsymbol{Q} = \boldsymbol{I} + \boldsymbol{C}$ konstrukciójú mátrixokat a bizonyítás további részében nevezzük $\boldsymbol{Q}$-típusú mátrixnak.

A $\boldsymbol{BA}$ szorzatot lépésenként alakítsuk át a $|\boldsymbol{BA}| = |\boldsymbol{QBA}|$ egyenlőségek felhasználásával, a $\boldsymbol{Q}$-hoz tartozó $\boldsymbol{C}$ mátrixot mindig alkalmasan megválasztva. A $\boldsymbol{QB}$ szorzással a $\boldsymbol{B}$ mátrixon a determináns kiszámításakor használt sorműveleteket hajtjuk végre, melyek ismétléseivel elérhető, hogy a $|\boldsymbol{B}|$ nem nulla elemei egyetlen elemi szorzatot képezzenek (5.2.9. Következmény). Jelöljük $\boldsymbol{Q}_1, \boldsymbol{Q}_2, ..., \boldsymbol{Q}_k$-val a felhasznált $\boldsymbol{Q}$-típusú mátrixokat és $\boldsymbol{B}_1$-gyel a kapott mátrixot, akkor $\boldsymbol{B}_1 = \boldsymbol{Q}_1\boldsymbol{Q}_2 \cdots \boldsymbol{Q}_k\boldsymbol{B}$, $|\boldsymbol{B}| = |\boldsymbol{B}_1|$ és $|\boldsymbol{BA}| = |\boldsymbol{B}_1\boldsymbol{A}|$.

Képezzük azt a $\boldsymbol{P}$ permutáló mátrixot (ld. 4.4.6. Definíció), amelyben az egyesek a $\boldsymbol{B}_1$ mátrix nem nulla elemeinek helyén állnak. Egy determináns a $\boldsymbol{P}^*$-gal való bal oldali beszorzása után előjelet válthat, hiszen $\boldsymbol{P}^*$ az inverziószámmal megegyező számú sorcserét létesít (ld. 4.4.7. Tétel). A determináns elemi szorzatának előjelszabálya alapján ugyanez az előjele a $|\boldsymbol{P}|$-nak, tehát $|\boldsymbol{P}|\boldsymbol{P}^*$-gal történő bal oldali beszorzás a determinás értékét nem változtatja meg.

Szorozzuk be $|\boldsymbol{P}|\boldsymbol{P}^*$-gal balról a $\boldsymbol{B}_1\boldsymbol{A}$ mátrixot. A $\boldsymbol{B}_1$ mátrix a $\boldsymbol{P}^*$-gal történő bal oldali beszorzás után diagonálissá válik. Jelöljük a főátló elemeit

$d_1, d_2, \ldots, d_n$-nel, akkor ezen számok szorzata kiemelhető a determinánsból:

$$|\boldsymbol{B}\mathbf{A}| = |\,|\boldsymbol{P}|\boldsymbol{P} * \boldsymbol{B}_1\boldsymbol{A}| = d_1 d_2 ... d_n |\boldsymbol{P}| \cdot |\boldsymbol{A}| = d_1 d_2 ... d_n |\boldsymbol{P}| \cdot |\boldsymbol{A}| =$$
$$= |\boldsymbol{B}_1| \cdot |\boldsymbol{A}| = |\boldsymbol{B}| \cdot |\boldsymbol{A}|,$$

ugyanis $d_1 d_2 \ldots d_n |\boldsymbol{P}|$ a $\boldsymbol{B}_1$ mátrix egyetlen elemi szorzata a megfelelő előjellel ellátva. □

5.5.4. Következmény. $|\boldsymbol{A}^{-1}| = \frac{1}{|\boldsymbol{A}|}$.

Bizonyítás. $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$-re alkalmazzuk az 5.5.3 Tételt. □

6. Kérdés. Készítsünk el az $\boldsymbol{S}$ mátrixot az $\boldsymbol{A}$ mátrix előjeles aldeterminánsaiból (ld. 5.4. fejezet). Mennyi $\boldsymbol{S}$ determinánsa?

A hasonló mátrixok determinánsa egyenlő. Algebrailag ez könnyen bizonyítható, hiszen

$$|\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}| = |\boldsymbol{C}^{-1}||\boldsymbol{A}||\boldsymbol{C}| = \frac{|\boldsymbol{A}||\boldsymbol{C}|}{|\boldsymbol{C}|} = |\boldsymbol{C}|.$$

A forgatások mátrixának a determinánsa 1. A forgatásra megadott kanonikus alakból ez rögtön látható.

A tükrözések mátrixának a determinánsa 1 vagy -1 (és ez nem csak ortogonális tükrözésekre igaz!). Az állítás következik a tükrözés $\boldsymbol{A}^2 = \boldsymbol{I}$ definíciójából. A tükrözés kanonikus alakjából pedig az látszik, hogy a determináns $+1$, ha a vetítősík dimenziója páros, és -1, ha páratlan.

A determináns módot ad az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ vektorok által kifeszített parallelepipedon *előjeles térfogatának* a definiálására is, legyen ez egyszerűen $|\boldsymbol{A}|$. Természetesen a térfogat előjele függ az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ vektorok sorrendjétől. Önmagában az előjeles köbtartalomnak nem sok értelme van, hiszen mennyi egy gömb köbtartalma, pozitív, vagy negatív? A kérdés így értelmetlen. Az előjeles köbtartalomnak a lineáris transzformációval kapcsolatban van értelme.

A 4.6.2. Tétel úgy interpretálható, hogy az az $\boldsymbol{A}\boldsymbol{x}$ ortogonális transzformáció, melyre $|\boldsymbol{A}| = 1$, mindig előállítható forgatások sorozatával, míg $|\boldsymbol{A}| = -1$ esetében nem. Másképpen fogalmazva, $|\boldsymbol{A}| = 1$ esetén egy tetszőleges alakzat és transzformáltja mozgatással (forgatások sorozatával) fedésbe hozhatók, míg $|\boldsymbol{A}| = -1$ esetében nem teljesül, hogy az alakzat minden minden pontja és annak transzformáltja fedésbe kerüljön. Hétköznapi szóhasználattal: a bal cipőből mozgatással nem lesz jobb cipő. Az ok: egyiknek negatív, a másiknak pozitív a köbtartalma.

7. Kérdés. Mennyi az n-dimenziós kúp (kúpszerű test) köbtartalma? Igaz-e, hogy alap-köbtartalom szorozva magasság osztva n? (Osszuk fel az n-dimenziós egységkockát $(n-1)$-dimenziós kocka alapú egybevágó gúlákra, azaz kúpszerű testekre!)

5.6. A szimplex

A legegyszerűbb n-dimenziós test – legalábbis a neve szerint. A kétdimenziós háromszög és a háromdimenziós tetraéder általánosítása (feltételezzük, hogy $n \geq 2$).

Induljunk ki az 5.5. paragrafusban definiált parallelepipedonból. Felvettük az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ lineárisan független vektorokat, és a

$$P = \{\boldsymbol{x} : \boldsymbol{x} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_n\boldsymbol{a}_n,\ 0 \leq c_i \leq 1\}$$

halmazt neveztük az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ vektorok által meghatározott parallelepipedonnak. Az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ vektorok egyben affin függetlenek is, tehát az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ pontokon át $(n-1)$-dimenziós sík fektethető; messük ketté a parallelepipedont ezzel a síkkal, és az origót tartalmazó része lesz az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ vektorok által meghatározott szimplex.

A szimplex definícióját közvetlenül is meg lehet adni, ugyanis az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ pontokon átfektetett sík $S = \{\boldsymbol{x} : \boldsymbol{x} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_n\boldsymbol{a}_n,\ \sum_{i=1}^{n} c_i = 1\}$ alakú.

5.6.1. Definíció. Az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ vektorok által meghatározott *szimplex* az a T halmaz, melyre

$$T = \{\boldsymbol{x} : \boldsymbol{x} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_n\boldsymbol{a}_n,\ \sum_{i=1}^{n} c_i \leq 1,\ c_i \geq 0\}.$$

Ha az $\boldsymbol{a}$-csúcsú szimplexről beszélünk, akkor az a $T + \boldsymbol{a}$ halmazt jelenti. A kijelölt csúccsal szemközti $(n-1)$-dimenziós oldallapot alaplapnak tekintjük.

A definíció szerint T nem más, mint a $\mathbf{0}$, $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ pontok konvex burka.

8. Kérdés. Igaz-e, hogy T bármely két csúcsát él köti össze?

Az elmondottakból adódóan a szimplex kúpszerű test, pontosabban gúla (ld. 5. fejezet, 7. kérdés), melynek az alaptartománya az

$$A = \{\boldsymbol{x} : \boldsymbol{x} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_n\boldsymbol{a}_n,\ \sum_{i=1}^{n} c_i = 1,\ c_i \geq 0\}$$

sokszög, az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ pontok konvex burka, csúcsa pedig az origó.

5.6.2. Tétel. A szimplexnél az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_n$ pontok bármelyike választható a szimplex kijelölt csúcsának.

Bizonyítás. Ha $\boldsymbol{a}_n$-et választjuk a csúcsnak, akkor az $\boldsymbol{a}_n$-ből kiinduló $\boldsymbol{a}_1 - \boldsymbol{a}_n$, $\boldsymbol{a}_2 - \boldsymbol{a}_n$, $\ldots$, $\boldsymbol{a}_{n-1} - \boldsymbol{a}_n$, $-\boldsymbol{a}_n$ vektorok határozzák meg a szimplexet, de ekkor az $\boldsymbol{a}_n$ csúcs az origóba kerül. Ha a testet $\boldsymbol{a}_n$ -nel eltoljuk, akkor

T-vel megegyező halmazt kapunk. Ezt bizonyítandó, képezzük az $\boldsymbol{a}_1 - \boldsymbol{a}_n$, $\boldsymbol{a}_2 - \boldsymbol{a}_n, \ldots, \boldsymbol{a}_{n-1} - \boldsymbol{a}_n, - \boldsymbol{a}_n$ vektorok által meghatározott T_1 szimplexet:

$$T_1 = \{\boldsymbol{x} : \boldsymbol{x} = d_1(\boldsymbol{a}_1 - \boldsymbol{a}_n) + d_2(\boldsymbol{a}_2 - \boldsymbol{a}_n) + ... \\ + d_{n-1}(\boldsymbol{a}_{n-1} - \boldsymbol{a}_n) - d_n\boldsymbol{a}_n, \sum_{i=1}^{n} d_i \leq 1,\ d_i \geq 0\} =$$

$$= \{\boldsymbol{x} : \boldsymbol{x} = d_1\boldsymbol{a}_1 + ... + d_{n-1}\boldsymbol{a}_{n-1} - (d_1 + ... + d_n)\boldsymbol{a}_n, \sum_{i=1}^{n} d_i \leq 1,\ d_i \geq 0\}.$$

T_1-et $\boldsymbol{a}_1$-gyel eltolva

$$T_1 + \boldsymbol{a}_n = \{\boldsymbol{x} : \boldsymbol{x} = d_1\boldsymbol{a}_1 + ... + d_{n-1}\boldsymbol{a}_{n-1} + (1 - (d_1 + ... + d_n))\boldsymbol{a}_n, \\ \sum_{i=1}^{n} d_i \leq 1,\ d_i \geq 0\},$$

amiről könnyen belátható, hogy megegyezik a T halmazzal.

Legyen

$$c_i = d_i, \text{ ha } i = 1, 2, \ldots, n-1, \text{ és} \\ c_n = 1 - (d_1 + d_2 + \ldots + d_n).$$

Ha $d_i \geq 0$ és $\sum d_i \leq 1$, akkor $c_i \geq 0$, és $\sum c_i = 1 - d_n \leq 1$. Ha viszont $c_i \geq \geq 0$ és $\sum c_i \leq 1$, akkor $d_i \geq 0$, és $\sum d_i = 1 - c_n \leq 1$. Vagyis az együtthatókra vonatkozó feltételek megegyeznek. □

5.6.3. Tétel. Az n-dimenziós szimplex alaplapja $(n-1)$-dimenziós szimplex.

Bizonyítás. A bizonyítás az előzővel lényegében azonos. A szimplex alaplapja az

$$A = \{\boldsymbol{x} : \boldsymbol{x} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_n\boldsymbol{a}_n, \sum_{i=1}^{n} c_i = 1,\ c_i \geq 0\}$$

halmaz. Az alaplap $\boldsymbol{a}_n$ csúcspontjából kiinduló $\boldsymbol{a}_1 - \boldsymbol{a}_n$, $\boldsymbol{a}_2 - \boldsymbol{a}_n, \ldots$, $\boldsymbol{a}_{n-1} - \boldsymbol{a}_n$ vektorok az

$$T' = \{\boldsymbol{x} : \boldsymbol{x} = d_1(\boldsymbol{a}_1 - \boldsymbol{a}_n) + d_2(\boldsymbol{a}_2 - \boldsymbol{a}_n) + ... + d_{n-1}(\boldsymbol{a}_{n-1} - \boldsymbol{a}_n), \\ \sum_{i=1}^{n-1} d_i \leq 1,\ d_i \geq 0\}$$

0-csúcsú szimplexet határozzák meg. Meg kell mutatni, hogy az T'-t $\boldsymbol{a}_n$-nel eltolva,

$$T' + \boldsymbol{a}_n = \{\boldsymbol{x} : \boldsymbol{x} = d_1\boldsymbol{a}_1 + d_2\boldsymbol{a}_2 + ... + d_{n-1}\boldsymbol{a}_{n-1} + (1 - (d_1 + ... + d_{n-1}))\boldsymbol{a}_n, \\ \sum_{i=1}^{n-1} d_i \leq 1,\ d_i \geq 0\},$$

az A-t kapjuk.

Legyen

$$c_i = d_i,\ \text{ha } i = 1, 2, \ldots, n-1,\ \text{és}$$
$$c_n = 1 - (d_1 + d_2 + \ldots + d_{n-1}),$$

akkor a $d_i \geq 0$, $\sum d_i \leq 1$, és a $c_i \geq 0$, $\sum c_i = 1$ ekvivalens tulajdonságpárok. $\square$

Az n-dimenziós szimplexnek $n+1$ csúcsa van, bármely három csúcs meghatároz egy (kétdimenziós) háromszög-lapot, tehát $\binom{n+1}{3}$ kétdimenziós lapja van. Bármely négy csúcs meghatároz egy háromdimenziós, tetraéder alakú oldallapot, ezek száma $\binom{n+1}{4}$, és így tovább.

A 7. Kérdésben meghatároztuk egy speciális, kocka alapú gúla térfogatát. A gúla alaplapja egységkocka, magassága pedig az egyik egységnyi hosszúságú oldalél, ekkor a térfogat $\frac{1}{n}$-nek adódott. Helyezzük el a gúlát koordináta-rendszerben úgy, hogy az alaplapot az első $n-1$ koordináta egységvektor feszítse ki, és a magassága az n-edik koordinátatengelyre essen. Alkalmazzuk a gúlára az

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & \cdots & & a_1 \\ & 1 & 0 & & & a_2 \\ & & & \ddots & & \vdots \\ & 0 & & & 1 & a_{n-1} \\ & & & & & 1 \end{pmatrix}$$

transzformációt. Ha $\boldsymbol{x} = (x_1, x_2, ..., x_{n-1}, 0)$, akkor $\boldsymbol{Ax} = \boldsymbol{x}$, tehát a gúla alaplapját a transzformáció helybenhagyja, nem változtatja meg. Ha $\boldsymbol{x} = (0, 0, \ldots, 0, 1)$, akkor $\boldsymbol{Ax} = \boldsymbol{x} + \boldsymbol{a}$, ahol $\boldsymbol{a} = (a_1, a_2, ..., a_{n-1}, 0)$, vagyis a csúcs az alaplappal párhuzamosan $\boldsymbol{a}$-val eltolódik. A gúla térfogata a transzformáció során nem változik meg, mert $|\boldsymbol{A}| = 1$. Ha $m > 0$, $a > 0$ és a

$$\boldsymbol{B} = \begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & a & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & m \end{pmatrix}$$

transzformációt alkalmazzuk, akkor a gúla magassága m-szeresére, az alaplap kocka éle a-szorosára változik, de $|\boldsymbol{B}| = a^{n-1}m$, tehát a térfogat $a^{n-1}m$ -szeresére változik. Tetszőleges kocka alapú gúla térfogata tehát $\frac{a^{n-1}m}{n}$. Mivel tetszőleges alaplap kockákkal megközelíthető, bármely gúla köbtartalma kiszámítható úgy, hogy az alaplap $(n-1)$-dimenziós köbtartalmát szorozzuk a magassággal és osztjuk n-nel.

Számoljuk ki a koordináta egységvektorok által kifeszített szimplex t_n térfogatát! Vegyük a $(0, 0, \ldots, 0, 1)$ pontot a gúla csúcspontjának, akkor az alaplap az első $n-1$ koordináta egységvektor által meghatározott $(n-1)$-dimenziós szimplex. Az előző képletet alkalmazva $t_n = \frac{1}{n}t_{n-1}$, majd ezt ismételten felhasználva: $t_n = \frac{1}{n(n-1)}t_{n-2} = ... = \frac{1}{n(n-1)...3}t_2 = \frac{1}{n!}$.

Az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ vektorok által meghatározott szimplex a koordináta egységvektorok által meghatározott szimplexből az $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ lineáris transzformációval származtatható, tehát köbtartalma az előző $||\boldsymbol{A}||$-szorosa ($||\boldsymbol{A}||$ az $\boldsymbol{A}$ determinánsának az abszolút értéke). Kimondható tehát a következő tétel:

6.5.4. Tétel. Az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ vektorok által meghatározott szimplex térfogata $\frac{1}{n!}||\boldsymbol{A}||$.

9. Kérdés. Előállítunk n darab véletlen számot (a $(0, 1)$ intervallumban egyeletes eloszlással és függetlenül). Mi a valószínűsége, hogy monoton növekedő sorozatot alkotnak? Igaz-e ugyanez, ha tetszőleges eloszlású véletlen mintát nézünk? Ha addig kérjük a véletlen számokat, amíg ki nem derül a belőlük készített monoton növekedő sorozat pontos hossza, várhatóan hány véletlen számot kell előállítani?

Létezik-e n-dimenziós szabályos szimplex? Van-e a szabályos tetraédernek n-dimenziós általánosítása? A kérdés valójában az, hogy megadhatók-e az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ egyenlő hosszúságú vektorok úgy, hogy bármely kettő 60^o-os szöget zárjon be egymással. A vektorok konkrét megadása nem túl kellemes feladat, de ilyen vektorok létezését az 1. fejezet 10. kérdése tisztázza. Ott az $\mathbf{R}^n$ $(n-1)$-dimenziós alterében adunk meg $n-1$ darab vektort úgy, hogy páronként 60^o-os szöget zárnak be egymással. Ez $(n-1)$ dimenzió esetén megoldja a létezési feladatot.

Az n-dimenziós szabályos szimplex tehát létezik, bármely két csúcspontját él köti össze, minden éle egyenlő hosszúságú, minden kétdimenziós oldallapja szabályos háromszög, minden háromdimenziós oldallapja szabályos tetraéder, és így tovább.

Határozzuk meg az egységoldalú szabályos szimplex térfogatát! Tegyük fel, hogy az $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ vektorok határozzák meg a szabályos szimplexet, és legyen $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$, akkor a feladat az $|\boldsymbol{A}|$ kiszámítására redukálódik. E helyett az $|\boldsymbol{A}|^2 = |\boldsymbol{A}^*\boldsymbol{A}|$ determinánst számoljuk ki. A $\boldsymbol{B} = \boldsymbol{A}^*\boldsymbol{A}$ mátrix az $\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle$ elemekből áll, ami $i = j$ esetén 1, $i \neq j$ esetén $\frac{1}{2}$. Célszerű a mátrixot beszorozni 2^n-nel, akkor mátrix főátlójában 2-esek, máshol 1-esek állnak:

$$2^n \boldsymbol{B} = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{pmatrix}.$$

A determináns kiszámolása érdekében vonjuk ki a második sorból az elsőt, majd fejtsük ki a második sor szerint. Az első determináns könnyen kiszámíthatóan 1-et ad, míg a második ugyanazt a feladatot adja vissza, ezért

$$d_n = 2^n |\boldsymbol{B}| \ = \ 1 + d_{n-1}.$$

Mivel $d_2 = 3$, $d_n = n + 1$. Visszatérve $|\boldsymbol{A}| = \sqrt{|\boldsymbol{B}|} = \sqrt{\frac{n+1}{2^n}}$, tehát az egységoldalú szabályos szimplex térfogata

$$V_n = \frac{\sqrt{n+1}}{n!\sqrt{2^n}}.$$

10. Kérdés. Határozzuk meg az n-dimenziós egységkocka origóból kiinduló $(n-1)$-dimenziós lapátlói által meghatározott szimplex térfogatát! Szabályos-e ez a szimplex?

11. Kérdés. Mutassuk meg, hogy a koordináta egységvektorok által meghatározott szimplex alaplapja szabályos szimplex.

5.7. Cramer-szabály

Térjünk vissza a lineáris egyenletrendszerekre. A legegyszerűbb esetet jelenti az, amikor ugyanannyi egyenlet, mint ismeretlen van, és az egyenletrendszer mátrixa teljesrangú. Képletekkel elmondva, keressük az $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ egyenlet megoldását, amikor $\boldsymbol{A}$ $n \times n$-es mátrix és $|\boldsymbol{A}| \neq 0$. Ekkor az $\boldsymbol{A}$ mátrix inverze létezik, és szorozzuk meg az egyenletrendszer mindkét oldalát balról $\boldsymbol{A}^{-1}$-gyel: $\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$, azaz

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}.$$

Előállítottuk tehát az egyenletrendszer megoldását, feltéve, hogy az inverz mátrixot ismerjük. Az inverz mátrix kiszámítása azonban munkaigényes feladat, a megoldást valamivel könnyebb alakban találja nekünk a Cramer-szabály.

5.7.1. Tétel (Cramer-szabály). Ha az $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ egyenletrendszerben $\boldsymbol{A}$ $n \times n$-es mátrix és $|\boldsymbol{A}| \neq 0$, akkor az egyenletrendszer egyértelműen megoldható, és az $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ megoldás az

$$x_k = \frac{|\boldsymbol{A}_k|}{|\boldsymbol{A}|} \quad (k = 1, 2, \ldots, n)$$

alakban adható meg, ahol az $\boldsymbol{A}_k$ mátrix az $\boldsymbol{A}$-ból úgy kapható meg, hogy $\boldsymbol{A}$ k-adik oszlopát $\boldsymbol{b}$-re cseréljük.

Bizonyítás. Használjuk az inverz mátrix $\boldsymbol{A}^{-1} = \frac{\boldsymbol{S}^*}{|\boldsymbol{A}|}$ előállítását (5.4.1. Tétel), akkor

$$\boldsymbol{x} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{S}^*\boldsymbol{b}.$$

Írjuk fel a k-adik koordinátára vonatkozó összefüggést. $\boldsymbol{S}^*\boldsymbol{b}$ k-adik koordinátája az $\boldsymbol{S}$ k-adik oszlopa skalárisan szorozva a $\boldsymbol{b}$-vel:

$$x_k = \frac{1}{|\boldsymbol{A}|}\sum_{i=1}^{n} A_{ik}b_i.$$

A képletben lévő összeg azonban a k-adik oszlopra alkalmazott kifejtési tétel miatt annak a mátrixnak a determinánsa, amelyben $\boldsymbol{A}$ k-adik oszlopát $\boldsymbol{b}$-re kicseréltük. □

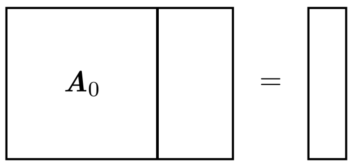
A numerikus számolás a Cramer-szabállyal még mindig eléggé nehézkes. Azt szokták mondani, hogy a Cramer-szabály legfőbb érdeme, hogy a megadott feltételek mellett biztosítja az egyértelmű megoldást, de ezt korábban, a dimenzió-tétel (2.5.3. Tétel) kapcsán már megtudtuk.

12. Kérdés. Az $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, $|\boldsymbol{A}| \neq 0$ egyenletrendszerrel kapcsolatban mit kapunk, ha az $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ jelölést felhasználva kiszámítjuk az $\frac{1}{|\boldsymbol{A}|}|(\boldsymbol{a}_1 - \boldsymbol{a}_2, \boldsymbol{b}, \boldsymbol{a}_3, ..., \boldsymbol{a}_n)|$ kifejezést?

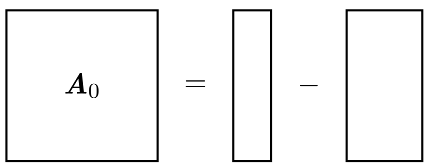
Tekintsük át az általános alakú lineáris egyenletrendszer algebrai megoldási lehetőségeit. A Cramer-szabály itt csak illusztratív jelentőségű, bármely más algebrai megoldással (behelyettesítés, egyenlő együtthatók módszere, stb.) pótolható, és rendszerint célszerű is pótolni.

1. eset. Ha az egyenletrendszer $\boldsymbol{A}$ mátrixa $n \times m$-es és $m > n$ (több ismeretlen van, mint egyenlet), és az $\boldsymbol{A}$ mátrix rangja n, akkor válasszunk ki n darab lineárisan független oszlopvektort, a többi oszlopot a hozzátartozó ismeretlenekkel szorozva vigyük át a jobb oldalra és olvasszuk bele a $\boldsymbol{b}$-be. A jobb oldalra átkerülő ismeretleneknek tetszőleges értéket adva az egyeletrendszer a Cramer-szabállyal megoldható, de mivel minden értékadáshoz tartozik egy megoldás, az egyenletrendszernek végtelen sok megoldása lesz. Az általános megoldás úgy kapható, hogy a jobb oldali ismeretleneket paramétereknek tekintve oldjuk meg a feladatot. A szabadon választható paraméterek száma $m - n$.

Szemléltessük grafikusan a leírt eljárást. A felírt egyenletrendszert most csak téglalapokkal szemléltetve:



átrendezve:

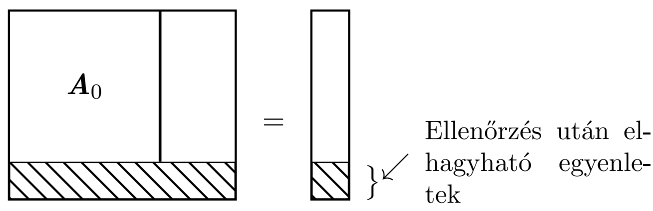


Megjegyzendő, hogy $\boldsymbol{A}_0$, a lineárisan független oszlopvektorokból képezett mátrix, nem feltétlenül az első n oszlopból képezhető, ilyenkor a fenti szemléltetés átrendezéssel érhető el.

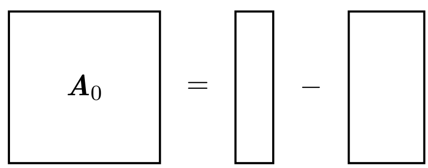
2. (általános) eset. Ha az egyenletrendszer $\boldsymbol{A}$ mátrixa $n \times m$-es, ahol n és m tetszőleges, és $\boldsymbol{A}$ rangja r, akkor válasszunk ki $\boldsymbol{A}$-ból egy $r \times r$-es $\boldsymbol{A}_0$ részmátrixot, melynek rangja továbbra is r. Ez megtehető úgy, hogy kiválsztunk először r lineárisan független oszlopvektort, majd a belőlük képezett mátrixból, amelynek a rangja továbbra is r, kiválsztunk r lineárisan független sorvektort (ld. 2.5.5. rangszám-tétel). A tényleges kiválasztást a 2.6.-ban leírt eljárás megkönnyíti.

Ha $r < n$, akkor ellenőrizzük a megoldhatóság feltételét (3.2.1. Tétel), ha nincs megoldás, akkor további teendő nincs, ha van megoldás, akkor azok az egyenletek, melyek együtthatóit $\boldsymbol{A}_0$-ba nem választottuk be, elhagyhatók, a kapott megoldások ezeket automatikusan ki fogják elégíteni. Ez az állítás a rangszám-tételből következik, ugyanis a $\boldsymbol{b}$-beli elemmel kibővített, elhagyásra ítélt sorvektorok már lineárisan kifejezhetők a kiválasztott sorvektorokkal. A többi egyenletre az 1. esetben leírt eljárás alkalmazható. A megoldások száma $m = r$ esetén egy, $m > r$ esetén végtelen sok, melyek $m - r$ db. *szabad változó* függvényeként adhatók meg.

Szemléltessük ezt is az előző módon:



átrendezve:



Természetesen ezek a grafikus szemléltetések itt is csak átrendezés után valósulhatnak meg.

Az általános esetet képletekkel is megfogalmazhatjuk. A megoldhatóság ellenőrzése után a felesleges egyenletek elhagyhatók, és a megmaradt egyenletek $\boldsymbol{A}_0\boldsymbol{x}_1 + \boldsymbol{B}\boldsymbol{x}_2 = \boldsymbol{b}_1$ alakra rendezhetők, ahol $\boldsymbol{A}_0$ invertálható mátrix, $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ pedig az ismeretleneket tartalmazó vektor. Ebből az általános megoldás: $\boldsymbol{x}_1 = \boldsymbol{A}_0^{-1}\boldsymbol{b}_1 - \boldsymbol{A}_0^{-1}\boldsymbol{B}\boldsymbol{x}_2$, és itt az $\boldsymbol{x}_2$ tartalmazza a szabad változókat.

5.8. Az MsExcel felhasználása mátrixműveleteknél

Mátrixműveleteink numerikus kivitelezését az Excel használata megkönnyítheti, vagy az ellenőrzését segítheti. A műveletek elvégezhetőségét a tömbök méretére vonatkozó feltételek szabják meg, ha ez nem teljesül, akkor az Excel hibát jelez.

Vektorműveleteknél a vektorok összeadása és a számmal történő szorzása a mátrixműveletek speciális eseteként végezhető el. Külön függvény készíti el azonban a skalárszorzatot és a norma négyzetét. A matematikai és trigonometriai függvények között található a SZORZATÖSSZEG(tömb1;

tömb2) függvény, mely a tömb1-ben és a tömb2-ben megadott vektorok skalárszorzatát, a NÉGYZETÖSSZEG(tömb) függvény a tömbben megadott vektor normájának a négyzetét számolja ki.

Mátrix összeadás és számmal történő szorzás egyszerűen tömbökkel végzett műveletekként elvégezhetők. Pl. $\boldsymbol{A} + 2\boldsymbol{B}$ számítása: ha $\boldsymbol{A}$ az A1:D3 tömbben, a $\boldsymbol{B}$ az A4:D6 tömbben van elhelyezve, akkor ki kell jelölni egy 3×4-es tartományt, ahová az eredményt elhelyezi, és az tartomány aktív cellájába = A1:D3 + 2*A4:D6 írandó. *Ne felejtsük el, ha a számolás eredménye tömb, akkor ezt Shift + Control-lal jelezni az Enter lenyomása alatt!* Ha az eredménytartományt rosszul jelöljük ki, akkor a feleslegesen kijelölt cellákba #HIÁNYZIK felirat jelenik meg, a kijelölt tartományon kívüli adatok pedig hiányozni fognak.

A további mátrixműveletekre beépített függvény áll a rendelkezésünkre. A matematikai és trigonometriai függvények között található az MSZORZAT(tömb1; tömb2) függvény a mátrixok szorzásának, az INVERZ.MÁTRIX (tömb) a mátrix inverzének a kiszámítására. A mátrix csoportban van a TRANSZPONÁLÁS(tömb) függvény a transzponált kiszámítására. Valamennyi függvény eredménye tömb, tehát az előzőekben leírtakat (Shift + Control + Enter, vagy Shift + Control + Kész) kell alkalmazni.

Determináns számolására az MDETERM(tömb) függvény használható. Az eredmény skalár, tehát a Shift + Control lenyomása felesleges.

Ha inverz mátrixot számolunk, és az eredeti mátrix egész számokból épül fel, akkor az inverz mátrix az esetek nagy részében törtszámokból áll, amit az Excel tizedes törtként ad meg. Mivel $\boldsymbol{A}^{-1} = \frac{\boldsymbol{S}^*}{|\boldsymbol{A}|}$, és $\boldsymbol{S}$ egészszámokból áll, ajánlatos először az $|\boldsymbol{A}|$-t kiszámítani, majd az $|\boldsymbol{A}|\cdot\boldsymbol{A}^{-1}$-et, így megkaphatjuk közönséges tört alakban az eredményt.

5.9. Válaszok a kérdésekre

1. $D = \begin{vmatrix} 0 & 0 & 5 & 0 \\ 2 & 0 & 0 & -1 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 0 & 12 \end{vmatrix} =?$

Egyetlen nem nulla elemi szorzat képezhető: $5\cdot2\cdot4\cdot12 = 480$. Az összekötések közül kettő halad jobbra-felfelé: a 2 – 5 és a 4 – 5. Ennek megfelelően $D = 480$.

Más meggondolással: a sorok szerint rendezve az oszlopindexek: 3, 1, 2, 4; az inverziók száma 2, mert a 3 megelőzi az 1-et és a 2-t, és több inverzió nincs. Ezért az inverziók száma páros, az elemi szorzat előjele pozitív.

2. $\boldsymbol{A}$ nem négyzetes mátrix. $\boldsymbol{AA}^*$ és $\boldsymbol{A}^*\boldsymbol{A}$ közül melyiknek a determinánsa biztosan nulla?

$\boldsymbol{A}$ mérete legyen $n \times m$, ahol $n > m$. Az értékkészlet definíciója alapján $R_{AA*} \subset R_A$, tehát $\mathrm{rang}(\boldsymbol{AA}^*) \leq \mathrm{rang}(\boldsymbol{A}) \leq m$. Az $\boldsymbol{AA}^*$ $n \times n$-es mátrix tehát nem teljesrangú, determinánsa biztosan nulla. (Az $\boldsymbol{A}^*\boldsymbol{A}$-ról ugyanez nem állítható.)

3. Ki lehet-e számolni a

$$D = \begin{vmatrix} 1 & 1 & 2 & -1 \\ 2 & & 4 & 0 \\ 2 & 4 & -1 & 3 \\ 3 & 6 & 1 & 2 \end{vmatrix}$$

determináns számszerű értékét, ha a hiányzó számot nem tudjuk elolvasni?

Jelöljük a hiányzó számot x-szel, és fejtsük ki a determinánst a második sor szerint:

$$D = -2\begin{vmatrix} 1 & 2 & -1 \\ 4 & -1 & 3 \\ 6 & 1 & 2 \end{vmatrix} + x\begin{vmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \\ 3 & 1 & 2 \end{vmatrix} - 4\begin{vmatrix} 1 & 1 & -1 \\ 2 & 4 & 3 \\ 3 & 6 & 2 \end{vmatrix} =$$

$$= -2 \cdot 5 + x \cdot 0 - 4 \cdot (-5) = 10.$$

(A jelenséget az okozza, hogy $A_{22} = 0$.)

4. Egy számokkal megadott mátrixban egy elemet x-re cserélünk, és a determinánsát $f(x)$-szel jelöljük. Hogyan lehet kiszámolni $f'(x)$-et? Hogyan lehet ugyanezt kiszámítani, ha két elemet cserélünk ki x-re?

Az x sora szerint fejtsük ki a determinánst, az eredmény: konstans + $A_{ij} \cdot x$, ennek a deriváltja A_{ij}, vagyis a derivált értéke a kicserélt elemhez tartozó előjeles aldetermináns.

A második kérdés kissé nehezebb. A determinánsba először u-t és v-t írjunk be, majd helyettesítsünk $u = u(x) = x$-et és $v = v(x) = x$-et. A deriváltját a kétváltozós összetett függvény deriválási szabályával („láncszabály") határozzuk meg. A két elem, melyet u-val ill. v-vel helyettesítünk, legyen a_{ij} és a_{kl}, akkor a derivált: $A_{ij} + A_{kl}$, hiszen a parciális deriváltat az előbbi lépésben meghatároztuk. (Természetesen most már A_{ij} és A_{kl} függ x-től.)

5. Igaz-e, hogy az $\boldsymbol{S}$ előjeles aldeterminánsaiból képezett mátrix elemei arányosak az $\boldsymbol{A}$ azonos pozícióban lévő elemeivel?

Tudjuk, hogy $\boldsymbol{A}^{-1} = \frac{\boldsymbol{S}^*}{|\boldsymbol{A}|}$, majd mindkét oldal inverzét képezve: $\boldsymbol{A} = |\boldsymbol{A}| \cdot \boldsymbol{S}^{*-1} = \frac{|\boldsymbol{A}|}{|\boldsymbol{S}|}\boldsymbol{T} = a\boldsymbol{T}$, ahol $\boldsymbol{T}$ az $\boldsymbol{S}$ előjeles aldeterminánsaiból képezett mátrix. (Felhasználtuk, hogy $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$.) Az $|\boldsymbol{S}|$ értékét a 6. Kérdésben kiszámítjuk, az eredmény ide behelyettesíthető.

6. Készítsünk el az $\boldsymbol{S}$ mátrixot az $\boldsymbol{A}$ mátrix előjeles aldeterminánsaiból. Mennyi $\boldsymbol{S}$ determinánsa?

5.5.4. alapján $|\boldsymbol{A}^{-1}| = \frac{1}{|\boldsymbol{A}|}$, másrészt $|\boldsymbol{A}^{-1}| = \left|\frac{\boldsymbol{S}*}{|\boldsymbol{A}|}\right| = \frac{|\boldsymbol{S}|}{|\boldsymbol{A}|^n}$, amiből $|\boldsymbol{S}| = |\boldsymbol{A}|^{n-1}$.

7. Mennyi az n-dimenziós kúp (kúpszerű test) köbtartalma? Igaz-e, hogy alap-köbtartalom szorozva magasság osztva n? (Osszuk fel az n-dimenziós egységkockát $(n-1)$-dimenziós kocka alapú egybevágó gúlákra!)

Mit nevezünk kúpszerű testnek? A kúpszerű testet egy $(n-1)$-dimenziós tartomány, az alaplap és egy, az alaplap síkján kívüli pont határozza meg. Jelöljük az alaplap pontjainak a halmazát T-vel, és adott pont legyen $\boldsymbol{x}_0$, akkor a kúpszerű test az

$$\{\boldsymbol{x} : \lambda\boldsymbol{x}_0 + (1-\lambda)\boldsymbol{y}, \boldsymbol{y} \in T, \quad 0 \leq \lambda \leq 1\}$$

halmaz.

A kocka belső pontjait az origóból vetítsük rá az $(n-1)$-dimenziós oldallapokra. Az $\boldsymbol{x}$ pont kivetítése az a $c\boldsymbol{x}$ pont lesz, amelynek minden koordinátája ≤ 1, de legalább egy koordináta eggyé válik. Ha a k-adik koordináta válik eggyé, akkor arra az oldallapra kerül, amelyet a k-adik koordináta 1 volta jellemez. Az origó és az ilyen oldallapok egyike egy-egy gúlát határoz meg. Nyilvánvaló, hogy a kocka minden pontja valamelyik gúlához hozzátartozik, és összesen n gúla keletkezik. Mivel a k-adik gúla pontjait az jellemzi, hogy a pontok k-adik koordinátája a legnagyobb, ortogonális transzformációval az egyik gúla a másikba átvihető (még 1 determinánsú ortogonális transzformációval is, vagyis mozgatással fedésbe hozhatók).

Egy ilyen gúla térfogata tehát $\frac{1}{n}$.

A gúla térfogata az alapterület additív függvénye, a magasságnak lineáris függvénye, ebből minden kúpszerű testre kapjuk az „alap-köbtartalom szorozva magasság osztva n” szabályt.

8. Igaz-e, hogy a T szimplex bármely két csúcsát él köti össze?

Igaz. Egyrészt a konvexitás miatt az összekötő szakasz a szimplexhez tartozik, és bármely $\boldsymbol{x}$ pontja a két végpont $\boldsymbol{a}_i$ és $\boldsymbol{a}_j$ konvex lineáris kombinációja: $\boldsymbol{x} = \lambda\boldsymbol{a}_i + (1 - \lambda)\boldsymbol{a}_j$, $\lambda \in (0, 1)$. Tegyük fel, hogy $\boldsymbol{x}$ egy két-

(vagy több-) dimenziós oldallap belső pontja, akkor $\boldsymbol{x}$ belső pontja ennek az oldallapnak, vagyis $\boldsymbol{x}$ valamely, erre a lapra eső környezete is T-hez tartozik. Vegyünk fel ebben a környezetben egy $\boldsymbol{x}+\boldsymbol{r}$ pontot, mely nem esik az $\boldsymbol{a}_i$ és $\boldsymbol{a}_j$ pontokat összekötő szakaszra, akkor $\boldsymbol{x}+\boldsymbol{r}$ baricentrikus koordinátákkal felírható

$$\boldsymbol{x}+\boldsymbol{r}=\sum_{k=0}^{n} c_k\boldsymbol{a}_k$$

alakban, ahol $\boldsymbol{a}_0=\boldsymbol{0}$, $\sum\limits_{k=0}^{n} c_k=1$, $c_k\geq 0$ és van olyan $c_l>0$, melyre $l\neq i$ és $l\neq j$. Ebből

$$\boldsymbol{x}-\boldsymbol{r}=2(\lambda\boldsymbol{a}_i+(1-\lambda)\boldsymbol{a}_j)-\sum_{k=0}^{n} c_k\boldsymbol{a}_k,$$

ami nem pontja T-nek, hiszen negatív (baricentrikus) koordinátája is van, de a fenti környezethez tartozik. $\boldsymbol{x}$ tehát nem lehet belső pontja az oldallapnak.

9. Előállítunk n darab véletlen számot (a $(0,\,1)$ intervallumban egyenletes eloszlással és függetlenül). Mi a valószínűsége, hogy monoton növekedő sorozatot alkotnak? Igaz-e ugyanez, ha tetszőleges eloszlású véletlen mintát nézünk? Ha addig kérjük a véletlen számokat, amíg ki nem derül a belőlük készített monoton növekedő sorozat pontos hossza, várhatóan hány véletlen számot kell előállítani?

Az n darab véletlen számból készített vektor egyenletes eloszlású az E egységkockában (egyenletes eloszlások direkt szorzata), tehát az egységkocka egy részhalmazának a valószínűsége megegyezik a térfogatával. Ki kell számítani annak az $A\subset E$ halmaznak a térfogatát, amely rendelkezik azzal a tulajdonsággal, hogy minden pontjának koordinátái monoton növekedőek. Azt állítjuk, hogy A az

$$\begin{aligned}\boldsymbol{a}_1&=(1,1,1,\ldots,1),\\ \boldsymbol{a}_2&=(0,1,1,\ldots,1),\\ \boldsymbol{a}_3&=(0,0,1,\ldots,1),\\ &\cdots\\ \boldsymbol{a}_n&=(0,0,\ldots,0,1)\end{aligned}$$

vektorok által kifeszített szimplex.

Egyrészt az $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, …, $\boldsymbol{a}_n$ vektorok nyilván hozzátartoznak A-hoz, és tetszőleges nemnegatív c_1, c_2, …, c_n számokra, melyekre $\sum c_i\leq 1$ a $\boldsymbol{x}=$

$= c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_n\boldsymbol{a}_n$ vektor is hozzátartozik A-hoz, mert koordinátái monoton növekedőek, és $\boldsymbol{x}$ n-edik koordinátája is legfeljebb 1. Másrészt ha $\boldsymbol{x} = (x_1, x_2, ..., x_n) \in A$, akkor $\boldsymbol{x}$ a $c_i = x_i - x_{i-1}$ ($x_0 = 0$) választás mellett előállítható az $\boldsymbol{x} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + ... + c_n\boldsymbol{a}_n$ alakban, ahol $c_i \geq 0$, és $\sum c_i = x_n \leq 1$.

Ha $\boldsymbol{A} = (\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \ldots,\ \boldsymbol{a}_n)$, akkor $|\boldsymbol{A}| = 1$, tehát a szimplex térfogata $\frac{1}{n!}$, és ez egyben a keresett valószínűség is. (Könnyen mondhatjuk, hogy az $x_1, x_2, ..., x_n$ számok minden permutációja egyenlő valószínű, tehát egy bizonyos permutáció előfordulásának a valószínűsége $\frac{1}{n!}$, de mivel minden permutáció előfordulásának a valószínűsége nulla, ez a gondolatmenet így nem bizonyító erejű.)

Ha $(x_1, x_2, ..., x_n)$ az $F(x)$ eloszlásfüggvénnyel rendelkező sokaságból választott minta, és F folytonos, akkor $y_i = F^{-1}(x_i)$ transzformációval, mely a monotonitást megtartja, elérhető, hogy a mintaelemek (0, 1)-ben egyenletes eloszlásúak legyenek, tehát a monoton növekedés valószínűsége továbbra is $\frac{1}{n!}$. Ha F nem folytonos, az állítás nem marad érvényben.

A monoton növekedő sorozat hosszát jelöljük a N valószínűségi változóval. Várható értékének a kiszámításához a valószínűségszámítás jól ismert képletét használjuk:

$$E(N+1) = 1 + \sum_{k=1}^{\infty} P(N \geq k) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} = e.$$

10. Határozzuk meg az n-dimenziós egységkocka origóból kiinduló $(n-1)$-dimenziós lapátlói által meghatározott szimplex térfogatát! Szabályos-e ez a szimplex?

A lapátló vektorok egyik koordinátája 0, a többi 1 (1. fejezet 11. Kérdés). Ha $\boldsymbol{a} = (0,\ 1,\ \ldots,\ 1)$ és $\boldsymbol{b} = (1,\ 0,\ 1,\ \ldots,\ 1)$, akkor $\langle \boldsymbol{a},\ \boldsymbol{b}\rangle = n-2$, $||\boldsymbol{a}|| = ||\boldsymbol{b}|| = \sqrt{n-1}$, vagyis $\cos\alpha = \frac{n-2}{n-1}$, ami általában különbözik 0,5-től, és csak $n = 3$ esetén kapunk szabályos szimplexet.

A térfogat kiszámítása a

$$d = \begin{vmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{vmatrix}$$

determináns kiszámításán múlik. Adjuk hozzá az összes többi sorvektort az utolsóhoz, ekkor

$$d = (n-1)\begin{vmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{vmatrix} = (-1)^{n-1}(n-1)$$

adódik. A keresett térfogat tehát $V = \frac{n-1}{n!}$.

11. Mutassuk meg, hogy a koordináta egységvektorok által meghatározott szimplex alaplapja szabályos szimplex.

Válasszuk ki egy csúcsot tetszőlegesen, legyen ez pl. az $\boldsymbol{a}_n$. Az ebből a csúcsból kiinduló, alaplaphoz tartozó élvektorok az 1. fejezet 10. kérdésének a vektorai, tehát bármely kettő 60^o-os szöget zár be és egyenlő normájúak.

Ez a példa a szabályos szimplex leggyakoribb előfordulása.

12. Az $\boldsymbol{Ax} = \boldsymbol{b}$, $|\boldsymbol{A}| \neq 0$ egyenletrendszerrel kapcsolatban mit kapunk, ha az $\boldsymbol{A} = (\boldsymbol{a}_1, \ \boldsymbol{a}_2, \ \ldots, \ \boldsymbol{a}_n)$ jelölést felhasználva kiszámítjuk az $\frac{1}{|\boldsymbol{A}|}|(\boldsymbol{a}_1 - \boldsymbol{a}_2, \boldsymbol{b}, \boldsymbol{a}_3, ..., \boldsymbol{a}_n)|$ kifejezést?

A determinánsok alaptulajdonságait használjuk fel (5.2.6. és 5.2.4. Tulajdonság):

$$\frac{1}{|\boldsymbol{A}|}|(\boldsymbol{a}_1 - \boldsymbol{a}_2, \boldsymbol{b}, \boldsymbol{a}_3, ..., \boldsymbol{a}_n)| =$$

$$= \frac{1}{|A|}(|(\boldsymbol{a}_1, \boldsymbol{b}, \boldsymbol{a}_3, ..., \boldsymbol{a}_n)| + |(-\boldsymbol{a}_2, \boldsymbol{b}, \boldsymbol{a}_3, ..., \boldsymbol{a}_n)|) =$$

$$= \frac{1}{|\boldsymbol{A}|}(|(\boldsymbol{a}_1, \boldsymbol{b}, \boldsymbol{a}_3, ..., \boldsymbol{a}_n)| + |\boldsymbol{b}, \boldsymbol{a}_2, \boldsymbol{a}_3, ..., \boldsymbol{a}_n|) = x_2 + x_1,$$

vagyis a fenti képlettel az egyenletrendszer megoldására vonatkozóan az első két koordináta összegét kapjuk meg. (Ha csak $x_1 + x_2$-re van szükség, akkor egy determináns kiszámítása megtakarítható.)

6. fejezet

Sajátértékek

6.1. Sajátvektor és sajátérték definíciója

A lineáris transzformációk áttekintését, ehhez az alkalmas bázis kiválasztását segíti a sajátvektor meghatározása. A sajátvektor a térnek olyan irányát jelöli ki, amely irányban a lineáris transzformáció egyszerűen skalárral való szorzással valósul meg. Ilyen irány természetesen nem mindig van.

6.1.1. Definíció. Az $\boldsymbol{Ax}$: $\mathbf{R}^n \to \mathbf{R}^n$ lineáris transzformációnak a $\boldsymbol{v} \neq \mathbf{0}$ vektor *sajátvektora*, ha $\boldsymbol{Av} = \lambda\boldsymbol{v}$. A λ számot a $\boldsymbol{v}$ sajátvektorhoz tartozó *sajátértéknek* nevezzük.

A sajátvektor és a sajátérték összetartozó fogalmak, így beszélhetünk a λ sajátértékhez tartozó $\boldsymbol{v}$ sajátvektorról is. Érdekes módon inkább a sajátértékek játsszák az elsődleges szerepet, értékük nem választható tetszőlegesen. A sajátértékek halmazát az $\boldsymbol{A}$ spektrumának is nevezik. Ha $\boldsymbol{v}$ a λ-hoz tartozó sajátvektor, akkor $c\boldsymbol{v}$, ahol $c \neq 0$, is a λ-hoz tartozó sajátvektor.

Ha $\boldsymbol{v}$ sajátvektora $\boldsymbol{A}$-nak λ sajátérték mellett, akkor sajátvektora $\boldsymbol{A}^2$-nek is, hiszen $\boldsymbol{A}^2\boldsymbol{v} = \boldsymbol{A}\lambda\boldsymbol{v} = \lambda^2\boldsymbol{v}$, és a hozzátartozó sajátérték λ^2. Általánosítsuk, és képezzük $\boldsymbol{A}$-nak egy polinomját. $p(x)$ legyen egyváltozós polinom, akkor képezhetjük $\boldsymbol{A}$ hatványaival és ezek lineáris kombinációival a $p(\boldsymbol{A})$ mátrixot, vagy lineáris transzformációt. Nyilván $p(\boldsymbol{A})\boldsymbol{v} = p(\lambda)\boldsymbol{v}$, vagyis $\boldsymbol{v}$ $p(\boldsymbol{A})$-nak is sajátvektora $p(\lambda)$ sajátérték mellett. Ha $\boldsymbol{A}$ inverze létezik, akkor $\boldsymbol{v}$ $\boldsymbol{A}^{-1}$-nek is sajátvektora, ugyanis $\boldsymbol{Av} = \lambda\boldsymbol{v}$-ből $\boldsymbol{v} = \lambda\boldsymbol{A}^{-1}\boldsymbol{v}$, vagy másképpen írva $\boldsymbol{A}^{-1}\boldsymbol{v} = \lambda^{-1}\boldsymbol{v}$, tehát a hozzátartozó sajátérték λ^{-1} (ha $\boldsymbol{A}$ inverze létezik, λ nulla nem lehet!).

A sajátérték és a sajátvektor a lineáris transzformáció tulajdonsága, és mint ilyen a választott bázistól nem függhet. Valóban $\boldsymbol{C}^{-1}\boldsymbol{AC}$-nek és $\boldsymbol{A}$-nak a sajátértékei megegyeznek, a sajátvektorai pedig új koordinátákat kapnak (ld. 4.1.2. és 4.1.4. Tétel). Alkalmazzuk a $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ transzformációt a $\boldsymbol{C}^{-1}\boldsymbol{v}$ vektorra, ahol $\boldsymbol{v}$ az $\boldsymbol{A}$ sajátvektora, akkor

$$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}(\boldsymbol{C}^{-1}\boldsymbol{v}) = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{C}^{-1}\boldsymbol{v},$$

amiből látszik, hogy $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$-nek $\boldsymbol{C}^{-1}\boldsymbol{v}$ a sajátvektora és a hozzátartozó sajátérték marad a korábbi λ.

6.1.2. Tétel. A λ szám akkor és csak akkor sajátértéke $\boldsymbol{A}$-nak, ha gyöke az $|\boldsymbol{A}$ - $\lambda\boldsymbol{I}| = 0$ ún. *karakterisztikus egyenletnek.* Ha λ k-szoros gyöke a karakterisztikus egyenletnek, akkor *k-szoros sajátértékről* beszélünk, vagy azt mondjuk, hogy λ *multiplicitása* (vagy algebrai multiplicitása) k.

Bizonyítás. Adott λ számhoz sajátvektor akkor és csak akkor létezik, ha az $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}$ egyenletnek van nullvektortól különböző megoldása. Az $n \times n$-es homogén lineáris egyenletrendszernek akkor és csak akkor van nemtriviális megoldása, ha az egyenletrendszer mátrixa nem teljesrangú, azaz determinánsa nulla (ld. 3.2. fejezet). □

Mivel a karakterisztikus egyenlet n-edfokú egyenlet, következik, hogy az $n \times n$-es $\boldsymbol{A}$ mátrixnak legfeljebb n sajátértéke van. A sajátvektoroknál a konstansszorosa nem jelent lényegesen különböző sajátvektort, minden sajátértékhez a konstans szorzótól eltekintve legalább egy sajátvektor tartozik. A lényegesen különböző sajátvektorok száma azonban lehet végtelen is.

A gyökök és együtthatók összefüggéséből következik, hogy a karakterisztikus egyenlet konstans tagja, $|\boldsymbol{A}|$ a sajátértékek szorzata (a komplex sajátértékeket is felhasználva és a sajátértékeket multiplicitással véve).

1. Kérdés. Hány lineárisan független sajátvektora van az

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

mátrixnak?

6.1.3. Tétel. A különböző sajátértékekhez tartozó sajátvektorok lineárisan függetlenek.

Bizonyítás. A bizonyítás a vektorok száma szerinti teljes indukcióval történik. Egy vektor, mivel nem nullvektor, mindig lineárisan független rendszert képez. Tegyük fel, hogy a $\lambda_1, \lambda_2, ..., \lambda_{k-1}$ sajátértékekhez tartozó $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_{k-1}$ sajátvektorok lineárisan függetlenek, de $\boldsymbol{v}_k$, a λ_k-hoz tartozó sajátvektor, már függ tőlük, azaz

$$\boldsymbol{v}_k = \sum_{i=1}^{k-1} c_i \boldsymbol{v}_i,$$

ahol a c_1, c_2, ..., c_{k-1} számok nem mind nullák. Alkalmazzuk az $\boldsymbol{A}$ transzformációt, és használjuk fel, hogy sajátértékekről van szó, akkor

$$\lambda_k \boldsymbol{v}_k = \sum_{i=1}^{k-1} c_i \lambda_i \boldsymbol{v}_i.$$

Az első egyenlet λ_k-szorosát vonjuk le a másodikból:

$$0 = \sum_{i=1}^{k-1} c_i (\lambda_i - \lambda_k) \boldsymbol{v}_i.$$

A lineáris függetlenség miatt $c_i(\lambda_i - \lambda_k) = 0$ minden i-re, de a sajátértékek különbözősége miatt ekkor $c_i = 0$ minden i-re, ami ellentmond a róluk tett feltevésnek. □

Diagonális mátrixnak nevezzük azt a mátrixot, melynek csak a főátlójában állhatnak nemnulla elemek. Ha a főátló elemei rendre d_1, d_2, ..., d_n, akkor a diagonális mátrix által létrehozott transzformáció a k-adik bázisvektor irányában d_k-szoros nyújtást (zsugorítást, tükrözést) hoz létre, tehát a transzformáció szerkezete viszonylag egyszerű. Ha egy mátrix a bázis transzformációjával diagonális mátrixszá alakítható, akkor azt mondjuk, hogy *diagonizálható*. Másképpen kifejezve $\boldsymbol{A}$ akkor és csak akkor diagonizálható, ha hasonló egy diagonális mátrixhoz.

6.1.4. Tétel. Az $n \times n$-es $\boldsymbol{A}$ mátrix akkor és csak akkor diagonizálható, ha $\boldsymbol{A}$-nak van n darab lineárisan független sajátvektora.

Bizonyítás. $\Leftarrow$ Ha van n darab lineárisan független sajátvektor nem feltétlenül különböző λ_1, λ_2, ..., λ_n sajátértékkel, akkor válasszuk ezeket a bázis vektoroknak. A transzformáció az új bázisvektorok irányában λ_i-szeres torzulást hoz létre, ami diagonális mátrixszal írható le.

$\Rightarrow$ Ha diagonizálható, azaz $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{D}$, ahol $\boldsymbol{D}$ olyan diagonális mátrix, melynek a főátlójában a d_1, d_2, ..., d_n elemek állnak, akkor $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}\boldsymbol{e}_k = \boldsymbol{D}\boldsymbol{e}_k = d_k\boldsymbol{e}_k$, és ebből $\boldsymbol{A}\boldsymbol{C}\boldsymbol{e}_k = d_k\boldsymbol{C}\boldsymbol{e}_k$, vagyis $\boldsymbol{C}\boldsymbol{e}_k$ $\boldsymbol{A}$ sajátvektora. $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, ..., $\boldsymbol{e}_n$ lineárisan függetlenek, a $\boldsymbol{C}$ teljesrangú mátrix, tehát $\boldsymbol{C}\boldsymbol{e}_1$, $\boldsymbol{C}\boldsymbol{e}_2$, ..., $\boldsymbol{C}\boldsymbol{e}_n$ is lineárisan függetlenek (4.1.1. Tétel). □

A bizonyításból még több derül ki: a diagonális elemei a sajátértékek.

Az 1. Kérdésben szereplő mátrix nem diagonizálható, hiszen nincs három lineárisan független sajátvektora.

2. Kérdés. Az $\boldsymbol{A}$ mátrix sajátvektorai (2, 1, 1), (1, -1, 1) és (1, 1, 2), a hozzátartozó sajátértékek rendre 1, 2, és 1. Meghatározzák-e az adatok $\boldsymbol{A}$-t? Ha igen, számítsuk is ki!

6.2. A komplex sajátérték esete

A karakterisztikus egyenlet megoldásaként gyakran komplex sajátérték adódik. A sajátvektor kiszámításának sincs akadálya, csak éppen komplex számokból álló vektort kapunk. Ezt eddigi valós vektorterünkben értelmezni nem tudjuk, úgy, hogy látszólag semmi értelme nincs evvel foglalkozni. Ez nem így van! Fogjuk fel inkább úgy a tényeket, hogy segédeszközként felhasználhatók valós n-dimenziós terek vizsgálatakor.

3. Kérdés. Igaz-e, hogy az ortogonális mátrixok minden valós vagy komplex sajátértékének az abszolút értéke 1?

Ismerkedjünk meg a komplex számokból álló n-dimenziós tér alapfogalmaival. A teret a komplex számokból álló szám-n-esek alkotják, vagyis n-dimenziós vektorok, amelyeknek koordinátái komplex számok. A norma itt is definiálható, ha $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$, akkor

$$||\boldsymbol{a}|| = \sqrt{|a_1|^2 + |a_2|^2 + ... + |a_n|^2}.$$

A vektorok összeadása, számmal való szorzása ugyanúgy képezhető, a skalárszorzatnál van lényeges eltérés. $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ és $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ skalárszorzata

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \sum_{i=1}^{n} a_i \bar{b}_i,$$

ahol $\bar{b}_i$ a b_i komplex szám konjugáltja, ha $b_i = \alpha_1 + i\alpha_2$, akkor $\bar{b}_i = \alpha_1 - i\alpha_2$. Ennek következtében megszűnik a skalárszorzat kommutatív tulajdonsága és $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \overline{\langle \boldsymbol{b}, \boldsymbol{a} \rangle}$ lép helyébe. Erre azért van szükség, hogy megmaradjon az $\langle \boldsymbol{a}, \boldsymbol{a} \rangle = ||\boldsymbol{a}||^2$ összefüggés.

Ha $\boldsymbol{A}$ az $\mathbf{R}^n$ lineáris transzformációja, és λ ennek komplex sajátértéke, akkor $\bar{\lambda}$ is sajátérték, ugyanis ha fennáll, hogy $\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}$ valamilyen komplex $\boldsymbol{v}$ vektorral, akkor mindkét oldal konjugáltját véve $\boldsymbol{A}\bar{\boldsymbol{v}} = \bar{\lambda}\bar{\boldsymbol{v}}$, ami mutatja, hogy $\bar{\lambda}$-hez a $\bar{\boldsymbol{v}}$ komplex sajátvektor tartozik.

4. Kérdés. Mutassuk meg, hogy az ortogonális mátrixok esetén a különböző sajátértékekhez tartozó sajátvektorok ortogonálisak. Igaz-e ez komplex sajátértékekre, sajátvektorokra is?

Egy adott $\boldsymbol{v}$ valós sajátvektor meghatároz (generál) egy alteret $\mathbf{R}^n$-ben, melynek az a tulajdonsága, hogy ha $\boldsymbol{x}$ az altérhez tartozik, akkor $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}c\boldsymbol{v} = c\lambda\boldsymbol{v}$ szintén az altérhez tartozik. Az ilyen tulajdonságú altereket $\boldsymbol{A}$-ra invariáns altérnek nevezzük.

6.2.1. Definíció. Az L altér $\boldsymbol{A}$-ra nézve *invariáns altér*, vagy röviden $\boldsymbol{A}$ invariáns altere, ha bármely $\boldsymbol{x} \in L$-re $\boldsymbol{A}\boldsymbol{x} \in L$.

A valós sajátértékekhez a sajátvektor mindig meghatároz egy egydimenziós invariáns alteret. Az egydimenziós invariáns altér minden eleme ugyanahhoz a sajátértékhez tartozó sajátvektor.

Komplex sajátérték esetén $\mathbf{R}^n$ ilyen egydimenziós invariáns altere nem adható meg, de a konjugált sajátérték párhoz megadható egy kétdimenziós invariáns altér.

6.2.2. Tétel. Az $\boldsymbol{A}$ mátrix minden λ és $\bar{\lambda}$ komplex sajátérték párjához megadható egy $\boldsymbol{A}$-ra nézve invariáns kétdimenziós altér. Az alteret λ-hoz tartozó komplex sajátvektor valós része és imaginárius része generálja. Ha λ (és így $\bar{\lambda}$ is) megoldása az $x^2 + ax + b = 0$ valós együtthatós egyenletnek, akkor az invariáns altér minden $\boldsymbol{v}$ elemére $(\boldsymbol{A}^2 + a\boldsymbol{A} + b\boldsymbol{I})\boldsymbol{v} = \mathbf{0}$.

Bizonyítás. A λ sajátértékhez meghatározható egy $\boldsymbol{z}$ komplex sajátvektor. Legyen $\boldsymbol{u}_1 = \mathrm{Re}\,\boldsymbol{z}$ és $\boldsymbol{u}_2 = \mathrm{Im}\,\boldsymbol{z}$ (komplex vektor valós és imaginárius részét koordinátánként képezzük). Az $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ vektorpár fogja generálni az invariáns alteret.

$\boldsymbol{u}_1$ és $\boldsymbol{u}_2$ lineárisan függetlenek, mert $\boldsymbol{u}_2 = c\boldsymbol{u}_1$ esetén $\boldsymbol{z} = (1 + ic)\boldsymbol{u}_1$, és $\boldsymbol{z}$-vel együtt $\boldsymbol{u}_1$ is sajátvektor lenne, de komplex sajátértéknek nem lehet valós sajátvektora.

Az $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ vektorok által generált altér invariáns, mert $\lambda = a + ib$ jelölés mellett

$$\boldsymbol{A}\boldsymbol{u}_1 = \boldsymbol{A}\frac{\boldsymbol{z} + \bar{\boldsymbol{z}}}{2} = \frac{1}{2}(\lambda\boldsymbol{z} + \bar{\lambda}\bar{\boldsymbol{z}}) = \mathrm{Re}(\lambda\boldsymbol{z}) =$$

$$= \mathrm{Re}((a + ib)(\boldsymbol{u}_1 + i\boldsymbol{u}_2)) = a\boldsymbol{u}_1 - b\boldsymbol{u}_2,$$

és

$$\boldsymbol{A}\boldsymbol{u}_2 = \boldsymbol{A}\frac{\boldsymbol{z} - \bar{\boldsymbol{z}}}{2i} = \frac{1}{2i}(\lambda\boldsymbol{z} - \bar{\lambda}\bar{\boldsymbol{z}}) = \mathrm{Im}(\lambda\boldsymbol{z}) =$$

$$= \mathrm{Im}((a + ib)(\boldsymbol{u}_1 + i\boldsymbol{u}_2)) = b\boldsymbol{u}_1 + a\boldsymbol{u}_2,$$

vagyis a generáló elemek transzformáltjai továbbra is az altér elemei.

A $\boldsymbol{z}$ komplex sajátvektorra

$$(\boldsymbol{A}^2 + a\boldsymbol{A} + b\boldsymbol{I})\boldsymbol{z} = (\lambda^2 + a\lambda + b)\boldsymbol{z} = 0.$$

Nézzük mindkét oldal valós részét, majd imaginárius részét, és látjuk, hogy $\boldsymbol{u}_1$ és $\boldsymbol{u}_2$ kielégíti az $(\boldsymbol{A}^2 + a\boldsymbol{A} + b\boldsymbol{I})\boldsymbol{v} = \mathbf{0}$ egyenletet. A linearitás miatt viszont tetszőleges lineáris kombinációjuk is kielégíti. $\square$

Azon $\boldsymbol{v}$ vektorok halmaza, melyre $(\boldsymbol{A}^2 + a\boldsymbol{A} + b\boldsymbol{I})\boldsymbol{v} = \mathbf{0}$, vagyis az $\boldsymbol{A}^2 + a\boldsymbol{A} + b\boldsymbol{I}$ magtere, általában lehet bővebb, mint a generált altér.

5. Kérdés. Igazoljuk, hogy az

$$\boldsymbol{A} = \begin{pmatrix} \cos\varphi & -\sin\varphi & \cdots & 0 \\ \sin\varphi & \cos\varphi & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

mátrixszal megadott forgatás (ld. 4.6.) sajátértékei az 1, $e^{i\varphi}$, $e^{-i\varphi}$ számok, ha $n \geq 3$. Igaz-e, hogy a komplex sajátértékekhez tartozó invariáns altér a forgatás „tengelyének" ortogonális kiegészítő altere?

6.3. A többszörös sajátérték esete

Az 1. Kérdés egyszerű példája több szempontból is elgondolkoztató. A mátrixnak egy sajátértéke van, de az háromszoros gyöke a karakterisztikus polinomnak. Sajátvektora azonban szintén csak egy van, tehát nincs remény arra, hogy a sajátvektorokból bázist építsünk. Az a gondolat is hibás, hogy a sajátvektor által generált altér kiegészítő alterét véve ott újra lejátsszuk a sajátvektor megkeresését, mert a kiegészítő altér általában nem lesz invariáns altér.

A feladat adott, nem sajátvektort kell keresni, hanem a sajátértékhez tartozó invariáns alteret.

Adott sajátértékhez mindig van legalább egy sajátvektor. A többszörös sajátérték esetén nem további saját vektorokat keresünk, hiszen az nem biztos hogy létezik, hanem ún. másodlagos sajátvektorokat, ami a sajátvektor fogalmának általánosítása.

6.3.1. Definíció. Adott λ sajátértékhez és $\boldsymbol{v}_1$ sajátvektorhoz az olyan $\boldsymbol{v}_2$ vektort, melyre $\boldsymbol{v}_1$ és $\boldsymbol{v}_2$ lineárisan függetlenek és valamilyen a konstanssal az

$$\boldsymbol{A}\boldsymbol{v}_2 = \lambda\boldsymbol{v}_2 + a\boldsymbol{v}_1$$

összefüggés áll fenn, a λ-hoz tartozó *másodlagos sajátvektornak* nevezzük. További másodlagos sajátvektorok rekurzív módon definiálhatók. Adott a λ-hoz tartozó $\boldsymbol{v}_1$ sajátvektor és legyenek $\boldsymbol{v}_2$, $\boldsymbol{v}_3$, $\ldots$, $\boldsymbol{v}_{k-1}$ a λ sajátértékhez tartozó másodlagos sajátvektorok. Ha $\boldsymbol{v}_k$ lineárisan független a $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\ldots$, $\boldsymbol{v}_{k-1}$ vektoroktól és teljesül az

$$\boldsymbol{A}\boldsymbol{v}_k = \lambda\boldsymbol{v}_k + \boldsymbol{u}_k$$

összefüggés, ahol $\boldsymbol{u}_k$ eleme a $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\ldots$, $\boldsymbol{v}_{k-1}$ vektorok által generált altérnek, akkor $\boldsymbol{v}_k$-t *másodlagos sajátvektornak* nevezzük. Ha az összefüggés

$\boldsymbol{u}_k = \mathbf{0}$-val áll fenn, akkor $\boldsymbol{v}_k$ sajátvektor a korábbi értelemben, vagy megkülönböztetésül $\boldsymbol{v}_k$ *elsődleges* sajátvektor. A $\boldsymbol{v}_1$, $\boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ vektorok rendszerét, ahol $\boldsymbol{v}_1$ mindig elsődleges, a többi első vagy másodlagos λ-hoz tartozó sajátvektor, λ-hoz tartozó *másodlagos sajátvektor-rendszernek* nevezzük.

Mivel $\boldsymbol{v}_k$ definíciója hivatkozik a korábbi $\boldsymbol{v}_i$ vektorokra, mindig másodlagos sajátvektor sorozatról vagy rendszerről kell beszélnünk.

Könnyen látható, hogy a $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\ldots$, $\boldsymbol{v}_k$ a λ-hoz tartozó másodlagos sajátvektor-rendszer által generált altér $\boldsymbol{A}$-ra nézve invariáns, hiszen

$$\boldsymbol{A}\sum_{i=1}^{k} c_i\boldsymbol{v}_i = \sum_{i=1}^{k} c_i(\lambda\boldsymbol{v}_i + \boldsymbol{u}_i),$$

és a jobb oldal az altérbeli elemeknek a lineáris kombinációja.

6.3.2. Tétel. k elemű λ_0-hoz tartozó másodlagos sajátvektor-rendszer akkor és csak akkor létezik, ha λ_0 a karakterisztikus egyenletnek legalább k-szoros gyöke. Ha λ_0 pontosan k-szoros gyök, akkor a másodlagos sajátvektorok rendszere által generált altér az $(\boldsymbol{A} - \lambda_0\boldsymbol{I})^k\boldsymbol{x}$ transzformáció magtere.

Bizonyítás. A bizonyításban többször is felhasználjuk, hogy a bázistranszformáció a karakterisztikus egyenletet nem változtatja meg.

1. $\Rightarrow$ Ha létezik a másodlagos sajátvektor-rendszer, akkor megfelelő a_1, b_1, b_2, c_1, c_2, c_3,… számokkal

$$\boldsymbol{A}\boldsymbol{v}_1 = \lambda_0\boldsymbol{v}_1,$$

$$\boldsymbol{A}\boldsymbol{v}_2 = \lambda_0\boldsymbol{v}_2 + a_1\boldsymbol{v}_1,$$

$$\boldsymbol{A}\boldsymbol{v}_3 = \lambda_0\boldsymbol{v}_3 + b_1\boldsymbol{v}_2 + b_2\boldsymbol{v}_1,$$

$$\ldots .$$

Válasszuk a bázist úgy, hogy az első k bázisvektor $\boldsymbol{v}_1$, $\boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ legyen, akkor a transzformáció mátrixa

$$\left(\begin{array}{cccc:c} \lambda_0 & a_1 & b_2 & \cdots & \\ 0 & \lambda_0 & b_1 & \cdots & \\ 0 & 0 & \lambda_0 & & \boldsymbol{B} \\ & & & \ddots & \\ \hdashline & & \mathbf{0} & & \end{array}\right)$$

hiszen az eredeti básisvektorokat a transzformáció a fenti egyenleteknek megfelelően transzformálja. Ennek a mátrixnak viszont a λ_0 legalább k-szoros sajátértéke.

$\Leftarrow$ $\boldsymbol{v}_2$ *konstrukciója.* Ha λ_0 k-szoros sajátérték, akkor a $\boldsymbol{v}_1$ sajátvektor létezik. Transzformáljuk a bázist oly módon, hogy az első bázisvektor $\boldsymbol{v}_1$ legyen (ezt ortogonális transzformációval is megtehetjük a későbbiek kedvéért). Az $\boldsymbol{A}$ mátrix alakja ebben a bázisban:

$$\boldsymbol{A} = \left(\begin{array}{c:ccc} \lambda_0 & a_1 & b_2 & \cdots \\ \hdashline 0 & & & \\ \vdots & & \boldsymbol{A}_1 & \\ 0 & & & \end{array}\right),$$

ahol $\boldsymbol{A}_1$-ben λ_0 $(k-1)$-szeres sajátérték $(k \geq 2)$. $\boldsymbol{A}_1$-nek tehát λ_0 sajátértéke, a hozzá tartozó $(n-1)$-dimenziós sajátvektort jelöljük $\tilde{\boldsymbol{v}}$-mal, és legyen $\boldsymbol{v}_2 = \begin{pmatrix} x \\ \tilde{\boldsymbol{v}} \end{pmatrix}$

$$\boldsymbol{A}\boldsymbol{v}_2 = \begin{pmatrix} 0 & a_1 & b_2 & \cdots \\ 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 0 \end{pmatrix} \boldsymbol{v}_2 + \begin{pmatrix} \lambda_0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \boldsymbol{A}_1 & \\ 0 & & & \end{pmatrix} \boldsymbol{v}_2 =$$

$$= \begin{pmatrix} c \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda_0 x \\ \lambda_0 \tilde{\boldsymbol{v}} \end{pmatrix} = c\boldsymbol{v}_1 + \lambda_0 \boldsymbol{v}_2,$$

tehát teljesíti a $\boldsymbol{v}_2$-től elvárt egyenletet. A lineáris függetlenség is teljesül, hiszen $\tilde{\boldsymbol{v}} \neq \boldsymbol{0}$.

$\boldsymbol{v}_3$ *konstrukciója.* Transzformáljuk tovább a bázist oly módon, hogy az első bázisvektor $\boldsymbol{v}_1$ maradjon és a második $\boldsymbol{v}_2$ legyen. Az $\boldsymbol{A}$ mátrix alakja ebben a bázisban:

$$\boldsymbol{A} = \left(\begin{array}{cc:cc} \lambda_0 & a_1 & b_2 & \cdots \\ 0 & \lambda_0 & b_1 & \cdots \\ \hdashline 0 & 0 & \boldsymbol{A}_2 & \\ 0 & 0 & & \end{array}\right).$$

$\boldsymbol{A}_2$-ben λ_0 $(k-2)$-szeres sajátérték $(k \geq 3)$. $\boldsymbol{A}_2$-nek tehát λ_0 sajátértéke, a hozzá tartozó $(n-2)$-dimenziós sajátvektort jelöljük $\tilde{\boldsymbol{v}}$-mal, és legyen

$$\boldsymbol{v}_3 = \begin{pmatrix} x \\ y \\ \tilde{\boldsymbol{v}} \end{pmatrix}.$$

$$\boldsymbol{A}\boldsymbol{v}_3 = \begin{pmatrix} 0 & a_1 & b_2 & \cdots & \\ 0 & 0 & b_1 & \cdots & \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & \cdots & 0 & \end{pmatrix} \boldsymbol{v}_3 + \begin{pmatrix} \lambda_0 & 0 & \cdots & 0 \\ 0 & \lambda_0 & 0 & \cdots \\ \vdots & & \boldsymbol{A}_2 & \\ 0 & 0 & & \end{pmatrix} \boldsymbol{v}_3 =$$

$$= \begin{pmatrix} c \\ d \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda_0 x \\ \lambda_0 y \\ \lambda_0 \tilde{\boldsymbol{v}} \end{pmatrix} = c\boldsymbol{v}_1 + d\boldsymbol{v}_2 + \lambda_0 \boldsymbol{v}_3,$$

tehát $\boldsymbol{v}_3$ is teljesíti a tőle elvárt egyenletet, és lineárisan független a $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ rendszertől.

A további másodlagos sajátvektorok konstrukciója ugyanígy történik.

2. Jelöljük M_k-val a $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ által generált alteret, és M-el az $(\boldsymbol{A} - \lambda_0 \boldsymbol{I})^k \boldsymbol{x}$ magterét. Jelöljük továbbá $\boldsymbol{B}$-vel az $\boldsymbol{A} - \lambda_0 \boldsymbol{I}$ mátrixot.

Indukcióval belátjuk, hogy $\boldsymbol{B}^j \boldsymbol{v}_j = \boldsymbol{0}$, $(1 \le j \le k)$. $j = 1$-re $\boldsymbol{B}\boldsymbol{v}_1 = \boldsymbol{0}$ a sajátérték definícióját jelenti. Feltételezzük, hogy $\boldsymbol{B}^i \boldsymbol{v}_i = \boldsymbol{0}$, ha $i < j$, akkor a másodlagos sajátérték definíciója miatt $\boldsymbol{B}\boldsymbol{v}_j = \boldsymbol{u}_j$, ahol $\boldsymbol{u}_j$ a $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\ldots$, $\boldsymbol{v}_{j-1}$ által generált altér eleme. Így $\boldsymbol{B}^j \boldsymbol{v}_j = \boldsymbol{B}^{j-1} \boldsymbol{u}_j$, ahol $\boldsymbol{u}_j$ lineáris kombinációja a $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{j-1}$ vektoroknak és minden tagra alkalmazható az indukciós feltétel.

A bizonyított állításból adódóan $\boldsymbol{B}^k \boldsymbol{v}_j = \boldsymbol{0}$, és, mivel M_k minden eleme a $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ vektorok lineáris kombinációja, adódik, hogy $M_k \subset M$.

A fordított állítás – $M_k \supset M$ – bizonyításához válasszuk meg a bázist úgy, hogy az első k bázisvektor $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ legyen. Ekkor $\boldsymbol{B}$ az

$$\left(\begin{array}{cccc|c} 0 & a_1 & b_2 & \cdots & \\ 0 & 0 & b_1 & \cdots & \\ 0 & 0 & 0 & & \boldsymbol{C} \\ & & & \ddots & \\ \hline & & \boldsymbol{0} & & \boldsymbol{B}_1 \end{array}\right)$$

alakot ölti. Tudjuk, hogy $\boldsymbol{B}$-nek a 0 k-szoros sajátértéke, tehát $\boldsymbol{B}_1$-nek a 0 már nem sajátértéke. Válasszunk egy $\boldsymbol{x} \in M$ elemet, és tegyük fel, hogy $\boldsymbol{x} \notin M_k$. Ebben a bázisban ez azt jelenti, hogy az $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$

vektorból képezett $\tilde{\boldsymbol{x}} = (x_{k+1}, x_{k+2}, ..., x_n) \neq \mathbf{0}$. A $\boldsymbol{B}_1$ teljesrangú mátrix, tehát akkor $\boldsymbol{B}_1\tilde{\boldsymbol{x}} \neq \mathbf{0}$, és $\boldsymbol{B}_1^k\tilde{\boldsymbol{x}} \neq \mathbf{0}$. Ebből $\boldsymbol{B}^k\boldsymbol{x} \neq \mathbf{0}$ következik, ami ellentmond $\boldsymbol{x} \in M$-nek. □

Megjegyzések a tételhez.

1. Az M-et úgy definiáltuk, hogy a $\boldsymbol{B}^k\boldsymbol{x}$ magtere. A $\boldsymbol{B}^N\boldsymbol{x}$ magtere $N > k$ esetén csak bővebb lehetne, de az előző bizonyítás vége $\boldsymbol{B}^N\boldsymbol{x}$ magterére is megismételhető, tehát $\boldsymbol{B}^N\boldsymbol{x}$ magtere is M, ha $N \geq k$.

2. Ha λ_0 k-szoros sajátérték, akkor pontosan k elemből áll a másodlagos sajátértékek teljes rendszere. Azt, hogy k-elemű rendszer létezik, a tétel kimondja, az hogy k-nál több elemű rendszer nem létezik, a tétel második állításából következik. Ha lenne ugyanis $\boldsymbol{v}_{k+1}$ $(k+1)$-edik másodlagos sajátérték, akkor $\boldsymbol{B}\boldsymbol{v}_{k+1} = \sum\limits_{i=1}^{k} c_k\boldsymbol{v}_k$, amiből $\boldsymbol{B}^{k+1}\boldsymbol{v}_{k+1} = \mathbf{0}$. De $\boldsymbol{B}^{k+1}$ magtere nem bővebb mint $\boldsymbol{B}^k$ magtere, ami M_k, így $\boldsymbol{v}_{k+1}$ nem lehet lineárisan független a $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, …, $\boldsymbol{v}_k$ vektoroktól.

3. A másodlagos sajátértékek rendszere még konstans szorzótól eltekintve sem egyértelmű. Jelöljük M_j-vel a $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, …, $\boldsymbol{v}_j$ vektorok által generált alteret $(0 < j \leq k)$, és legyen $M_0 = \{\mathbf{0}\}$. Bármely j-re $\boldsymbol{v}_j$ felcserélhető az $M_j - M_{j-1}$ halmaz bármely $\boldsymbol{w}_j$ elemével. Mivel $\boldsymbol{w}_j = \sum\limits_{i=1}^{j} c_i\boldsymbol{v}_i$, tehát $\boldsymbol{B}\boldsymbol{w}_j = \sum\limits_{i=1}^{j} c_i\boldsymbol{B}\boldsymbol{v}_i = \sum\limits_{i=2}^{j} c_i\boldsymbol{u}_i \in M_{j-1}$, továbbá $\boldsymbol{w}_j$ lineárisan független a $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, …, $\boldsymbol{v}_{j-1}$ vektoroktól, tehát $\boldsymbol{w}_j$ másodlagos sajátvektor.

4. A $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, …, $\boldsymbol{v}_k$ vektorok felhasználásával (a Gram–Schmidt-ortogonalizáció segítségével, 1.4.4. Tétel) ortogonális és 1-re normált $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, …, $\boldsymbol{w}_k$ vektorrendszer is készíthető. Az előző megjegyzés alapján ez is másodlagos sajátvektor-rendszer lesz.

5. Az M_j altér $1 \leq j < k$ esetén szintén nem biztos, hogy egyértelmű. Az M_k altér azonban a tétel állítása szerint egyértelműen meghatározott, mivel adott transzformáció magtere.

6. Az M_k altér $\boldsymbol{A}$-ra invariáns. Legyen $\boldsymbol{x} \in M_k$, akkor $(\boldsymbol{A} - \lambda\boldsymbol{I})^k\boldsymbol{x} = \mathbf{0}$, ezért $\boldsymbol{A}(\boldsymbol{A} - \lambda\boldsymbol{I})^k\boldsymbol{x} = (\boldsymbol{A} - \lambda\boldsymbol{I})^k\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$, vagyis $\boldsymbol{A}\boldsymbol{x} \in M_k$.

6.3.3. Tétel. Ha $\boldsymbol{A}$ minden sajátértéke valós, akkor $\boldsymbol{A}$ ortogonális transzformációval felső háromszögmátrixszá alakítható, ahol a főátló alatt csupa nulla áll. A főátlóban a sajátértékek sorakoznak, minden sajátérték annyiszor szerepel, amennyi a multiplicitása.

Bizonyítás. Válasszuk ki a λ_1 sajátértéket, és a 3. megjegyzés alapján vegyük fel a hozzátartozó ortogonális és normált másodlagos maximális sajátvektor-rendszert. Transzformáljuk a bázist oly módon, hogy az új bázis első koordinátái ezek a vektorok legyenek. Ez ortogonális transzformációval is megtehető, ha a sajátvektor-rendszert ortogonálisnak választjuk. $\boldsymbol{A}$ transzformáció mátrixa ebben a koordináta-rendszerben az

$$\left(\begin{array}{cc|cc} \lambda_1 & a_1 & \cdots & \cdots \\ 0 & \lambda_1 & \cdots & \cdots \\ \hline 0 & 0 & & \\ 0 & 0 & & \boldsymbol{A}_1 \end{array}\right)$$

alakot ölti. Az $\boldsymbol{A}_1$ mátrix örökli $\boldsymbol{A}$ λ_1-től különböző sajátértékeit. Folytassuk az eljárást a λ_2 sajátértékkel megtartva az előző sajátértékhez tartozó bázisvektorokat, akkor

$$\left(\begin{array}{cccc|c} \lambda_1 & a_1 & a_2 & \cdots & \\ 0 & \ddots & b_1 & \cdots & \boldsymbol{C} \\ 0 & \cdots & \lambda_2 & & \\ 0 & 0 & 0 & \ddots & \\ \hline & \mathbf{0} & & & \boldsymbol{A}_2 \end{array}\right)$$

alakú lesz a mátrixunk. A további folytatás evidens, és elvezet a tételben jelzett célhoz. □

6. Kérdés. Ha a mátrix nemnegatív elemekből áll, és minden sorában az elemek összege 1, akkor sztochasztikus mátrixnak nevezik (az ún. átmenet valószínűségek mátrixa). Igaz-e, hogy az $n \times n$-es $\boldsymbol{A}$ sztochasztikus mátrixnak $\lambda = 1$ mindig sajátértéke? Azt mondjuk, hogy az $\boldsymbol{Ax} = \boldsymbol{x}$ egyenlenek $\boldsymbol{x}$ sztochasztikus megoldása, ha $\boldsymbol{x}$ minden koordinátája nemnegatív és a koordináták összege 1. Igaz-e, hogy ha $\lambda = 1$ egyszeres sajátérték, akkor az $\boldsymbol{Ax} = \boldsymbol{x}$ egyenletnek csak triviális sztochasztikus megoldása van, azaz $\boldsymbol{x} = \frac{1}{n}\mathbf{1}$.

6.4. Szimmetrikus mátrixok kanonikus alakja

A négyzetes $\boldsymbol{A}$ mátrix szimmetrikus, ha $\boldsymbol{A}^* = \boldsymbol{A}$. Szimmetrikus mátrixokra az elmondottak jelentősen egyszerűbb alakot öltenek, és váratlan megállapítások tehetők.

Szimmetrikus mátrixok összege, lineáris kombinációja nyilván továbbra is szimmetrikus, két szimmetrikus mátrix szorzata azonban nem. Szimmetrikus mátrix marad egy szimmetrikus mátrix ortogonális transzformáltja, ugyanis

$$(\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C})^* = \boldsymbol{C}^*\boldsymbol{A}^*(\boldsymbol{C}^{-1})^* = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}.$$

6.4.1. Tétel. Szimmetrikus mátrix minden sajátértéke valós.

Bizonyítás. Legyen $\boldsymbol{x}$ a λ sajátértékhez tartozó (esetleg komplex) sajátvektor, akkor

$$\lambda||\boldsymbol{x}||^2 = \lambda\langle\boldsymbol{x},\boldsymbol{x}\rangle = \langle\lambda\boldsymbol{x},\boldsymbol{x}\rangle = \langle\boldsymbol{A}\boldsymbol{x},\boldsymbol{x}\rangle = \langle\boldsymbol{x},\boldsymbol{A}^*\boldsymbol{x}\rangle = \langle\boldsymbol{x},\boldsymbol{A}\boldsymbol{x}\rangle =$$

$$= \langle\boldsymbol{x},\lambda\boldsymbol{x}\rangle = \bar{\lambda}\langle\boldsymbol{x},\boldsymbol{x}\rangle = \bar{\lambda}||\boldsymbol{x}||^2,$$

amiből $\lambda = \bar{\lambda}$, vagyis λ valós. □

6.4.2. Tétel. Különböző sajátértékekhez tartozó sajátvektorok ortogonálisak.

Bizonyítás. Legyen λ és μ két különböző sajátérték, $\boldsymbol{v}$ és $\boldsymbol{u}$ a hozzájuk tartozó sajátvektorok, akkor a szimmetria miatt $\langle\boldsymbol{A}\boldsymbol{v},\ \boldsymbol{u}\rangle = \langle\boldsymbol{v},\ \boldsymbol{A}\boldsymbol{u}\rangle$, vagyis $\lambda\langle\boldsymbol{u},\ \boldsymbol{v}\rangle = \mu\langle\boldsymbol{u},\ \boldsymbol{v}\rangle$, amiből $(\lambda-\mu)\langle\boldsymbol{u},\ \boldsymbol{v}\rangle = 0$. Mivel $\lambda \neq \mu$, következik, hogy $\langle\boldsymbol{u},\ \boldsymbol{v}\rangle = 0$. □

6.4.3. Tétel. Szimmetrikus mátrix és csak a szimmetrikus mátrix az, amelyik ortogonális transzformációval diagonizálható.

Bizonyítás. $\Rightarrow$ A 6.3.3. Tétel miatt ortogonális transzformációval felső háromszög mátrixszá alakítható, de a transzformálással szimmetrikus marad, tehát automatikusan diagonális mátrix lesz.

$\Leftarrow$ A diagonális mátrix szimmetrikus, és ortogonális transzformáltja is az marad. □

A 6.4.3. Tétel következményeként jegyezzük meg, hogy szimmetrikus mátrix esetén van olyan ortogonális bázis, melynek elemei sajátvektorok. A diagonizálást létrehozó mátrix oszlopvektorai ilyen bázisvektorok.

7. Kérdés. Diagonizálható-e az $\begin{pmatrix} 5 & -1 & -2 \\ 4 & 0 & -2 \\ 4 & -1 & -1 \end{pmatrix}$ mátrix? Ha igen, számoljuk is ki (jól használhatjuk az Excelt, az 5.7.-ben leírtak szerint)! Ortogonális transzformációval diagonizálható-e?

8. Kérdés. Igaz-e, hogy szimmetrikus mátrix csak ortogonális mátrixszal diagonizálható?

6.5. Direkt összeg felbontás

Láttuk, hogy a különböző sajátértékekhez tartozó sajátvektorok lineárisan függetlenek (6.1.3. Tétel). Igaz-e hasonló állítás a másodlagos sajátvektorokra? Tudunk-e bázist építeni másodlagos sajátvektorokból? Ezekre a kérdésekre keressük a választ az alábbiakban.

Ahhoz, hogy a szóhasználatot megkönnyítsük, bevezetünk két fogalmat.

6.5.1. Definíció. *Diszjunktnak* nevezzük két *alteret*, ha egyetlen közös elemeük a $\mathbf{0}$. A λ sajátértékhez tartozó másodlagos sajátérték-rendszer által generált alteret *kanonikus altérnek* fogjuk nevezni.

Mint a 6.3.2. Tételben láttuk, a λ-hoz tartozó kanonikus altér az $(\boldsymbol{A} - \lambda_0\boldsymbol{I})^k\boldsymbol{x}$ lineáris transzformáció magtere, ahol k a λ multiplicitása.

6.5.2. Segédtétel. Legyenek $\lambda_1, \lambda_2, \ldots, \lambda_k$ az $\boldsymbol{A}$ különböző saját-értékei, $\boldsymbol{v}$ pedig a λ_k-hoz tartozó sajátvektor. Ekkor $\boldsymbol{v}$ nem lehet eleme a $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ értékekhez tartozó kanonikus alterek által generált altérnek. Másképpen megfogalmazva, ha a $\boldsymbol{B}_i = \boldsymbol{A} - \lambda_i\boldsymbol{I}$ jelölés mellett $\boldsymbol{C} = \boldsymbol{B}_1\boldsymbol{B}_2\ldots\boldsymbol{B}_{k-1}$, akkor $\boldsymbol{v}$ semmilyen p-re nem eleme a $\boldsymbol{C}^p$ magterének.

Bizonyítás. A második állításból az első már következik. Ha p értéke nagyobb, mint a maximális multiplicitás, akkor minden másodlagos sajátvektor benne van $\boldsymbol{C}^p$ magterében, így lineáris kombinációjuk is, tehát a generált altér része a magtérnek. Elég tehát a második állítást bizonyítani.

Vegyük észre, hogy a $\boldsymbol{B}_i$ mátrixok szorzata felcserélhető, és számoljuk ki $\boldsymbol{C}^p\boldsymbol{v}$-t!

$$\boldsymbol{C}\boldsymbol{v} = \boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-2}(\boldsymbol{A} - \lambda_{k-1}\boldsymbol{I})\boldsymbol{v} = \boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-2}(\lambda_k - \lambda_{k-1})\boldsymbol{v} = ... =$$

$$= \boldsymbol{v}\prod_{i=1}^{k-1}(\lambda_k - \lambda_i) = \alpha\,\boldsymbol{v},$$

ahol a sajátértékek különbözősége miatt $\alpha \neq 0$. Így bármely p-re $\boldsymbol{C}^p\boldsymbol{v} = \alpha^{\,p}\boldsymbol{v} \neq \mathbf{0}$. $\square$

6.5.3. Tétel. Legyenek $\lambda_1, \lambda_2, \ldots, \lambda_k$ az $\boldsymbol{A}$ különböző sajátértékei. A $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ értékekhez tartozó kanonikus alterek által generált altér és a λ_k-hoz tartozó kanonikus altér diszjunktak.

Bizonyítás. Be fogjuk bizonyítani, hogy $(\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-1})^p$ és $\boldsymbol{B}_k^p$ magterei diszjunktak minden p-re. Ez az állítás a tétel állítását tartalmazza.

A bizonyítást p szerinti teljes indukcióval végezzük el. $p = 1$-re, ha valamely $\boldsymbol{x} \neq \mathbf{0}$-ra $\boldsymbol{B}_k\boldsymbol{x} = \mathbf{0}$, akkor $\boldsymbol{x}$ λ_k-hoz tartozó sajátérték, és az előző segédtétel szerint nem lehet $\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-1}$ magterének eleme. Tegyük fel most, hogy $p-1$-re igaz az állítás, azaz ha valamilyen $\boldsymbol{x}$-re $(\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-1})^{p-1}\boldsymbol{x} = \mathbf{0}$ és $\boldsymbol{B}_k^{p-1}\boldsymbol{x} = \mathbf{0}$, akkor tudjuk, hogy $\boldsymbol{x} = \mathbf{0}$, és bebizonyítjuk az állítást p-re, azaz, hogy a $(\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-1})^p\boldsymbol{x} = \mathbf{0}$ és a $\boldsymbol{B}_k^p\boldsymbol{x} = \mathbf{0}$ összefüggésekből következik, hogy $\boldsymbol{x} = \mathbf{0}$.

Ha $(\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-1})^p\boldsymbol{x} = \mathbf{0}$, akkor $(\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-1})^{p-1}\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_k\boldsymbol{x} = \mathbf{0}$, továbbá, ha $\boldsymbol{B}_k^p\boldsymbol{x} = \mathbf{0}$, akkor $\boldsymbol{B}_k^{p-1}\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_k\boldsymbol{x} = \mathbf{0}$. Az indukciós feltétel miatt tehát $\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_k\boldsymbol{x} = \mathbf{0}$.

Mivel $\boldsymbol{B}_k^p\boldsymbol{x} = \mathbf{0}$, keressük meg azt a legkisebb q értéket, melyre $\boldsymbol{B}_k^q\boldsymbol{x} = \mathbf{0}$. Feltehető, hogy $q \geq 2$, mert $q = 0$ azt jelenti, hogy $\boldsymbol{x} = 0$, és ezt akarjuk bizonyítani, $q = 1$ pedig azt, hogy $\boldsymbol{x}$ λ_k-hoz tartozó sajátvektor, ami a segédtétel miatt nem lehet $(\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-1})^p$ magterében, csak, ha $\boldsymbol{x} = 0$. Legyen $\boldsymbol{v} = \boldsymbol{B}_k^{q-1}\boldsymbol{x}$, és mivel $\boldsymbol{B}_k\boldsymbol{v} = \mathbf{0}$, $\boldsymbol{v}$ a λ_k-hoz tartozó sajátvektor. A $\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_k\boldsymbol{x} = \mathbf{0}$ összefüggés alapján $\boldsymbol{B}_1\boldsymbol{B}_2...\boldsymbol{B}_{k-1}\boldsymbol{v} = \mathbf{0}$, ami ellentmond a segédtételnek. □

6.5.4. Következmény. Az egyes $\lambda_1, \lambda_2, \ldots, \lambda_k$ különböző sajátértékekhez tartozó kanonikus alterekben kiválasztva a másodlagos sajátvektorok egy-egy rendszerét, összességükben lineárisan független vektorrendszert alkotnak.

Bizonyítás. k szerinti teljes indukcióval bizonyítjuk. $k = 1$ esetén a másodlagos sajátvektorokat úgy készítettük el, hogy lineárisan függetlenségük teljesüljön. Legyenek a λ_k-hoz tartozó kiválasztott másodlagos sajátvektorok $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_l$, a többiek pedig $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m$. A

$$\sum_{i=1}^{l} c_i\boldsymbol{v}_i = \sum_{i=1}^{m} d_i\boldsymbol{w}_i$$

összefüggés azt jelenti, hogy a λ_k-hoz tartozó kanonikus altérnek és a többi sajátértékhez tartozó alterek által generált altérnek van közös pontja, de – 6.5.2. miatt – ez csak a $\mathbf{0}$ lehet. Mivel $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_l$ lineárisan függetlenek, $c_1 = c_2 = \ldots = c_l = 0$. Az indukciós feltétel szerint a $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m$ vektorok is lineárisan függetlenek, tehát $d_1 = d_2 = \ldots = d_m = 0$. Mindez pedig a teljes vektorrendszer lineáris függetlenségét jelenti. □

6.5.5. Definíció. Az L altér az $L_1, L_2, \ldots, L_k$ *alterek direkt összege*, ha bármely $\boldsymbol{x} \in L$ egyértelműen írható fel $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 + \ldots + \boldsymbol{x}_k$ ($\boldsymbol{x}_i \in L_i$ minden i-re) alakban. Jelölése:

$$L = L_1 \oplus L_2 \oplus \ldots \oplus L_k.$$

Láttuk, hogy tetszőleges L altérrel $\mathbf{R}^n = L \oplus \bar{L}$, ahol $\bar{L}$ az L kiegészítő altere. $k = 2$ esetén $L = L_1 \oplus L_2$ akkor és csak akkor teljesül, ha L_1 és L_2 diszjunkt alterek és generálják L-et. Ha $k > 2$, akkor (a generálás mellett) a páronként diszjunkt tulajdonság nem elég.

9. Kérdés. $\mathbf{R}^3$-ban adható-e példa arra, hogy diszjunkt, generáló alterek direkt összegként nem állítják elő $\mathbf{R}^3$-at?

6.5.6. Tétel. Az L alteret generáló $L_1, L_2, \ldots, L_k$ alterekre az $L = L_1 \oplus L_2 \oplus \ldots \oplus L_k$ direkt összeg előállítás akkor és csak akkor áll fenn, ha a tetszőlegesen kiválasztott $\boldsymbol{x}_i \in L_i$, $\boldsymbol{x}_i \neq \mathbf{0}$ $(i = 1, 2, \ldots, k)$ elemek lineárisan függetlenek.

Bizonyítás. $\Rightarrow$ Tegyük fel, hogy a kiválasztott $\boldsymbol{x}_i \in L_i$, $i = 1, 2, ..., k$ elemek nem lineárisan függetlenek, azaz $\sum\limits_{i=1}^{k} c_i \boldsymbol{x}_i = \mathbf{0}$, úgy, hogy nem minden együttható nulla. Legyen $\boldsymbol{x} = \sum\limits_{i=1}^{k} \boldsymbol{x}_i$, akkor $\boldsymbol{x}$ előállítása nem egyértelmű, mert $\boldsymbol{x} = \sum\limits_{i=1}^{k} (\boldsymbol{x}_i + c_i \boldsymbol{x}_i)$ az előzőtől különböző előállítás.

$\Leftarrow$ Ha valamely $\boldsymbol{x} \in L$ előállítása nem lenne egyértelmű, azaz $\boldsymbol{x} = \sum\limits_{i=1}^{k} \boldsymbol{x}_i$ és $\boldsymbol{x} = \sum\limits_{i=1}^{k} \boldsymbol{x}'_i$ egyaránt teljesülne, akkor $\sum\limits_{i=1}^{k} (\boldsymbol{x}_i - \boldsymbol{x}'_i) = \mathbf{0}$ ellentmondana az $\boldsymbol{x}_1 - \boldsymbol{x}'_1, \boldsymbol{x}_2 - \boldsymbol{x}'_2, ..., \boldsymbol{x}_k - \boldsymbol{x}'_k$ elemek lineáris függetlenségének. □

Tegyük fel a paragrafus további részében, hogy az $\boldsymbol{A}$ lineáris transzformáció minden sajátértéke valós. λ_1, λ_2, $\ldots \lambda_k$ jelölje a különböző sajátértékeket, és M_1, M_2, $\ldots$, M_k legyenek az egyes sajátértékekhez tartozó kanonikus alterek.

6.5.7. Tétel. Ha $\boldsymbol{A}$ minden sajátértéke valós, akkor

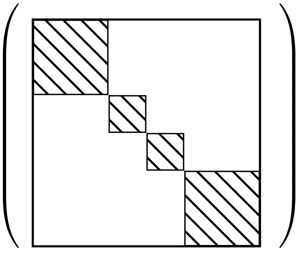
$$\mathbf{R}^n = M_1 \oplus M_2 \oplus \ldots \oplus M_k.$$

Bizonyítás. Minden egyes kanonikus altérben felvehető a sajátérték multiplicitásával egyező számú másodlagos sajátvektor-rendszer. Mivel a multiplicitások összege n, a kiválasztott sajátvektorok n elemű, lineárisan független vektorrendszert alkotnak (lásd 6.5.4. Következmény), ami egyben bázis, tehát generálja $\mathbf{R}^n$-et.

Megmutatjuk, hogy az alterekből tetszőlegesen kiválasztott $\boldsymbol{x}_i \in L_i$, $i = 1, 2, ..., k$ elemek lineárisan függetlenek. Ha a lineáris függetlenség nem teljesülne, és fennállna a $\sum\limits_{i=1}^{k} c_i \boldsymbol{x}_i = 0$ lineáris összefüggés, akkor minden $\boldsymbol{x}_i$ az M_i-n belül lineárisan kifejezhető M_i-ban lévő másodlagos sajátvektorokkal, és a másodlagos sajátvektorok között is létezne egy lineáris összefüggés, ami lehetetlen. □

Minden kanonikus M_i altérben válasszunk egy bázist: $\boldsymbol{u}_{i1}, \boldsymbol{u}_{i2}, ..., \boldsymbol{u}_{in_i}$, a vektorok összessége n elemű lesz, hiszen $n_1 + n_2 + ... + n_k = n$, és lineárisan független vektorokból áll. Tegyük fel ugyanis, hogy a kiválasztott vektorokra fennáll, hogy $\sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} c_{ij} \boldsymbol{u}_{ij} = \mathbf{0}$, akkor a $\boldsymbol{z}_i = \sum\limits_{j=1}^{n_i} c_{ij} \boldsymbol{u}_{ij} \in M_i$ részösszegek, a 6.5.6. Tétel miatt, ha nem $\mathbf{0}$-k, akkor lineárisan függetlenek. A részletösszegek összege azonban $\mathbf{0}$, így lineárisan függetlenek nem lehetnek, tehát minden ilyen részletösszeg $\mathbf{0}$. Minden részletösszeg M_i-beli lineárisan független vektorok lineáris kombinációja, ami csak csupa nulla együtthatóval lehet nulla. Tehát minden c_{ij}nulla, vagyis a vektorrendszer lineárisan független. A kiválasztott vektorrendszer tehát $\mathbf{R}^n$ bázisát képezi.

Rendezzük a báziselemeket $\boldsymbol{u}_{11}, \boldsymbol{u}_{12}, ..., \boldsymbol{u}_{1n_1}, \boldsymbol{u}_{21}, ..., \boldsymbol{u}_{2n_2}, ..., \boldsymbol{u}_{k1}, \boldsymbol{u}_{k2}, ..., \boldsymbol{u}_{kn_k}$ formában, és írjuk fel az $\boldsymbol{A}$ transzformáció mátrixát a most előállított bázisra vonatkozóan. Az $\boldsymbol{A}$ mátrix oszlopvektorait az $\boldsymbol{u}_{ij}$ bázisvektorok transzformáltjainak a koordinátái képezik. Az M_i altér $\boldsymbol{A}$-ra invariás altér (6.3.2. Tétel, 5. Megjegyzés), tehát $\boldsymbol{A}\boldsymbol{u}_{ij}$ az $\boldsymbol{u}_{i1}, \boldsymbol{u}_{i2}, ..., \boldsymbol{u}_{in_i}$ báziselemek lineáris kombinációja, ezekhez a bázisvektorokhoz tartozó koordinátái tetszőleges valós számok lehetnek, míg a többi koordinátája nulla. Az $\boldsymbol{A}$ mátrixa erre a bázisra vonatkozóan ún. kvázidiagonális alakot ölt, amit alább szemléltetünk.

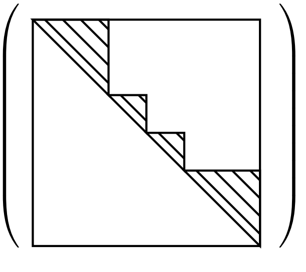


A vonalkázott részeken lehet nem nulla elem, a többi elem nulla.

Válasszuk meg speciálisan a bázisunkat: minden M_i altérben a másodlagos sajátértékek egy rendszerét vegyük fel. A kapott rendszert jelöljük most a

$$\boldsymbol{v}_{11}, \boldsymbol{v}_{12}, ..., \boldsymbol{v}_{1n_1}, \boldsymbol{v}_{21}, ..., \boldsymbol{v}_{2n_2}, ..., \boldsymbol{v}_{k1}, \boldsymbol{v}_{k2}, ..., \boldsymbol{v}_{kn_k}$$

jelekkel. Mivel $\boldsymbol{A}\boldsymbol{v}_{ij}$ most a $\boldsymbol{v}_{i1}, \boldsymbol{v}_{i2}, ..., \boldsymbol{v}_{ij}$ báziselemek lineáris kombinációja, az $\boldsymbol{A}$ mátrix alakja ebben a bázisban kvázidiagonál háromszögmátrix lesz, az alakját az előzőhöz hasonlóan szemléltetjük:



6.6. Jordan-alak

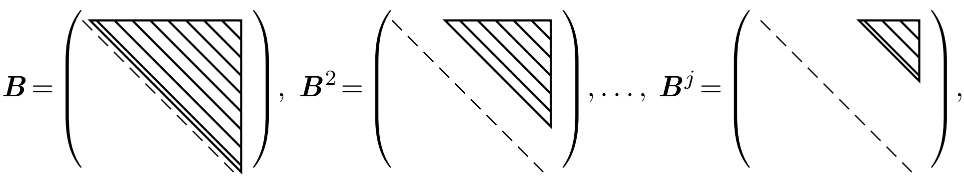
Szándékunk továbbra is az $\boldsymbol{A}$ lineáris transzformáció struktúrájának a felderítése. Ebben a részben is feltételezzük, hogy $\boldsymbol{A}$ minden sajátértéke valós. Az előző paragrafusban láttuk, hogy a tér felbontható a különböző sajátértékekhez tartozó kanonikus alterek direkt összegére. Ez azt jelenti, hogy a transzformáció tulajdonságait kisebb dimenziós alterek vizsgálatára redukálhatjuk, ráadásul ezen az altéren a transzformációnak egyetlen sajátértéke van, melynek multiplicitása megegyezik az altér dimenziójával.

Ez a gondolatmenet vezet minket oda, hogy feltételezzük, hogy λ az $\mathbf{R}^n$-en értelmezett $\boldsymbol{A}$ transzformációnak n-szeres sajátértéke. Vezessük be a $\boldsymbol{B} = \boldsymbol{A} - \lambda\boldsymbol{I}$ transzformációt, melynek a 0 n-szeres sajátértéke.

6.6.1. Tétel. A 0 a $\boldsymbol{B}$-nek akkor és csak akkor n-szeres sajátértéke, ha $\boldsymbol{B}^n = \mathbf{0}$.

Bizonyítás. $\Rightarrow$Válasszuk meg a tér bázisát oly módon, hogy $\boldsymbol{B}$ mátrixa felső háromszögmátrix legyen (6.3.3. Tétel). Ebben a mátrixban a főátló alatt és a főátlóban minden elem nulla.

Indukcióval bizonyítjuk, hogy a $\boldsymbol{B}^k = (c_{ij})$ mátrixban minden olyan c_{ij} elem nulla, melyre $j - i \le k - 1$. Szemléltetve a $\boldsymbol{B}$, $\boldsymbol{B}^2$, ..., $\boldsymbol{B}^j$ mátrixok sorozata az alábbi módon néz ki:



vagyis a nullától különböző elemek (vonalkázva jelölve) „eltávoznak” $\boldsymbol{B}^j$-ből. A $\boldsymbol{B}^{k+1} = (d_{ij})$ mátrix d_{ij} eleme a $\boldsymbol{B}^k$ mátrix i-edik sorának és a $\boldsymbol{B}$ mátrix j-edik oszlopának a skalárszorzata. $\boldsymbol{B}^k$ mátrix i-edik sorának az első $i+k-$ -1 eleme nulla, a $\boldsymbol{B}$ mátrix j-edik oszlopának minden eleme, az első $j-1$ kivételével, szintén nulla. A skalárszorzat nulla lesz, ha $j - 1 \le i + k - 1$, vagyis ha $j - i \le k$, és ezt akartuk bizonyítani.

Ha $k = n$, akkor minden szóbajövő $(i,\ j)$ számpárra teljesül, hogy $j - i \le n - 1$, tehát a $\boldsymbol{B}^n$ minden eleme nulla. Ha $\boldsymbol{B}^n = \mathbf{0}$ valamilyen bázis mellett, akkor $\boldsymbol{B}^n$ minden bázisban $\mathbf{0}$ mátrix.

$\Leftarrow$ Ha $\boldsymbol{B}$-nek van nem nulla (esetleg komplex) λ sajátértéke, akkor $\boldsymbol{B}\boldsymbol{z} = \lambda\boldsymbol{z}$, valamilyen, esetleg komplex $\boldsymbol{z}$ vektorra. Ekkor $\boldsymbol{B}^n z = \lambda^n \boldsymbol{z} \ne$ $\ne \mathbf{0}$, ami ellentmond $\boldsymbol{B}^n = \mathbf{0}$-nak. $\square$

Definíció 6.6.2. Azt a mátrixot, melyre $\boldsymbol{B}^n = \mathbf{0}$ *nilpotens* mátrixnak nevezzük. Ha $\boldsymbol{B}^n = \boldsymbol{0}$, akkor azt a legkisebb k számot, melyre $\boldsymbol{B}^k = \boldsymbol{0}$, a $\boldsymbol{B}$ *indexének* nevezzük. Azt a legkisebb l számot, melyre $\boldsymbol{B}^l\boldsymbol{x} = \boldsymbol{0}$, az $\boldsymbol{x}$ *indexének* nevezzük. Ha $\boldsymbol{x} \neq \boldsymbol{0}$, akkor az $\boldsymbol{x}$-hez hozzárendelhetünk egy sajátvektort, $\boldsymbol{v} = \boldsymbol{B}^{l-1}\boldsymbol{x}$-et (ugyanis $\boldsymbol{Bv} = \boldsymbol{0}$), ezt a hozzárendelt sajátvektort $\boldsymbol{x}$ *gyökerének* nevezzük.

10. Kérdés. Igaz-e, hogy ha $\boldsymbol{v}$ adott sajátvektor, akkor azon $\boldsymbol{x}$ elemek halmaza, melyek gyökere $\boldsymbol{v}$, a $\boldsymbol{0}$-ral kiegészítve $\boldsymbol{B}$-re invariáns alteret ad?

Végezzük el a következő konstrukciót. Az $\boldsymbol{x} \in \mathbf{R}^n$ elemek közül keressük meg a legnagyobb indexűt (pontosabban válasszunk ki a legnagyobb indexűek közül egyet). Legyen ez $\boldsymbol{a}_1$, indexét pedig jelöljük $(p_1 + 1)$-gyel. Válasszuk ki a másodlagos sajátvektorok egy sorozatát a következőképpen: $\boldsymbol{w}_{10} = \boldsymbol{v} = \boldsymbol{B}^{p_1}\boldsymbol{a}_{1,}\ \boldsymbol{w}_{11} = \boldsymbol{B}^{p_1-1}\boldsymbol{a}_{1,} ..., \boldsymbol{w}_{1p_1} = \boldsymbol{a}_1$. $\boldsymbol{v}$ nyilván $\boldsymbol{a}_1$ gyökere. Ha van $\boldsymbol{v}$-től lineárisan független más sajátvektor, akkor folytassuk a másodlagos sajátvektorok konstrukcióját az alább leírtak szerint.

Tegyük fel, hogy kiválasztottuk már a $\boldsymbol{w}_{10}, \boldsymbol{w}_{11}, ..., \boldsymbol{w}_{1p_1}, \boldsymbol{w}_{20}, ..., \boldsymbol{w}_{2p_2}, ...,$ $\boldsymbol{w}_{i-10}, ..., \boldsymbol{w}_{i-1p_{i-1}}$ rendszert, és $\boldsymbol{w}_{10}, \boldsymbol{w}_{20}, ..., \boldsymbol{w}_{i-10}$ nem generálja az összes sajátvektor által képezett alteret, akkor kiválaszthatunk maximális indexű és a $\boldsymbol{w}_{10}, \boldsymbol{w}_{20}, ..., \boldsymbol{w}_{i-10}$ vektoroktól lineárisan független gyökerű $\boldsymbol{a}_i$ vektort. Ezzel elkészíthetjük a $\boldsymbol{w}_{i0} = \boldsymbol{v} = \boldsymbol{B}^{p_i}\boldsymbol{a}_i$, $\boldsymbol{w}_{i1} = \boldsymbol{B}^{p_i-1}\boldsymbol{a}_{i,} ..., \boldsymbol{w}_{ip_i} =$ $= \boldsymbol{a}_i$ másodlagos sajátvektorokból álló folytatást. Az eljárás addig folytatható, amíg van gyökérként választható lineárisan független sajátvektor.

Ezzel az eljárással megkonstruálható a másodlagos sajátvektoroknak a

$$\boldsymbol{w}_{10}, \boldsymbol{w}_{11}, ..., \boldsymbol{w}_{1p_1},\ \boldsymbol{w}_{20}, ..., \boldsymbol{w}_{2p_2},\ ..., \boldsymbol{w}_{k0}, ..., \boldsymbol{w}_{kp_k}$$

rendszere, ahol k a sajátvektorok alterének a dimenziója. A rendszert röviden $\{\boldsymbol{w}\}$-vel fogjuk jelölni.

6.6.3. Tétel. A $\{\boldsymbol{w}\}$ vektorrendszer bázis.

Bizonyítás. 1. *A lineáris függetlenség igazolása.* Tegyük fel, hogy $\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{p_i} c_{ij}\boldsymbol{w}_{ij} = 0$. A legmagasabb indexű $c_{ij}\boldsymbol{w}_{ij}$ tag indexe legyen $l + 1$, akkor

$$\boldsymbol{B}^l\sum_{i=1}^{k}\sum_{j=1}^{p_i} c_{ij}\boldsymbol{w}_{ij} = \sum_{i=1}^{k} c_{il}\boldsymbol{B}^l\boldsymbol{w}_{il} = \sum_{i=1}^{k} c_{il}\boldsymbol{w}_{i0} = 0.$$

(Itt a rendszer esetleg nem létező $\boldsymbol{w}_{il}$ elemeit is felhasználjuk, a bizonyítás kedvéért legyen $\boldsymbol{w}_{il} = \mathbf{1}$ és $c_{il} = 0$, ha $l > p_i$.) A $\boldsymbol{w}_{10}, \boldsymbol{w}_{20}, ..., \boldsymbol{w}_{k0}$ sajátértékek lineárisan függetlenek, tehát minden c_{il} szorzó nulla, ez ellentmond annak, hogy a maximális index l+ 1.

2. *A generálás bizonyítása.* Válasszunk ki tetszőlegesen egy $\boldsymbol{x} \in \mathbf{R}^n$ elemet. A bizonyítás $\boldsymbol{x}$ indexe szerinti teljes indukcióval történik. Ha $\boldsymbol{x}$ indexe

1, akkor $\boldsymbol{x}$ sajátvektor, ami a $\boldsymbol{w}_{10}, \boldsymbol{w}_{20}, ..., \boldsymbol{w}_{k0}$ vektorok lineáris kombinációjával előállítható. Tegyük fel, hogy ha egy elem indexe kisebb, mint $l + 1$, akkor a $\{\boldsymbol{w}\}$ rendszer generálja. Legyen most $\boldsymbol{x}$ indexe $l + 1$, akkor

$$\boldsymbol{B}^l x = \sum_{i=1}^{k} c_i \boldsymbol{w}_{i0} = \sum_{p_i \geq l} c_i \boldsymbol{B}^l \boldsymbol{w}_{il} + \sum_{p_i < l} c_i \boldsymbol{w}_{i0}.$$

Átrendezve

$$\boldsymbol{B}^l \left(\boldsymbol{x} - \sum_{p_i \geq l} c_i \boldsymbol{w}_{il} \right) = \boldsymbol{B}^l \boldsymbol{y} = \sum_{p_i < l} c_i \boldsymbol{w}_{i0}.$$

A $\{\boldsymbol{w}\}$ rendszer konstrukciója alapján azonban az l-nél nagyobb indexű a_i-k kiválasztása után már nem lesz olyan $\boldsymbol{y}$ elem a térben, melynek indexe $l + 1$, és gyökere lineárisan független a már kiválasztott gyökerektől, tehát $\boldsymbol{B}^l \boldsymbol{y} = \mathbf{0}$. $\boldsymbol{y}$ indexe tehát legfeljebb l, így a vektorrendszer generálja, és ekkor ugyanez igaz $\boldsymbol{x}$-re is. $\square$

A 6.6.3. Tétel folyománya az, hogy egyrészt $\sum_{i=1}^{k} (p_i + 1) = n$, másrészt – ha $k > 1$ darab sajátvektor létezik – a tér $\boldsymbol{A}$-ra invariáns halmazok direkt összegére bontható. A $\boldsymbol{w}_{i0}, \boldsymbol{w}_{i1}, ..., \boldsymbol{w}_{ip_i}$ vektorok által generált L_i altér nyilván invariáns altér, továbbá a bázis felbontása miatt $\mathbf{R}^n = L_1 \oplus L_2 \oplus \ldots \oplus L_k$. Minden L_i altérben pontosan egy lényegesen különböző sajátvektor van.

6.6.4. Következmény. Ha az $\boldsymbol{A}$ transzformációnak λ n-szeres sajátértéke, akkor megadható a másodlagos sajátvektorok $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ rendszere úgy, hogy

$$\boldsymbol{A}\boldsymbol{v}_i = \lambda \boldsymbol{v}_i + \varepsilon_i \boldsymbol{v}_{i-1},\ (i = 1, 2, \ldots, k),$$

ahol $\varepsilon_1 = 0$, és $\varepsilon_i = 0$ vagy 1.

A 6.6.4. Következményben megadott másodlagos vektorrendszer megegyezik a korábbi $\{\boldsymbol{w}\}$ rendszerrel.

6.6.5. Következmény (Jordan-alak). Ha az $\boldsymbol{A}$ transzformációnak λ n-szeres sajátértéke, akkor a bázis mindig megadható úgy, hogy $\boldsymbol{A}$ mátrixában a főátló elemei λ-k, a főátló feletti elemek 0-k vagy 1-esek, és a többi elem nulla. Ha $\boldsymbol{A}$ minden sajátértéke valós, akkor megadható a bázis oly módon, hogy a $\boldsymbol{A}$ főátlójában a sajátértékek sorakoznak a multiplicitásuknak megfelelő számban, a főátló feletti elemek 0-k vagy 1-esek, és a többi elem nulla.

Az első esetben a $\{\boldsymbol{w}\}$ rendszert választva bázisnak, $\boldsymbol{A}$ Jordan-alakot ölt. Második esetben a direktösszeg felbontás után minden összeadandóra el kell készíteni a Jordan-alakot.

Komplex környezetben minden lineáris transzformáció Jordan-alakra hozható.

11. Kérdés. Igaz-e, hogy az $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ mátrix tükröző transzformációval Jordan-alakra hozható? Határozzuk meg a másodlagos sajátvektorait!

6.7. A lineáris transzformáció normája

Mivel az $\mathbf{R}^n \to \mathbf{R}^n$ lineáris transzformációk ($n \times n$-es négyzetes mátrixok) lineáris teret alkotnak, felmerül a kérdés, hogy rendelhetünk e normát a transzformációkhoz. A normától a következő tulajdonságokat várjuk el (v.ö. 1.3. fejezet):

1. $||\boldsymbol{A}|| \geq 0$, és $||\boldsymbol{A}|| = 0$ akkor és csak akkor, ha $\boldsymbol{A} = \mathbf{0}$.
2. $||c\boldsymbol{A}|| = |c| \cdot ||\boldsymbol{A}||$.
3. $||\boldsymbol{A} + \boldsymbol{B}|| \leq ||\boldsymbol{A}|| + ||\boldsymbol{B}||$.

Másik ezzel összefüggő kérdés az, hogy korlátos-e a leképezés, vagyis korlátos halmaznak a képe is korlátos-e? Erre egyszerűen válaszolhatunk, igen. Becsüljük meg $||\boldsymbol{Ax}||^2$-et a Cauchy–Schwarz-egyenlőtlenséggel. Legyenek $\boldsymbol{A} = (a_{ij})$ sorvektorai $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$, akkor

$$\boldsymbol{Ax} = \begin{pmatrix} \langle \boldsymbol{a}_1, \boldsymbol{x} \rangle \\ \langle \boldsymbol{a}_2, \boldsymbol{x} \rangle \\ \vdots \\ \langle \boldsymbol{a}_n, \boldsymbol{x} \rangle \end{pmatrix},$$

és

$$||\boldsymbol{Ax}||^2 = \sum_{i=1}^{n} \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle^2 \leq \sum_{i=1}^{n} ||\boldsymbol{a}_i||^2 ||\boldsymbol{x}||^2 = ||\boldsymbol{x}||^2 \sum_{i,j} a_{ij}^2,$$

tehát, ha $\boldsymbol{x}$-ek halmaza (normában) korlátos, akkor $\boldsymbol{Ax}$-ek halmaza is. Ha az

$$||\boldsymbol{Ax}|| \leq K \cdot ||\boldsymbol{x}||$$

becslés minden $\boldsymbol{x} \in \mathbf{R}^n$-re teljesül, akkor a transzformációt *korlátosnak* nevezzük.

Az $\boldsymbol{A}$-hoz rendelt $\sqrt{\sum_{i,j} a_{ij}^2}$ érték tekinthető $\boldsymbol{A}$ normájának, teljesíti a normával szembeni elvárásokat, de ez egyszerűen az a_{ij} számokból készített n^2-dimenziós vektor normája, igy sok újdonságot nem várhatunk tőle. Másrészt az egyenlőtlenségben a K-nak ez az értéke meglehetősen rossz. A Cauchy–Schwarz-egyenlőtlenségben az egyenlőség csak úgy teljesülhet, ha $\boldsymbol{a}_i = c_i\boldsymbol{x}$, ez azt jelenti, hogy $\mathrm{rang}(\boldsymbol{A}) \leq 1$, vagyis 1-nél nagyobb rang esetén az egyenlőség nem érhető el semmilyen $\boldsymbol{x} \neq \mathbf{0}$ vektorra.

A legkisebb K konstans, melyre $||\boldsymbol{A}\boldsymbol{x}|| \leq K \cdot ||\boldsymbol{x}||$ minden $\boldsymbol{x}$-re teljesül, a

$$K = \sup_{\boldsymbol{z}} \frac{||\boldsymbol{A}\mathbf{z}||}{||\boldsymbol{z}||} = \sup_{\boldsymbol{z}} \left\| \boldsymbol{A}\frac{\boldsymbol{z}}{||\boldsymbol{z}||} \right\| = \sup_{||\boldsymbol{x}||=1} ||\boldsymbol{A}\boldsymbol{x}||$$

relációval adható meg. Ebből látható K-nak az a jelentése is, hogy az n-dimenzós tér egységgömbjének a torzulását adja meg. Ezt a számot rendeljük hozzá az $\boldsymbol{A}$ transzformációhoz normaként.

6.7.1. Definíció. $\boldsymbol{A}$ *normájának* a $||\boldsymbol{A}|| = \sup\limits_{||\boldsymbol{x}||=1} ||\boldsymbol{A}\boldsymbol{x}||$ számot tekintjük.

Bizonyítás. A definíció értelmes, mert – mint korábban láttuk – $||\boldsymbol{A}|| \leq \sqrt{\sum_{i,j} a_{ij}^2}$, tehát véges számérték. A normát definiáló tulajdonságok is teljesülnek: ha $||\boldsymbol{A}|| = 0$, akkor $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ minden $\boldsymbol{x}$-re, amiből $\boldsymbol{A} = \mathbf{0}$. A háromszög egyenlőtlenség az

$$\begin{aligned} ||\boldsymbol{A}+\boldsymbol{B}|| &= \sup ||(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x}|| \leq \sup(||\boldsymbol{A}\boldsymbol{x}|| + ||\boldsymbol{B}\boldsymbol{x}||) \leq \\ &\leq \sup ||\boldsymbol{A}\boldsymbol{x}|| + \sup ||\boldsymbol{B}\boldsymbol{x}|| = ||\boldsymbol{A}|| + ||\boldsymbol{B}|| \end{aligned}$$

képletsorból következik, ahol a szuprémum az $||\boldsymbol{x}|| = 1$ vektorokra vonatkozik. □

A definíció alapján könnyen látható, hogy $||\boldsymbol{A}\boldsymbol{B}|| \leq ||\boldsymbol{A}|| \cdot ||\boldsymbol{B}||$, ugyanis $||\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}|| \leq ||\boldsymbol{A}|| \cdot ||\boldsymbol{B}\boldsymbol{x}|| \leq ||\boldsymbol{A}|| \cdot ||\boldsymbol{B}|| \cdot ||\boldsymbol{x}||$.

Ortogonális bázistranszformációt elvégezve a transzformáció normája nem változik, hiszen az egységgömb torzulása koordináta-rendszertől független. Az állítás algebrailag is bizonyítható.

$\boldsymbol{C}$ legyen ortogonális mátrix, akkor $||\boldsymbol{A}|| = ||\boldsymbol{C}^{-1}\boldsymbol{A}||$, hiszen a normatartás miatt $||\boldsymbol{A}\boldsymbol{x}|| = ||\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{x}||$, továbbá

$$||\boldsymbol{A}|| = ||\boldsymbol{C}^{-1}\boldsymbol{A}|| = \sup_{||\boldsymbol{x}||=1} ||\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{x}|| = \sup_{||\boldsymbol{x}||=1} ||\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}\boldsymbol{C}^{-1}\boldsymbol{x}|| =$$

$$= \sup_{||y||=1} ||\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y}|| = ||\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}||.$$

6.7.2. Tétel. Szimmetrikus mátrix normája a sajátértékek abszolút értékének a maximuma.

Bizonyítás. Szimmetrikus mátrix ortogonális transzformációval $\boldsymbol{D}$ diagonális mátrixszá alakítható, ahol a diagonális elemei a $\lambda_1, \lambda_2, \ldots, \lambda_n$ sajátértékek. Legyen $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ és $||\boldsymbol{x}|| = 1$, akkor

$$||\boldsymbol{D}\boldsymbol{x}||^2 = \sum_{i=1}^{n} \lambda_i^2 x_i^2 \leq \max_i \lambda_i^2 ||\boldsymbol{x}||^2 = (\max_i |\lambda_i|)^2.$$

Ennek alapján $||\boldsymbol{D}|| \leq \max|\lambda_i|$. Ugyanakkor, ha $\boldsymbol{x} = \boldsymbol{e}_i$, akkor $||\boldsymbol{D}\boldsymbol{x}|| = |\lambda_i|$, tehát $||\boldsymbol{D}|| \geq |\lambda_i|$ minden i-re, tehát $||\boldsymbol{D}|| \geq \max|\lambda_i|$. Mindezek alapján $||\boldsymbol{A}|| = ||\boldsymbol{D}|| = \max|\lambda_i|$. $\square$

A következőkben szerepet kap az $\boldsymbol{A}^*\boldsymbol{A}$ mátrix. $\boldsymbol{A}^*\boldsymbol{A}$ szimmetrikus mátrix, mert – felhasználva, hogy $(\boldsymbol{A}\boldsymbol{B})^* = \boldsymbol{B}^*\boldsymbol{A}^*$ –

$$(\boldsymbol{A}^*\boldsymbol{A})^* = \boldsymbol{A}^*(\boldsymbol{A}^*)^* = \boldsymbol{A}^*\boldsymbol{A}.$$

Az $\boldsymbol{A}^*\boldsymbol{A}$ szimmetrikus, tehát sajátértékei valósak, sőt mi több, nemnegatívok. Legyen λ az $\boldsymbol{A}^*\boldsymbol{A}$ egyik sajátértéke, és $\boldsymbol{v}$ a hozzá tartozó sajátvektor, akkor

$$0 \leq ||\boldsymbol{A}\boldsymbol{v}||^2 = \langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{A}\boldsymbol{v} \rangle = \langle \boldsymbol{A}^*\boldsymbol{A}\boldsymbol{v}, \boldsymbol{v} \rangle = \langle \lambda\boldsymbol{v}, \boldsymbol{v} \rangle = \lambda ||\boldsymbol{v}||^2,$$

amiből $\lambda \geq 0$.

6.7.3. Tétel. Az $||\boldsymbol{A}||$ megegyezik $\boldsymbol{A}^*\boldsymbol{A}$ legnagyobb sajátértékének a négyzetgyökével.

Bizonyítás. Legyen $\boldsymbol{A}^*\boldsymbol{A}$ legnagyobb sajátértéke λ és a hozzátartozó sajátvektor $\boldsymbol{v}$. A 6.7.2 Tétel alapján tudjuk, hogy $||\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x}|| \leq \lambda||\boldsymbol{x}||$. A Cauchy–Schwarz-egyenlőtlenséget felhasználva

$||\boldsymbol{A}\boldsymbol{x}||^2 = \langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{A}\boldsymbol{x} \rangle = \langle \boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \rangle \leq ||\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x}|| \cdot ||\boldsymbol{x}|| \leq \lambda ||\boldsymbol{x}||^2$,
amiből $||\boldsymbol{A}|| \leq \sqrt{\lambda}$ adódik. Ha $\boldsymbol{x} = \boldsymbol{v}$, akkor ebben az egyenlőtlenségsorozatban végig egyenlőség áll fenn, tehát $||\boldsymbol{A}|| = \sqrt{\lambda}$. $\square$

12. Kérdés. Igaz-e, hogy minden vetítésnek (a $\mathbf{0}$-t és $\boldsymbol{I}$-t itt nem tekintve vetítésnek) pontosan két különböző sajátértéke van? Igaz-e, hogy minden vetítés normája 1?

13. Kérdés. Megegyeznek-e $\boldsymbol{A}\boldsymbol{B}$ és $\boldsymbol{B}\boldsymbol{A}$ sajátértékei? A sajátértékek multiplicitása is megegyezik? Igaz-e, hogy $||\boldsymbol{A}|| = ||\boldsymbol{A}^*||$?

A $\rho(\boldsymbol{A}) = \max_k |\lambda_k|$ értéket, ahol λ_k az $\boldsymbol{A}$ sajátértékeit jelöli, *spektrál sugárnak* nevezik. Könnyen igazolható, hogy $\rho(\boldsymbol{A}) \le ||\boldsymbol{A}||$. Jelöljük ugyanis a λ_k sajátértékhez tartozó egységnormájú sajátvektort $\boldsymbol{v}_k$-val, akkor

$$||\boldsymbol{A}|| = \max_{||\boldsymbol{x}||=1} ||\boldsymbol{A}\boldsymbol{x}|| \ge \max_k ||\boldsymbol{A}\boldsymbol{v}_k|| = \max_k |\lambda_k|\, ||\boldsymbol{v}_k|| = \max_k |\lambda_k| = \rho(\boldsymbol{A}).$$

Mátrixok konvergenciáját több módon lehet definiálni. A legtermészetesebb az elemenkénti konvergencia: $\boldsymbol{A}_k$ konvergál $\boldsymbol{A}$-hoz, ha $\boldsymbol{A}_k$ minden eleme az $\boldsymbol{A}$ azonos pozícióban lévő eleméhez tart. A másik lehetőség a normában való konvergencia: $\boldsymbol{A}_k$ konvergál $\boldsymbol{A}$-hoz normában, ha $||\boldsymbol{A}_k - \boldsymbol{A}|| \to 0$. A két konvergencia-fogalom ekvivalens.

6.7.4. Tétel. Legyen $\boldsymbol{A}_k = (a_{ij}^{(k)})$ és $\boldsymbol{A} = (a_{ij})$. $||\boldsymbol{A}_k - \boldsymbol{A}|| \to 0$ akkor és csak akkor, ha minden i-re és j-re $a_{ij}^{(k)} \to a_{ij}$.

Bizonyítás. Elég azt bizonyítani, hogy $||\boldsymbol{A}_k|| \to 0$ akkor és csak akkor, ha minden i-re és j-re $a_{ij}^{(k)} \to 0$.

$\Rightarrow$ Legyen $\boldsymbol{A}_k = (\boldsymbol{a}_1^{(k)}, \boldsymbol{a}_2^{(k)}, ..., \boldsymbol{a}_n^{(k)})$, akkor $||\boldsymbol{a}_i^{(k)}|| = ||\boldsymbol{A}_k \boldsymbol{e}_i|| \le ||\boldsymbol{A}_k|| \to$ $\to 0$, ami azt jelenti, hogy $\boldsymbol{a}_i^{(k)}$ minden eleme nullához tart, és ez minden i-re teljesül.

$\Leftarrow$ Tudjuk, hogy $||\boldsymbol{A}_k|| \le \sqrt{\sum_{i,j} (a_{ij}^{(k)})^2}$, és ha ez utóbbi nullához tart, akkor $||\boldsymbol{A}_k||$ is. $\square$

6.7.5. Következmény. Ha $\boldsymbol{A}^*\boldsymbol{A}$ minden sajátértéke egynél kisebb, akkor $\boldsymbol{A}^k \to \boldsymbol{0}$ abban az értelemben, hogy minden eleme nullához tart.

Bizonyítás. A 6.7.3. Tétel szerint $||\boldsymbol{A}|| = c < 1$. Legyen $\boldsymbol{x}$ tetszőleges vektor, akkor

$$||\boldsymbol{A}^k\boldsymbol{x}|| = ||\boldsymbol{A}(\boldsymbol{A}^{k-1}\boldsymbol{x})|| \le c||\boldsymbol{A}^{k-1}\boldsymbol{x}|| \le$$
$$\le c^2||\boldsymbol{A}^{k-2}\boldsymbol{x}|| \le ... \le c^{k-1}||\boldsymbol{A}\boldsymbol{x}|| \le c^k||\boldsymbol{x}||.$$

Ennek alapján $||\boldsymbol{A}^k|| \le c^k \to 0$, ami az előző tétel miatt adja az állítást. $\square$

14. Kérdés. Tudunk-e egyszerű példát adni arra, hogy $\boldsymbol{A}$ minden sajátértéke abszolút értékben egynél kisebb, és $||\boldsymbol{A}|| > 1$?

Fontos szerepe van a normának a lineáris egyenletrendszerek stabilitásvizsgálatánál. Az $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ egyenletrendszerben feltesszük a kérdést, hogy a $\boldsymbol{b}$ állandókból álló vektor $\Delta\boldsymbol{b}$-vel jelölt megváltozása mennyire befolyásolja a megoldást. Gazdasági gyakorlatban: a kapacitás bővítése mennyire hat a termelés megváltozására. Feltételezzük, hogy $\boldsymbol{A}$ invertálható mátrix.

Jelöljük az $\boldsymbol{x}$ megváltozását $\Delta\boldsymbol{x}$-szel, akkor $\boldsymbol{A}(\boldsymbol{x} + \Delta\boldsymbol{x}) = \boldsymbol{b} + \Delta\boldsymbol{b}$, amiből $\boldsymbol{x} + \Delta\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} + \boldsymbol{A}^{-1}\Delta\boldsymbol{b}$. Mivel $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$, $\Delta\boldsymbol{x} = \boldsymbol{A}^{-1}\Delta\boldsymbol{b}$. Normákkal megbecsülve

$$||\Delta\boldsymbol{x}|| \leq ||\boldsymbol{A}^{-1}|| \cdot ||\Delta\boldsymbol{b}||, \quad \text{és} \quad ||\boldsymbol{b}|| \leq ||\boldsymbol{A}|| \cdot ||\boldsymbol{x}||.$$

Szorozzuk össze a két egyenlőtlenséget, akkor

$$||\Delta\boldsymbol{x}|| \cdot ||\boldsymbol{b}|| \leq ||\boldsymbol{A}|| \cdot ||\boldsymbol{A}^{-1}|| \cdot ||\Delta\boldsymbol{b}|| \cdot ||\boldsymbol{x}||,$$

amiből

$$\frac{||\Delta\boldsymbol{x}||}{||\boldsymbol{x}||} \leq \ ||\boldsymbol{A}|| \cdot ||\boldsymbol{A}^{-1}|| \frac{||\Delta\boldsymbol{b}||}{||\boldsymbol{b}||}$$

felső becslést kapunk $\boldsymbol{x}$ relatív megváltozására.

Alsó becslés ugyanígy elérhető. Az $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ és a $\Delta\boldsymbol{b} = \boldsymbol{A}\Delta\boldsymbol{x}$ egyenletekből

$$||\boldsymbol{x}|| \leq ||\boldsymbol{A}^{-1}|| \cdot ||\boldsymbol{b}||, \quad \text{és} \quad ||\Delta\boldsymbol{b}|| \leq ||\boldsymbol{A}|| \cdot ||\Delta\boldsymbol{x}||,$$

ezeket összeszorozva

$$||\boldsymbol{x}|| \cdot ||\Delta\boldsymbol{b}|| \leq ||\boldsymbol{A}|| \cdot ||\boldsymbol{A}^{-1}|| \cdot ||\boldsymbol{b}|| \cdot ||\Delta\boldsymbol{x}||,$$

és

$$\frac{||\Delta\boldsymbol{x}||}{||\boldsymbol{x}||} \geq \ \frac{1}{||\boldsymbol{A}|| \cdot ||\boldsymbol{A}^{-1}||} \frac{||\Delta\boldsymbol{b}||}{||\boldsymbol{b}||}.$$

Látható tehát, hogy a relatív megváltozás az $||\boldsymbol{A}|| \cdot ||\boldsymbol{A}^{-1}||$ mennyiség nagyságán múlik mindkét irányban. (Nyilván $||\boldsymbol{A}|| \cdot ||\boldsymbol{A}^{-1}|| \geq ||\boldsymbol{A}\boldsymbol{A}^{-1}|| = = ||\boldsymbol{I}|| = 1$.)

6.8. Hatványsorok

Láttuk, hogy $||\boldsymbol{A}|| < 1$ esetén $\boldsymbol{A}^n \to \mathbf{0}$. Ennél azonban többet is állíthatunk. Ebből a célból először vizsgáljuk meg egy speciális alakú $\boldsymbol{B}$ mátrix hatványait. Legyen

$$\boldsymbol{B} = \begin{pmatrix} \alpha & & C & \\ & \alpha & & \\ & & \ddots & \\ & C & & \\ & & & \alpha \end{pmatrix},$$

vagyis $\boldsymbol{B}$ felső háromszög alakú mátrix, melyben a főátló feletti elemek azonosan C-vel egyenlők, és a főátló elemei is egyenlők. Előírjuk, hogy $C > 0$, és $0 < \alpha < 1$.

6.8.1. Segédtétel. $\boldsymbol{B}^m \to \mathbf{0}$ ha $m \to \infty$.

Bizonyítás. Belátjuk, hogy $\boldsymbol{B}^m$ az alábbi szerkezetű mátrix és megadjuk $r_k^{(m)}$ explicit előállítását, amelyből az állítás már leolvasható:

$$\boldsymbol{B}^m = \begin{pmatrix} \alpha^m & r_1^{(m)} & \cdots & r_{n-1}^{(m)} \\ 0 & \alpha^m & \ddots & \vdots \\ \vdots & & \ddots & r_1^{(m)} \\ 0 & 0 & \cdots & \alpha^m \end{pmatrix},$$

ahol

$$r_k^{(m)} = \alpha^m \sum_{j=1}^{k} \binom{k-1}{j-1} \binom{m}{j} D^j$$

és $D = \frac{C}{\alpha}$. Ha ezt tudjuk, akkor $r_k^{(m)} \to 0$, mert α^m m-nek egy k-adfokú polinomjával szorzódik.

Mivel

$$\boldsymbol{B}^m = \begin{pmatrix} \alpha^m & r_1^{(m)} & \cdots & r_{n-1}^{(m)} \\ 0 & \alpha^m & \ddots & \vdots \\ \vdots & & \ddots & r_1^{(m)} \\ 0 & 0 & \cdots & \alpha^m \end{pmatrix} = \boldsymbol{B} \begin{pmatrix} \alpha^{m-1} & r_1^{(m-1)} & \cdots & r_{n-1}^{(m-1)} \\ 0 & \alpha^{m-1} & \ddots & \vdots \\ \vdots & & \ddots & r_1^{(m-1)} \\ 0 & 0 & \cdots & \alpha^{m-1} \end{pmatrix},$$

a mátrixszorzás elvégzésével az

$$r_k^{(m)} = \alpha r_k^{(m-1)} + C(\alpha^{m-1} + r_1^{(m-1)} + r_2^{(m-1)} + ... + r_{k-1}^{(m-1)})$$

rekurzív egyenlet adódik. Ennek segítségével m-szerinti teljes indukcióval tudjuk bizonyítani az $r_k^{(m)}$-re megadott formulát. $m = 1$ esetén $r_k^{(1)} = \alpha D = C$, ami nyilván igaz. Tegyük fel a formula helyességét minden k mellett $m-1$-re, akkor

$$r_k^{(m)} = \alpha^m \sum_{j=1}^{k} \binom{k-1}{j-1} \binom{m-1}{j} D^j +$$

$$+\alpha^m D(1 + \sum_{i=1}^{k-1} \sum_{j=1}^{i} \binom{i-1}{j-1} \binom{m-1}{j} D^j) =$$

$$= \alpha^m \sum_{j=0}^{k-1} \binom{k-1}{j} \binom{m-1}{j+1} D^{j+1} +$$

$$+\alpha^m D(1+\sum_{j=1}^{k-1}\sum_{i=j}^{k-1}\begin{pmatrix} i-1 \\ j-1 \end{pmatrix}\begin{pmatrix} m-1 \\ j \end{pmatrix}D^j).$$

Mivel $\sum\limits_{i=j}^{k-1}\begin{pmatrix} i-1 \\ j-1 \end{pmatrix}=\begin{pmatrix} k-1 \\ j \end{pmatrix}$ a Pascal-háromszög elemi tulajdonságai miatt,

$$r_k^{(m)}=\alpha^m((m-1)D+D+\sum_{j=1}^{k-1}\begin{pmatrix} k-1 \\ j \end{pmatrix}\begin{pmatrix} m-1 \\ j+1 \end{pmatrix}D^{j+1}+$$

$$+\sum_{j=1}^{k-1}\begin{pmatrix} k-1 \\ j \end{pmatrix}\begin{pmatrix} m-1 \\ j \end{pmatrix}D^{j+1})=$$

$$=\alpha^m(mD+\sum_{j=1}^{k-1}\begin{pmatrix} k-1 \\ j \end{pmatrix}\begin{pmatrix} m \\ j+1 \end{pmatrix}D^{j+1}),$$

ami megfelel a bizonyítandó formulának. □

A 6.8.1. Segédtétel nem pusztán játszadozás a számokkal (binomiális együtthatókkal). Mondjuk azt, hogy az $\boldsymbol{A}=(a_{ij})$ mátrixnak a majoránsa valamely $\boldsymbol{B}=(b_{ij})$ mátrix, ha méreteik megegyeznek, és minden elemükre $|a_{ij}|\leq b_{ij}$. Könnyen látható, hogy ha $\boldsymbol{A}_1$ majoránsa $\boldsymbol{B}_1$ és $\boldsymbol{A}_2$ majoránsa $\boldsymbol{B}_2$, akkor $\boldsymbol{A}_1\boldsymbol{A}_2$-nek is majoránsa a $\boldsymbol{B}_1\boldsymbol{B}_2$.

6.8.2. Segédtétel. Ha $\boldsymbol{A}$ felső háromszög alakú mátrix és a főátló elemeinek (vagyis a sajátértékeinek) az abszolút értéke 1-nél kisebb, akkor $\boldsymbol{A}^m\to\boldsymbol{0}$. (A mátrix elemei komplex számok is lehetnek.)

Bizonyítás. A segédtételben szereplő $\boldsymbol{A}$ mátrix mindig majorálható a 6.8.1. Segédtételben szereplő $\boldsymbol{B}$ mátrixszal az α és a C alkalmas megválasztása esetén. Mivel $\boldsymbol{A}^m$-et a $\boldsymbol{B}^m$ majorálja, és $\boldsymbol{B}^m\to\boldsymbol{0}$, kapjuk, hogy $\boldsymbol{A}^m\to$ $\to\boldsymbol{0}$. □

6.8.3. Tétel. Adott $\boldsymbol{A}$ mátrix esetén $\boldsymbol{A}^m\to\boldsymbol{0}$ akkor és csak akkor teljesül, ha sajátértékei (a komplex sajátértékeket is beleértve) a komplex egységkör belsejébe esnek, azaz, ha spektrálsugara egynél kisebb.

Bizonyítás. $\Leftarrow$ Az $\boldsymbol{A}$ mátrix, komplex elemeket is megengedve, alkalmas transzformációval felső háromszög alakra hozható, vagyis elérhető, hogy $\boldsymbol{CAC}^{-1}$ már felső háromszög mátrix legyen. Mivel a transzformáció során a sajátértékek nem változnak, $\boldsymbol{CAC}^{-1}$ főátlójának elemei egynél kisebb abszolút értékűek. A 6.8.2. Segédtétel szerint $(\boldsymbol{CAC}^{-1})^m=\boldsymbol{CA}^m\boldsymbol{C}^{-1}\to\boldsymbol{0}$, de ekkor $\boldsymbol{A}^m=\boldsymbol{C}^{-1}(\boldsymbol{CA}^m\boldsymbol{C}^{-1})\boldsymbol{C}\to\boldsymbol{0}$.

$\Rightarrow$ Ha $\boldsymbol{A}^m \to \boldsymbol{0}$, akkor tetszőleges, rögzített $\boldsymbol{x}$-re $\boldsymbol{A}^m\boldsymbol{x} \to \boldsymbol{0}$. Ha $\boldsymbol{A}$-nak van $|\lambda| \geq 1$ sajátértéke, és $\boldsymbol{v}$ a hozzátartozó sajátvektor, akkor $\boldsymbol{A}^m\boldsymbol{v} = \lambda^m\boldsymbol{v}$, ami nem tart $\boldsymbol{0}$-hoz, így ellentmondást kapunk.□

A 6.8.3. Tétel lehetőséget ad mátrix elemekből álló hatványsorok definiálására. Legyen az $f(x)$ függvény az

$$f(x) = \sum_{k=0}^{\infty} a_k x^x$$

hatványsorral megadva, ahol feltételezzük, hogy a hatványsor konvergenciakörének a sugara r.

6.8.4. Definíció. A fenti $f(x)$ függvényre definiáljuk az $f(\boldsymbol{A})$ kifejezést az

$$f(\boldsymbol{A}) = \sum_{k=0}^{\infty} a_k \boldsymbol{A}^k$$

sorral (definíció szerint $\boldsymbol{A}^0 = \boldsymbol{I}$). A sor konvergens, ha $\boldsymbol{A}$ minden (valós vagy komplex) sajátértéke a hatványsor konvergencia-körének a belsejébe esik, vagy más szóval $\boldsymbol{A}$ spektrál sugara kisebb mint r. Ha a spektrál sugár nagyobb mint r, akkor a sor divergens.

Bizonyítás. Legyen $\boldsymbol{A}$ spektrál sugara $\rho < r$, és vegyünk fel egy c számot úgy, hogy $\rho < c < r$. Tudjuk, hogy a $\sum\limits_{k=0}^{\infty} a_k c^k$ sor abszolút konvergens, tehát $\sum\limits_{k=m}^{\infty} |a_k| c^k < \varepsilon$, ha $m > m_1$. Az $\frac{1}{c}\boldsymbol{A}$ sajátértékei az egységkörbe esnek, tehát $\left(\frac{1}{c}\boldsymbol{A}\right)^m \to 0$, így a mátrix minden eleme abszolút értékben 1-nél kisebb lesz, ha $m > m_2$. Ebből következik, hogy $\left|\sum\limits_{k=m}^{\infty} a_k c^k \left(\frac{1}{c}\boldsymbol{A}\right)^k\right| < \varepsilon$, ha $m > \max(m_1, m_2)$, vagyis a $\sum\limits_{k=0}^{\infty} a_k \boldsymbol{A}^k$ sor konvergens.

Ha $S_m(\boldsymbol{A}) = \sum\limits_{k=0}^{m} a_k \boldsymbol{A}^k \to \boldsymbol{C}$, akkor minden $\boldsymbol{x}$-re $S_m(\boldsymbol{A})\boldsymbol{x} \to \boldsymbol{C}\boldsymbol{x}$. Ha $\rho > r$, és $\boldsymbol{v}$ olyan sajátvektor, amelyhez tartozó λ sajátértékre $|\lambda| = \rho$, akkor $S_m(\boldsymbol{A})\boldsymbol{v} = S_m(\lambda)\boldsymbol{v}$, és ez utóbbi nem konvergens, hiszen $|\lambda| > r$. □

A 6.8.4. Definíció felhasználásával mátrixok különböző függvényeit tudjuk definiálni, például értelmezhetjük a $\sin\boldsymbol{A}$, az $e^{\boldsymbol{A}}$ és (bizonyos feltételek mellett) a $\log\boldsymbol{A}$ mátrixokat.

15. Kérdés. Igaz-e, hogy minden négyzetes $\boldsymbol{A}$ mátrixra

$$\sin^2\boldsymbol{A} + \cos^2\boldsymbol{A} = \boldsymbol{I}?$$

Legyen $f(x)$ a fenti, hatványsorral megadott függvény és $\boldsymbol{A}$ spektrál sugara legyen kisebb, mint f konvergencia sugara. Ha $\boldsymbol{A}$ sajátértékei a λ_1, $\lambda_2, \ldots, \lambda_n$ számok, akkor $S_m(\boldsymbol{A}) = \sum\limits_{k=0}^{m} a_k \boldsymbol{A}^k$ sajátértékei az $S_m(\lambda_i)$ számok $(i = 1, 2, \ldots, n)$, ugyanis vegyünk egy λ_i-hez tartozó $\boldsymbol{x}$ sajátvektort, akkor

$$S_m(\boldsymbol{A})x = \sum_{k=0}^{m} a_k \boldsymbol{A}^k \boldsymbol{x} = \sum_{k=0}^{m} a_k \lambda_i^k \boldsymbol{x} = S_m(\lambda_i)\boldsymbol{x}.$$

Határátmenettel kapjuk, hogy $f(\boldsymbol{A})$ sajátértékei az $f(\lambda_1)$, $f(\lambda_2), \ldots, f(\lambda_n)$ számok.

6.9. Kvadratikus alakok

Először a bilineáris kifejezésekről ejtsünk szót. Olyan $f(\boldsymbol{x}, \boldsymbol{y})$ valós értékű függvényről van szó ($\boldsymbol{x} \in \mathbf{R}^n$, $\boldsymbol{y} \in \mathbf{R}^n$), mely az egyik változó rögzített értéke mellett a másikban, és a másik rögzítése esetén az elsőben lineáris függvény. Könnyen látható, hogy általános alakja, ha $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ és $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$, akkor $f(\boldsymbol{x}, \boldsymbol{y}) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} a_{ij} x_i y_j$. A tömörebb mátrixos felírásnál jelöljük $\boldsymbol{A}$-val az $\boldsymbol{A} = (a_{ij})$ $n \times n$-es mátrixot.

6.9.1. Definíció. Az $f(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{y} \rangle$ kétváltozós függvényt ($\boldsymbol{x} \in \mathbf{R}^n$, $\boldsymbol{y} \in \mathbf{R}^n$) *bilineáris alaknak*, bilineáris függvénynek nevezzük.

Bilineáris alakok több helyen előfordulnak a matematikában (és a fizikában), ezek egyik legtipikusabb esete a mátrixjátékok elmélete.

Ha egy $\boldsymbol{C}$ bázistranszformációt alkalmazunk, akkor az $\boldsymbol{x}$ és az $\boldsymbol{y}$ vektorok koordinátái megváltoznak és $\boldsymbol{C}^{-1}\boldsymbol{x}$, ill. $\boldsymbol{C}^{-1}\boldsymbol{y}$ alakot öltenek (4.1.2. Tétel), ha azt akarjuk, hogy a bilineáris kifejezés értéke változatlan maradjon, akkor az $\boldsymbol{A}$ mátrixot meg kell változtatni. A változás azonban eltér a szokásos $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ transzformációs alaktól (4.1.4. Tétel).

6.9.2. Tétel. A $\boldsymbol{C}$ bázistranszformációt végrehajtva a bilineáris függvény alakja megváltozik és mátrixa $\boldsymbol{C}^*\boldsymbol{A}\boldsymbol{C}$ lesz.

Bizonyítás. Az $\boldsymbol{x}$ és az $\boldsymbol{y}$ vektorok a bázistranszformáció után $\boldsymbol{C}^{-1}\boldsymbol{x}$, ill. $\boldsymbol{C}^{-1}\boldsymbol{y}$ alakot öltenek (4.1.2. Tétel). Ha az új alak mátrixa $\boldsymbol{X}$, akkor azt kívánjuk meg, hogy $\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{y} \rangle = \langle \boldsymbol{C}^{-1}\boldsymbol{x}, \boldsymbol{X}\boldsymbol{C}^{-1}\boldsymbol{y} \rangle$ legyen. Ha $\boldsymbol{X} = \boldsymbol{C}^*\boldsymbol{A}\boldsymbol{C}$, akkor valóban $\langle \boldsymbol{C}^{-1}\boldsymbol{x}, \boldsymbol{X}\boldsymbol{C}^{-1}\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{C}^{-1*}\boldsymbol{C}^*\boldsymbol{A}\boldsymbol{C}\boldsymbol{C}^{-1}\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{y} \rangle$. $\square$

Megjegyezzük, hogy ortogonális transzformációknál, ahol $\boldsymbol{C}^{-1} = \boldsymbol{C}^*$, ott a két transzformált mátrix, a $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ és a $\boldsymbol{C}^*\boldsymbol{A}\boldsymbol{C}$ megegyezik.

A bilineáris függvénynél még gyakoribb a kvadratikus alak. Két dimenzióban a másodfokú görbék, három dimenzióban a másodfokú felületek egyenletének fő alkatrésze. A görbék, ill. felületek osztályozását (elliptikus, parabolikus, hiperbolikus) a másodfokú rész, a kvadratikus alak vizsgálata jelenti. Kvadratikus alakot kapunk a bilineáris függvényből, ha mindkét változónak $\boldsymbol{x}$-et választunk.

6.9.3. Definíció. A $q(\boldsymbol{x}) = \langle \boldsymbol{x},\ \boldsymbol{A}\boldsymbol{x} \rangle$ függvényt ($\boldsymbol{x} \in \mathbf{R}^n$) *kvadratikus alaknak*, kvadratikus függvénynek nevezzük.

Kvadratikus alakoknál az $\boldsymbol{A}$ mátrix megválasztása nem egyértelmű, hiszen $x_i x_j$ együtthatója $a_{ij} + a_{ji}$, és egy adott együttható többféle módon bontható fel az $a_{ij} + a_{ji}$ összegre. Egyértelművé tehető azonban, és ezt mindig érdemes megtenni, ha feltételezzük, hogy $\boldsymbol{A}$ szimmetrikus mátrix. Az egyértelműség szinte magától értetődőnek látszik, mégis látványos bizonyítása van.

6.9.4. Tétel. Minden $q(\boldsymbol{x}) = \langle \boldsymbol{x},\ \boldsymbol{A}_1\boldsymbol{x} \rangle$ kvadratikus alak felírható $q(\boldsymbol{x}) = \langle \boldsymbol{x},\ \boldsymbol{A}\boldsymbol{x} \rangle$ alakban, ahol $\boldsymbol{A}$ szimmetrikus mátrix. Ez a felírás egyértelmű.

Bizonyítás. Legyen $\boldsymbol{A} = \frac{1}{2}(\boldsymbol{A}_1 + \boldsymbol{A}_1^*)$ szimmetrikus mátrix, hiszen transzponáltja önmaga, akkor

$$\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{x} \rangle = \langle \boldsymbol{x}, \frac{1}{2}(\boldsymbol{A}_1 + \boldsymbol{A}_1^*)\boldsymbol{x} \rangle =$$
$$= \frac{1}{2}\langle \boldsymbol{x}, \boldsymbol{A}_1\boldsymbol{x} \rangle + \frac{1}{2}\langle \boldsymbol{x}, \boldsymbol{A}_1^*\boldsymbol{x} \rangle = \frac{1}{2}\langle \boldsymbol{x}, \boldsymbol{A}_1\boldsymbol{x} \rangle + + \frac{1}{2}\langle \boldsymbol{A}_1\boldsymbol{x}, \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \boldsymbol{A}_1\boldsymbol{x} \rangle.$$

Tegyük fel, hogy $q(\boldsymbol{x}) = \langle \boldsymbol{x},\ \boldsymbol{B}\boldsymbol{x} \rangle$, ahol $\boldsymbol{B}$ is szimmetrikus mátrix, akkor a két egyenlőséget kivonva egymásból $\langle \boldsymbol{x},\ (\boldsymbol{A} - \boldsymbol{B})\boldsymbol{x} \rangle = 0$ adódik minden $\boldsymbol{x}$-re. Legyen $\boldsymbol{D} = \boldsymbol{A} - \boldsymbol{B}$, akkor $\langle \boldsymbol{x},\ \boldsymbol{D}\boldsymbol{x} \rangle = 0$, minden $\boldsymbol{x}$-re, és $\boldsymbol{D}$ is szimmetrikus mátrix. Válasszuk $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{D}\boldsymbol{y}$-t, akkor

$$0 = \langle \boldsymbol{y} + \boldsymbol{D}\boldsymbol{y}, \boldsymbol{D}(\boldsymbol{y} + \boldsymbol{D}\boldsymbol{y}) \rangle =$$
$$= \langle \boldsymbol{y}, \boldsymbol{D}\mathbf{y} \rangle + \langle \boldsymbol{D}\boldsymbol{y}, \boldsymbol{D}^2\boldsymbol{y} \rangle + \langle \boldsymbol{D}\boldsymbol{y}, \boldsymbol{D}\boldsymbol{y} \rangle + \langle \boldsymbol{y}, \boldsymbol{D}^2\boldsymbol{y} \rangle = 2\langle \boldsymbol{D}\boldsymbol{y}, \boldsymbol{D}\boldsymbol{y} \rangle = 2||\boldsymbol{D}\boldsymbol{y}||^2$$

vagyis $\boldsymbol{D}\boldsymbol{y} = \mathbf{0}$ minden $\boldsymbol{y}$-ra, így $\boldsymbol{D} = \mathbf{0}$. Ez azt jelenti, hogy $\boldsymbol{A} = \boldsymbol{B}$. □

A továbbiakban feltételezzük, hogy a kvadratikus alakok mindig a szimmetrikus mátrixszal vannak megadva.

A 6.4.3. Tétel alapján minden szimmetrikus mátrix ortogonális transzformációval diagonálissá alakítható. Ennek a kvadratikus alakokra vonatkozó állítása csupán átfogalmazás.

6.9.5. Tétel. Ortogonális bázistranszformációval minden kvadratikus alak tiszta négyzetes alakra hozható, ahol az együtthatók a mátrix sajátértékei és minden sajátérték a multiplictásának megfelelő számban fordul elő. *Tiszta négyzetes alaknak* nevezzük azt a kvadratikus alakot, ahol $a_{ij} = 0$, ha $i \neq j$.

Az ortogonális transzformációt, ami a tiszta négyzetes alakot létrehozza, főtengely-transzformációnak is nevezik.

A kvadratikus alakokat az értékkészletük alapján szokás osztályozni. Mivel $q(\mathbf{0}) = 0$, a nulla helyet itt pillanatnyilag zárjuk ki az értelmezési tartományból.

6.9.6. Definíció. A $q(\boldsymbol{x})$ kvadratikus alakot

a) *pozitív definitnek* nevezzük, ha $q(\boldsymbol{x}) > 0$ minden $\boldsymbol{x} \neq 0$-ra.

b) *negatív definitnek* nevezzük, ha $q(\boldsymbol{x}) < 0$ minden $\boldsymbol{x} \neq 0$-ra.

c) *pozitív szemidefinitnek* nevezzük, ha nem pozitív definit, de $q(\boldsymbol{x}) \geq 0$.

d) *negatív szemidefinitnek* nevezzük, ha nem negatív definit, de $q(\boldsymbol{x}) \leq \leq 0$.

e) *indefinitnek* nevezzük, ha $q(\boldsymbol{x})$ pozitív és negatív értéket is felvesz.

A $q(\boldsymbol{x})$-hez tartozó $\boldsymbol{A}$ szimmetrikus mátrixra is ugyanezeket az elnevezéseket használjuk.

A kvadratikus alakok osztályozása a matematika számos területén felhasználásra kerül, a geometrián kívül például a többváltozós függvények konvexitásánál, az analízisben a szélsőérték számításnál. Az alábbiakban a definitség eldöntésére adunk szabályokat.

Tiszta négyzetes alakról a definit tulajdonság közvetlenül látszik. Pozitív definit, ha minden együttható pozitív, negatív definit, ha minden együttható negatív. Pozitív szemidefinit, ha pozitív és nulla együtthatók vannak, negatív szemidefinit, ha negatív és nulla együtthatók vannak, indefinit, ha pozitív és negatív együttható egyaránt előfordul. Mivel a tiszta négyzetes alak együtthatói a sajátértékek, ugyanezt el lehet mondani a sajátértékek vonatkozásában.

A sajátértékek meghatározása azonban elég nehéz és munkaigényes feladat, ennél egyszerűbb eldöntési szabályokra van szükség. Nevezzük a másodfokú alak *karakterisztikus sorozatának* az $\boldsymbol{A}$ mátrix bal felső sarokaldeterminánsaiból álló

$$A_0 = 1,\ A_1 = a_{11},\ A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix},\ A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix},$$

$$\ldots, A_n = |\boldsymbol{A}|$$

számsorozatot.

6.9.7. Tétel. $\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{x} \rangle$ akkor és csak akkor pozitív definit, ha a karakterisztikus sorozata pozitív elemekből áll. Akkor és csak akkor negatív definit, ha a karakterisztikus sorozat nullát nem tartalmaz és váltakozó előjelű.

Bizonyítás. Elég az első állítást bizonyítani, mert, ha $\boldsymbol{A}$ negatív definit, akkor $-\boldsymbol{A}$ pozitív definit, és a karakterisztikus sorozat váltakozó előjelűből pozitívvá válik.

$\Rightarrow$ Jelöljük $q_m(\tilde{\boldsymbol{x}})$-szel az első m változóval képezett kvadratikus alakot:

$$q_m(\tilde{\boldsymbol{x}}) = \sum_{i,j=1}^{m} a_{ij} x_i x_j.$$

$q_m(\tilde{\boldsymbol{x}})$ nyilván pozitív definit, hiszen $\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{x}\rangle$-ből a felesleges koordináták helyére nullát helyettesítve kapható meg. Pozitív definit mátrix determinánsa a sajátértékeinek a szorzata, tehát szintén pozitív, így $A_m > 0$, amit bizonyítani akartunk.

$\Leftarrow$ A változók száma szerinti teljes indukcióval bizonyítjuk. A karakterisztikus sorozat A_k determinánsának megfelelő kvadratikus alakot jelöljük $q_k(\boldsymbol{x})$-szel. $q_1(\boldsymbol{x}) = a_{11}x_1^2$, és valóban, ha $A_1 = a_{11} > 0$, akkor q_1 pozitív definit. Tegyük fel, hogy $q_k(\boldsymbol{x})$ pozitív definit, és $A_{k+1} > 0$. A $q_{k+1}(\boldsymbol{x})$ kvadratikus alak a $\boldsymbol{C} = (c_{ij})$ ortogonális transzformációval tiszta négyzetes alakra transzformálható, azaz $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}$ jelöléssel ($\boldsymbol{x} = (x_1, x_2, \ldots, x_{k+1})$, $\boldsymbol{y} = (y_1, y_2, \ldots, y_{k+1})$), $q_{k+1}(\boldsymbol{C}^*\boldsymbol{y}) = \sum_{i=1}^{k+1} \lambda_i y_i^2$, ahol a λ_i számok a q_{k+1} mátrixának a sajátértékei.

Indirekt módon feltételezzük, hogy a $\lambda_1, \lambda_2, \ldots, \lambda_{k+1}$ számok nem mind pozitívok, mivel a szorzatuk $A_{k+1} > 0$, legalább két negatívnak kell lennie köztük. Mivel a $\boldsymbol{C}$ transzformáció alkalmasan megválasztható, feltehetjük, hogy $\lambda_k < 0$ és $\lambda_{k+1} < 0$.

Válasszuk meg az α és a β számokat úgy, hogy $\alpha c_{kk+1} + \beta c_{k+1k+1} = 0$ legyen, de ne legyen mindkettő nulla, majd oldjuk meg a

$$\boldsymbol{C}\boldsymbol{x} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \alpha \\ \beta \end{pmatrix}$$

egyenletrendszert, melynek megoldása

$$\boldsymbol{x}_0 = \boldsymbol{C}^* \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \alpha \\ \beta \end{pmatrix}.$$

Az $\boldsymbol{x}_0$ $(k+1)$-edik koordinátája $x_{k+1} = \alpha c_{kk+1} + \beta c_{k+1k+1} = 0$ lesz az α és a β választása miatt. Helyettesítsünk be $q_{k+1}(\boldsymbol{x})$-be $\boldsymbol{x}_0$-t, akkor mivel $x_{k+1} =$

$= 0$, a $q_k(\boldsymbol{x})$ helyettesítési értékét kapjuk, ami a feltevés szerint pozitív. A transzformált alakba helyettesítve azonban $\lambda_k\alpha^2+\lambda_{k+1}\beta^2 < 0$-t kapunk, ami ellentmondás. Ennek következtében a $\lambda_1, \lambda_2, \ldots, \lambda_{k+1}$ számok pozitívok, és így $q_{k+1}(\boldsymbol{x})$ pozitív definit. □

A szemidefinit alak eldöntéséhez segédtételre van szükség.

6.9.8. Segédtétel. Ha az $\boldsymbol{A}$ szimmetrikus mátrix rangja r, akkor van a főátlóra szimmetrikusan elhelyezkedő $r \times r$-es teljesrangú részmátrixa.

Bizonyítás. Válasszunk ki r darab lineárisan független sort. A könnyebb jelölés kedvéért tegyük fel, hogy ez az első r darab sor (sor- és oszlopcserékkel és a szimmetria megtartásával ez elérhető). A szimmetria miatt az első r oszlop is lineárisan független maximális vektorrendszert alkot, a $k > r$ indexű oszlopok ezektől lineárisan függők. A $k > r$ indexű sorok elhagyásával ez a függőség változatlanul fennáll. Ha az $r \times n$-es részmátrixban az első r oszlop lineárisan összefüggő lenne, és a többi oszlop ezektől függő, akkor az $r \times n$-es részmátrix rangja r-nél kisebb lenne, de ez lehetetlen, hiszen r lineárisan független sora van (rangszám-tétel). Az első r oszlop tehát lineárisan független, így az első r sorból és oszlopból készített részmátrix teljesrangú. □

A szemidefinit kvadratikus alakoknál mindig van nulla sajátérték, tehát $\boldsymbol{A}$ rangja n-nél kisebb. Legyen $\operatorname{rang}(\boldsymbol{A}) = r$, akkor a segédtétel alapján kiválasztható olyan, a főátlóra szimmetrikusan elhelyezkedő $r \times r$-es részmátrix, melynek rangja r. Ehhez a részmátrixhoz tartozó, r változós kvadratikus alak típusa dönti el az eredeti kvadratikus alak típusát. A részmátrix kiválasztása nem egyértelmű, de bármely választás kötelezően ugyanazt az eredményt adja.

6.9.9. Tétel. Az $\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{x} \rangle$ kvadratikus alak akkor és csak akkor pozitív szemidefinit, ha $\operatorname{rang}(\boldsymbol{A}) = r < n$, és a főátlóra szimmetrikusan elhelyezkedő teljesrangú $r \times r$-es részmátrixhoz pozitív definit $q_r(\boldsymbol{x})$ kvadratikus alak tartozik. Negatív szemidefinit, ha a részmátrixhoz negatív definit kvadratikus alak tartozik.

Bizonyítás. Elég a pozitív szemidefinit esetet bizonyítani, mert ha $q(\boldsymbol{x})$ negatív szemidefinit, akkor $-q(\boldsymbol{x})$ pozitív szemidefinit.

$\Rightarrow$ Ha $q(\boldsymbol{x})$ szemidefinit, akkor van nulla sajátértéke, tehát $\boldsymbol{A}$ determinánsa nulla, ezért rangja n-nél kisebb. Ha $q(\boldsymbol{x})$ pozitív szemidefinit, akkor minden helyettesítési értéke nagyobb vagy egyenlő nullánál, ez igaz akkor is, ha a részmátrixnál nem választott változók értékének nullát adunk. Ekkor $q_r(\boldsymbol{x})$ helyettesítési értékeit kapjuk, $q_r(\boldsymbol{x})$ tehát pozitív szemidefinit. De szemidefinit nem lehet, mert mátrixa teljesrangú, tehát pozitív definit.

$\Leftarrow$ Tegyük fel most, hogy $q_r(\boldsymbol{x})$ pozitív definit, és, a könnyebb jelölés kedvéért, a változói az $x_1, x_2, \ldots, x_r$ legyenek. Az $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}$ ortogonális transzformációval hozzuk $q(\boldsymbol{x})$-et tiszta négyzetes alakra. A tiszta négyzetes alak együtthatói a sajátértékek, melyek közül csak r darab nem nulla. A $\boldsymbol{C}$-t úgy válasszuk meg, hogy az első r együttható legyen nullától különböző:

$$q(\boldsymbol{C}^*\boldsymbol{y}) = \sum_{i=1}^{r} \lambda_i y_i^2.$$

Bizonyítandó, hogy a $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_r$ számok pozitívok. Nulla nem lehet köztük, tegyük fel indirekt módon, hogy van köztük negatív, az általánosság megszorítása nélkül feltehető, hogy $\lambda_r < 0$. Az $\boldsymbol{y} = \boldsymbol{Cx}$ kifejezésben helyettesítsünk $x_{r+1} = x_{r+2} = ... = x_n = 0$-t, akkor $q(\boldsymbol{x})$ a q_r helyettesítési értékét adja meg. Válasszuk meg az $x_1, x_2, ..., x_r$ számokat úgy, hogy $y_1 = y_2 = ... = y_{r-1} = 0$ legyen. Mivel ez $r - 1$ darab r változós homogén lineáris egyenlet teljesülését kívánja meg, ennek mindig van nemtriviális megoldása. Ilyen választás mellett

$$q(\boldsymbol{x}) = \lambda_r y_r^2 \leq 0$$

a λ_r-re tett feltétel miatt. De ez a $q_r(\boldsymbol{x})$ helyettesítési értéke $\boldsymbol{x} \neq \boldsymbol{0}$ pontban, ami ellentmond $q_r(\boldsymbol{x})$ pozitív definitásának. $\square$

16. Kérdés. Mutassuk meg, hogy $\boldsymbol{A}^*\boldsymbol{A}$ pozitív definit vagy pozitív szemidefinit mátrix aszerint, hogy $\boldsymbol{A}$ teljesrangú mátrix vagy nem. Igaz-e, hogy minden pozitív definit vagy pozitív szemidefinit szimmetrikus mátrix $\boldsymbol{A}^*\boldsymbol{A}$ alakban írható fel?

17. Kérdés. Igaz-e, hogy egy vetítés akkor és csak akkor ortogonális, ha mátrixa pozitív szemidefinit?

6.10. Lineáris egyenletrendszer megoldása iterációval

Az alábbi, nagyon könnyen előállítható iteráció is ötletet ad az egyenletrendszer numerikus megoldására. Írjuk át az $\boldsymbol{Ax} = \boldsymbol{b}$ egyenletrendszert $\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} + \boldsymbol{b}$ alakba, és képezzük az

$$\boldsymbol{x}_{n+1} = (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x}_n + \boldsymbol{b}$$

iterációt valamely $\boldsymbol{x}_1$ vektorbólkiindulólag. Ha az eljárás konvergens, akkor $\boldsymbol{x} = \lim \boldsymbol{x}_n$ kielégíti az $\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} + \boldsymbol{b}$ egyenletet, tehát az egyenletrendszer megoldását adja. Tegyük fel a továbbiakban mindig, hogy $\boldsymbol{A}$ négyzetes mátrix és az egyenletrendszer megoldható. Ha $\boldsymbol{x}$ jelöli a megoldást, akkor a két egyenlet kivonásából, majd a kapott egyenlet ismételt alkalmazásából az

$$\boldsymbol{x}_n-\boldsymbol{x}=(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{x}_{n-1}-\boldsymbol{x})=(\boldsymbol{I}-\boldsymbol{A})^2(\boldsymbol{x}_{n-2}-\boldsymbol{x})=...=(\boldsymbol{I}-\boldsymbol{A})^n(\boldsymbol{x}_0-\boldsymbol{x})$$

összefüggést kapjuk. Tetszőleges $\boldsymbol{x}_0$-ból kiindulva az eljárás akkor és csak akkor konvergens, ha $(\boldsymbol{I}-\boldsymbol{A})^n \to 0$, vagyis, ha $\boldsymbol{I}-\boldsymbol{A}$ spektrál sugara 1-nél kisebb. Ez azonban nem mindig teljesül, tehát a módszer javítására van szükség.

Ha az iteráció minden $\boldsymbol{x}_0$-ra konvergens, akkor az egyenletnek csak egyetlen megoldása lehet, tehát $\boldsymbol{A}$ szükségképpen invertálható mátrix.

Ha λ_1, λ_2, ..., λ_n jelöli $\boldsymbol{A}$ sajátértékeit, akkor $\boldsymbol{I}-\boldsymbol{A}$ sajátértékei $1-\lambda_1$, $1-\lambda_2$, ..., $1-\lambda_n$, vagyis az $\boldsymbol{x}_{n+1}=(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_n+\boldsymbol{b}$ iteráció akkor konvergens, ha $\boldsymbol{A}$ sajátértékei az 1 pont körüli egységsugarú kör belsejébe esnek. Ha az egyenletet előbb c-vel beszorozzuk, és ezután hajtjuk végre a fenti eljárást, akkor a sajátértékek előzőleg c-vel beszorzódnak, tehát az előző kritérium elérhető mindig, ha minden i-re $\mathrm{Re}\lambda_i > 0$.

Ahhoz, hogy $\mathrm{Re}\lambda_i > 0$-t elérjük, szorozzuk meg az egyenletet $\boldsymbol{A}^*$-gal! Így kapjuk a következő állítást. A c szám megválasztásához az $||\boldsymbol{A}||$ felülről egyszerűen megbecsülhető.

6.10.1. Tétel. Ha $\boldsymbol{A}$ invertálható mátrix, és az $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ egyenletrendszer megoldását $\boldsymbol{x}_0$-lal jelöljük, akkor $c<\frac{2}{||\boldsymbol{A}||^2}$ esetén az

$$\boldsymbol{x}_{n+1}=(\boldsymbol{I}-c\boldsymbol{A}^*\boldsymbol{A})\boldsymbol{x}_n+c\boldsymbol{A}^*\boldsymbol{b}$$

iteráció konvergens és $\boldsymbol{x}_n \to \boldsymbol{x}_0$ bármely $\boldsymbol{x}_1$ kiinduló vektor esetén.

Bizonyítás. Mivel $\boldsymbol{A}$ invertálható, az $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^*\boldsymbol{b}$ megoldása azonos az $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ megoldásával. A $\boldsymbol{B}=\boldsymbol{A}^*\boldsymbol{A}$ mátrix pozitív definit (ld. 6.9. 16. Kérdés) és szimmetrikus, így minden sajátértéke pozitív. Jelölje $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ a $\boldsymbol{B}$ sajátértékeit, akkor $\boldsymbol{I}-c\boldsymbol{A}^*\boldsymbol{A}$ sajátértékei az $1-c\lambda_i$ számok lesznek. Mivel $0 < c\lambda_i \le c||\boldsymbol{A}||^2 < 2$, kapjuk, hogy $|1-c\lambda_i| < 1$, vagyis a 6.8.1. Tétel következtében $(\boldsymbol{I}-c\boldsymbol{A}^*\boldsymbol{A})^n \to 0$. Így a fejezet elején elmondott bizonyítás $\boldsymbol{A}$ helyett $c\boldsymbol{B}$-re megismételhető, vagyis az iteráció konvergens és a megoldáshoz konvergál. □

Megkérdezhetjük még, hogy milyen c érték mellett lesz a leggyorsabb a konvergencia. A 6.8.1. Tétel alapján a konvergencia sebessége exponenciális, ahol az exponenciális kifejezés alapja $q=\max(|1-c\lambda_i|)$. Jelen esetben $q=\max(1-c\lambda_1,\ c\lambda_n-1)$. q akkor minimális, azaz a konvergencia akkor a leggyorsabb, ha $c=\frac{2}{\lambda_1+\lambda_n}$, ekkor $q=\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}$. Röviden szólva a két szélső sajátérték ismeretében optimalizálni tudjuk a konvergencia sebességét.

6.11. Még egyszer az ortogonalizációról

Ha adott az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r$ lineárisan független vektorrendszer, akkor lineáris transzformációk sorozatával ortogonális és egységvektorokból álló vektorrendszert tudtunk készíteni a Gram–Schmidt-féle ortogonalizációs eljárással (1.4.4. és 1.5.2. Tétel). Kérdés, hogy ez a rekurzív eljárás egyetlen lépésben is elvégezhető-e, explicite megadható-e olyan $\boldsymbol{B}$ lineáris transzformáció, hogy a $\boldsymbol{Ba}_1, \boldsymbol{Ba}_2, \ldots, \boldsymbol{Ba}_r$ vektorok egységvektorok és ortogonálisak?

A kérdés összefügg az ortogonális vetítés megadásának a kérdésével (3.5.10. Tétel), hiszen ortogonális vektorok esetén az általuk kifeszített altérre történő ortogonális vetítés igen egyszerűen megadható, amint az az alábbi tételből látható.

6.11.1. Tétel. Ha az $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r$ vektorok egységvektorok és ortogonálisak, és

$$\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r),$$

akkor a vektorok által kifeszített altérre (vagyis $\boldsymbol{A}$ képterére) történő ortogonális vetítés mátrixa $\boldsymbol{AA}^*$.

Bizonyítás. A vektorokra tett feltételek alapján $\boldsymbol{a}_i\boldsymbol{a}_j = 0$, ha $i \neq j$, és $\boldsymbol{a}_i\boldsymbol{a}_j = 1$, ha $i = j$, és ennek alapján $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I}_r$, ahol $\boldsymbol{I}_r$ az r-dimenziós egységmátrix.

Az $\boldsymbol{AA}^*$ transzformáció vetítés, mert

$$(\boldsymbol{AA}^*)^2 = \boldsymbol{A}(\boldsymbol{A}^*\boldsymbol{A})\boldsymbol{A}^* = \boldsymbol{AI}_r\boldsymbol{A}^* = \boldsymbol{AA}^*.$$

A vetítés ortogonális, mert $\boldsymbol{AA}^*$ szimmetrikus mátrix (ld. 3.4.5. Tétel): $(\boldsymbol{AA}^*)^* = \boldsymbol{A}^{**}\boldsymbol{A}^* = \boldsymbol{AA}^*$.

$\boldsymbol{AA}^*$ képtere R_A, mert egyrészt triviálisan $R_{AA^*} \subset R_A$, másrészt $R_A = R_{AA^*A} \subset R_{AA^*}$. $\square$

Adjuk meg az $\boldsymbol{A}$ mátrixot az $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_0 \\ \boldsymbol{V} \end{pmatrix}$, alakban, ahol az $\boldsymbol{A}$ oszlopvektorai lineárisan függetlenek és $\boldsymbol{A}_0$ teljesrangú négyzetes mátrix. Az általános eset sorcserékkel mindig ilyen alakra hozható. Ha keressük az $\boldsymbol{A}$ képterére ortogonálisan vetítő mátrixot, a 3.5.10. Tétel mellett másik megoldásnak kínálkozik ortogonalizálás után a 6.11.1. Tétel alkalmazása. Az ortogonalizálás történhet a Gram–Schmidt-eljárással, vagy egy alkalmas mátrixszal történő jobb oldali szorzással. Ez utóbbi utat követjük az alábbiakban.

Az alábbiakban fontos szerepet fog játszani a 3.5.10. Tétel $\boldsymbol{C}$ és $\boldsymbol{D}$ mátrixa, melyek definíciója:

$\boldsymbol{C} = \boldsymbol{VA}_0^{-1}$ és $\boldsymbol{D} = (\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})^{-1}$.

Néhány apró, egyszerű segédtétellel közelítjük meg a problémát.

6.11.2. Segédtétel. $\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C}$ pozitív definit mátrix.

Bizonyítás. A hozzátartozó kvadratikus alak

$$\langle \boldsymbol{x}, (\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})\boldsymbol{x}\rangle = ||\boldsymbol{x}||^2 + \langle \boldsymbol{x}, \boldsymbol{C}^*\boldsymbol{C}\boldsymbol{x}\rangle =$$
$$= ||\boldsymbol{x}||^2 + \langle \boldsymbol{C}\boldsymbol{x}, \boldsymbol{C}\boldsymbol{x}\rangle = ||\boldsymbol{x}||^2 + ||\boldsymbol{C}\boldsymbol{x}||^2 > 0,$$

ha $\boldsymbol{x} \neq \boldsymbol{0}$, ami a pozitív definit tulajdonságot jelenti. □

6.11.3. Segédtétel. Pozitív definit mátrix inverze is pozitív definit.

Bizonyítás. Ha $\boldsymbol{G}$ pozitív definit, akkor $\langle \boldsymbol{x}, \boldsymbol{G}^{-1}\boldsymbol{x}\rangle = \langle \boldsymbol{G}\boldsymbol{y}, \boldsymbol{y}\rangle > 0$, ha $\boldsymbol{y} \neq \boldsymbol{0}$, vagyis, ha $\boldsymbol{x} \neq \boldsymbol{0}$. □

A 16. Kérdésben láthattuk, hogy minden szimmetrikus, pozitív definit $\boldsymbol{G}$ mátrix felírható $\boldsymbol{G} = \boldsymbol{H}\boldsymbol{H}^*$ alakban. Ez a felírás természetesen nem egyértelmű, hiszen ha $\boldsymbol{K}$ tetszőleges ortogonális mátrix, akkor $(\boldsymbol{H}\boldsymbol{K})(\boldsymbol{H}\boldsymbol{K})^* = \boldsymbol{H}\boldsymbol{K}\boldsymbol{K}^{-1}\boldsymbol{H}^* = \boldsymbol{H}\boldsymbol{H}^* = \boldsymbol{G}$. Más lehetőség azonban nincs, a felbontás tényezői jobb oldali ortogonális mátrixszorzótól eltekintve egyértelműek.

6.11.4. Segédtétel. Ha $\boldsymbol{G}$ pozitív definit és $\boldsymbol{G} = \boldsymbol{H}\boldsymbol{H}^*$ és $\boldsymbol{G} = \boldsymbol{H}_1\boldsymbol{H}_1{}^*$, akkor van olyan $\boldsymbol{K}$ ortogonális mátrix, hogy $\boldsymbol{H}_1 = \boldsymbol{H}\boldsymbol{K}$.

Bizonyítás. Mivel $\boldsymbol{G}$ pozitív definit, teljesrangú, ezért $\boldsymbol{H}_1$ is az, és így $\boldsymbol{H}_1^{-1}$ létezik. A feltétel alapján $\boldsymbol{H}\boldsymbol{H}^* = \boldsymbol{H}_1\boldsymbol{H}_1{}^*$, vagy átalakítva $\boldsymbol{H}^{-1}\boldsymbol{H}_1\boldsymbol{H}_1{}^*\boldsymbol{H}^{-1*} = (\boldsymbol{H}^{-1}\boldsymbol{H}_1)(\boldsymbol{H}^{-1}\boldsymbol{H}_1)^* = \boldsymbol{I}$. A $\boldsymbol{K} = \boldsymbol{H}^{-1}\boldsymbol{H}_1$ mátrix tehát ortogonális, és $\boldsymbol{H}_1 = \boldsymbol{H}\boldsymbol{K}$. □

Használjuk a fejezet elején bevezetett jelöléseket, legyen $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_0 \\ \boldsymbol{V} \end{pmatrix}$, $\boldsymbol{C} = \boldsymbol{V}\boldsymbol{A}_0^{-1}$ és $\boldsymbol{D} = (\boldsymbol{I} + \boldsymbol{C}^*\boldsymbol{C})^{-1}$.

6.11.5. Tétel. Adott $\boldsymbol{A}$ mellett vegyük a $\boldsymbol{D}$ mátrix tetszőleges $\boldsymbol{D} = \boldsymbol{F}\boldsymbol{F}^*$ felbontását. Az $\boldsymbol{A}\boldsymbol{B}$ mátrix oszlopvektorai akkor és csak akkor ortogonális egységvektorok, ha $\boldsymbol{B} = \boldsymbol{A}_0^{-1}\boldsymbol{F}$.

Bizonyítás. $\Leftarrow$ Az $\boldsymbol{B}$ adott választása mellett $\boldsymbol{A}\boldsymbol{B} = \begin{pmatrix} \boldsymbol{A}_0 \\ \boldsymbol{V} \end{pmatrix} \boldsymbol{A}_0^{-1}\boldsymbol{F} = \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{V}\boldsymbol{A}_0^{-1}\boldsymbol{F} \end{pmatrix} = \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{C}\boldsymbol{F} \end{pmatrix}$. Mivel

$$(\boldsymbol{F}^* \;\; \boldsymbol{F}^*\boldsymbol{C}^*) \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{C}\boldsymbol{F} \end{pmatrix} = \boldsymbol{F}^*\boldsymbol{F} + \boldsymbol{F}^*\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{F} = \boldsymbol{F}^*(\boldsymbol{I}_r + \boldsymbol{C}^*\boldsymbol{C})\boldsymbol{F} =$$

$$= \boldsymbol{F}^*\boldsymbol{D}^{-1}\boldsymbol{F} = \boldsymbol{F}^*\boldsymbol{F}^{*^{-1}}\boldsymbol{F}^{-1}\boldsymbol{F} = \boldsymbol{I}_r$$

az $\boldsymbol{A}\boldsymbol{B}$ oszlopvektorai ortogonális egységvektorok.

$\Rightarrow$ Ha valamilyen $\boldsymbol{F}$-re $\boldsymbol{AB}$ oszlopvektorai ortogonális egységvektorok, akkor

$$\boldsymbol{I}_r = (\boldsymbol{F}^* \quad \boldsymbol{F}^*\boldsymbol{C}^*)\begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{CF} \end{pmatrix} = \boldsymbol{F}^*\boldsymbol{F} + \boldsymbol{F}^*\boldsymbol{C}^*\boldsymbol{CF} =$$

$$= \boldsymbol{F}^*(\boldsymbol{I}_r + \boldsymbol{C}^*\boldsymbol{C})\boldsymbol{F} = \boldsymbol{F}^*\boldsymbol{D}^{-1}\boldsymbol{F},$$

és ebből $\boldsymbol{D}^{-1} = \boldsymbol{F}^{-1*}\boldsymbol{F}^{-1}$, azaz $\boldsymbol{D} = \boldsymbol{FF}^*$. $\square$

Az egy lépésben történő ortogonalizáció szép, de az $\boldsymbol{F}$ mátrix előállítása nem egyszerű, legalábbis a 16. Kérdésben megadott módon. Itt adunk egy másik előállítást, azonban ez is csak lépésenként elvégezhető konstrukció, kísértetiesen hasonlít a kézzel történő gyökvonásra.

Képezzük a $\boldsymbol{D}$ pozitív definit és szimmetrikus mátrixból a hozzátartozó $q(\boldsymbol{x}) = \langle \boldsymbol{x},\, \boldsymbol{Dx}\rangle$ kvadratikus alakot, ami a $\boldsymbol{D} = \boldsymbol{F}^*\boldsymbol{F}$ faktorizáció után a $q(\boldsymbol{x}) = \langle \boldsymbol{x},\, \boldsymbol{F}^*\boldsymbol{Fx}\rangle = \langle \boldsymbol{Fx},\, \boldsymbol{Fx}\rangle = ||\boldsymbol{Fx}||^2$ alakot ölti. Hozzuk $q(\boldsymbol{x})$-et az $||\boldsymbol{Fx}||^2$ alakra teljes négyzetre történő kiegészítéssel.

$\boldsymbol{D} = (d_{ij})$ jelölés mellett $d_{11} > 0$ a pozitív definit tulajdonság miatt, ezért

$$q(\boldsymbol{x}) = d_{11}(x_1^2 + 2x_1\sum_{i=2}^{n}\frac{d_{1i}}{d_{11}}x_i) + q_1(\tilde{\boldsymbol{x}}) = d_{11}(x_1 + \sum_{i=2}^{n}\frac{d_{i1}}{d_{11}}\,x_i)^2 + q_2(\tilde{\boldsymbol{x}}),$$

ahol $\tilde{\boldsymbol{x}} = (x_2, x_3, ..., x_n)$ és $q_2(\tilde{\boldsymbol{x}})$ $n-1$ változós pozitív definit kvadratikus alak. $q_2(\tilde{\boldsymbol{x}})$-re az eljárás folytatható és $n-1$ lépésben a

$$q(\boldsymbol{x}) = \sum_{i=1}^{n} y_i^2$$

tiszta négyzetes alakhoz jutunk, ahol $y_i = \sum\limits_{j=i}^{n} f_{ij}x_j$. Ezzel az $||\boldsymbol{Fx}||^2$ alakra hozás megtörtént és az $\boldsymbol{F}$ mátrix leolvasható: $\boldsymbol{F}$ felső háromszög mátrix, melynek nem nulla elemei az f_{ij} számok.

A módszer elsajátításához ajánljuk a következő gyakorlatot:

17. Kérdés. Melyek a $\boldsymbol{D}$ mátrix $\boldsymbol{D} = \boldsymbol{F}^*\boldsymbol{F}$ faktorizációjához tartozó tényezők, ha

$$\boldsymbol{D} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 8 & 10 \\ 2 & 10 & 16 \end{pmatrix}?$$

6.12. Válaszok a kérdésekre

1. Hány lineárisan független sajátvektora van az $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ mátrixnak?

A karakterisztikus egyenlet $(\lambda - 1)^3 = 0$, tehát $\lambda = 1$ háromszoros sajátérték, más sajátérték nincs. Az $\boldsymbol{x} = (x, y, z)$ sajátvektorokat az

$$\begin{aligned} x + y &= x \\ y - z &= y \\ z &= z \end{aligned}$$

egyenletrendszer adja. Ennek általános megoldása: $y = z = 0$, és x tetszőleges, ami a $(1, 0, 0)$ vektor többszöröseit jelenti. Más, konstans szorzótól eltekintve különböző megoldás nincs, tehát csak egy (lényegesen különböző) sajátvektor van.

2. Az $\boldsymbol{A}$ mátrix sajátvektorai $(2, 1, 1)$, $(1, -1, 1)$ és $(1, 1, 2)$, a hozzátartozó sajátértékek rendre 1, 2, és 1. Meghatározzák-e az adatok $\boldsymbol{A}$-t? Ha igen, számítsuk is ki!

A sajátvektorok meghatározzák a diagonalizálás $\boldsymbol{C}$ mátrixát, a sajátértékek a $\boldsymbol{D}$ diagonális mátrixot, az $\boldsymbol{A}$ mátrix ebből sorrendcserétől eltekintve egyértelműen meghatározható: $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{D}$, ebből $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{D}\boldsymbol{C}^{-1}$.

$$\boldsymbol{C} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \boldsymbol{C}^{-1} = \frac{1}{5}\begin{pmatrix} 3 & 1 & -2 \\ 1 & -3 & 1 \\ -2 & 1 & 3 \end{pmatrix}, \boldsymbol{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

Ebből

$$\boldsymbol{A} = \boldsymbol{C}\boldsymbol{D}\boldsymbol{C}^{-1} = \frac{1}{5}\begin{pmatrix} 6 & -3 & 1 \\ -1 & 8 & -1 \\ 1 & -3 & 6 \end{pmatrix}.$$

3. Igaz-e, hogy az ortogonális mátrixok minden valós vagy komplex sajátértékének az abszolút értéke 1?

Igen. Ha λ sajátérték és $\boldsymbol{v}$ a hozzátartozó sajátvektor, akkor $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, itt komplex sajátértéket és sajátvektort is megengedve. Az $\boldsymbol{A}$ komplex vektorokra is normatartó, mert $\boldsymbol{v} = \boldsymbol{x} + i\boldsymbol{y}$ esetén $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{A}\boldsymbol{x} + i\boldsymbol{A}\boldsymbol{y}$, és $||\boldsymbol{A}\boldsymbol{v}||^2 = ||\boldsymbol{A}\boldsymbol{x}||^2 + ||\boldsymbol{A}\boldsymbol{y}||^2 = ||\boldsymbol{x}||^2 + ||\boldsymbol{y}||^2 = ||\boldsymbol{v}||^2$. Az $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ egyenletből $||\boldsymbol{A}\boldsymbol{v}|| = |\lambda|||\boldsymbol{v}||$, és így $|\lambda| = 1$.

4. Mutassuk meg, hogy az ortogonális mátrixok esetén a különböző sajátértékekhez tartozó sajátvektorok ortogonálisak. Igaz-e ez komplex sajátértékekre, sajátvektorokra is?

Nézzük rögtön komplex környezetben. Tegyük fel, hogy λ sajátértékhez az $\boldsymbol{u}$ sajátvektor, a $\mu \neq \lambda$ sajátértékhez a $\boldsymbol{v}$ tartozik. Egyrészt $\langle \boldsymbol{Au},\, \boldsymbol{v}\rangle =$ $= \lambda\langle \boldsymbol{u},\, \boldsymbol{v}\rangle$, másrészt $\langle \boldsymbol{Au},\, \boldsymbol{v}\rangle = \langle \boldsymbol{u},\, \boldsymbol{A}^*\boldsymbol{v}\rangle = \langle \boldsymbol{u},\, \boldsymbol{A}^{-1}\boldsymbol{v}\rangle = \frac{1}{\bar{\mu}}\langle \boldsymbol{u},\, \boldsymbol{v}\rangle =$ $= \mu\langle \boldsymbol{u},\, \boldsymbol{v}\rangle$ (itt felhasználtuk, hogy $|\mu| = 1$, ld. 3. Kérdés). A két összefüggést összehasonlítva $\langle \boldsymbol{u},\, \boldsymbol{v}\rangle = 0$ adódik.

5. Igazoljuk, hogy az

$$\boldsymbol{A} = \begin{pmatrix} \cos\varphi & -\sin\varphi & \cdots & 0 \\ \sin\varphi & \cos\varphi & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

mátrixszal megadott forgatás (ld. 4.6.) sajátértékei az 1, $e^{i\varphi}$, $e^{-i\varphi}$ számok, ha $n \geq 3$. Igaz-e, hogy a komplex sajátértékekhez tartozó invariáns altér a forgatás „tengelyének" ortogonális kiegészítő altere?

A karakterisztikus egyenlet

$$(1-\lambda)^{n-2}[(\cos\varphi - \lambda)^2 + \sin^2\varphi] = (1-\lambda)^{n-2}(\lambda^2 - 2\lambda\cos\varphi + 1) = 0,$$

ennek megoldása $\lambda = 1$-en kívül $\lambda = \cos\varphi \pm i\sin\varphi = e^{\pm i\varphi}$. Az $e^{-i\varphi}$-hez tartozó komplex sajátvektor $(1,\, i,\, 0,\, \ldots,0)$, ennek valós és imaginárius része: $(1,\, 0,\, 0,\, \ldots,\, 0)$ valamint $(0,\, 1,\, 0,\, \ldots,\, 0)$ vektorok. A forgatás „tengelye" a 2-nél magasabb indexű koordináta egységvektorok által kifeszített altér, erre ezek a vektorok merőlegesek. Mivel $(n-2)$-dimenziós altér ortogonális kiegészítője 2-dimenziós, a $(1,\, 0,\, 0,\, \ldots,\, 0)$ és a $(0,\, 1,\, 0,\, \ldots,\, 0)$ vektorok által kifeszített altér az ortogonális kiegészítő altér.

6. Ha a mátrix nemnegatív elemekből áll, és minden sorában az elemek összege 1, akkor sztochasztikus mátrixnak nevezik (az ún. átmenet valószínűségek mátrixa). Igaz-e, hogy az $n \times n$-es $\boldsymbol{A}$ sztochasztikus mátrixnak $\lambda = 1$ mindig sajátértéke? Azt mondjuk, hogy az $\boldsymbol{Ax} = \boldsymbol{x}$ egyenlenek $\boldsymbol{x}$ sztochasztikus megoldása, ha $\boldsymbol{x}$ minden koordinátája nemnegatív és a koordináták összege 1. Igaz-e, hogy ha $\lambda = 1$ egyszeres sajátérték, akkor az $\boldsymbol{Ax} = \boldsymbol{x}$ egyenletnek csak triviális sztochasztikus megoldása van, azaz $\boldsymbol{x} = \frac{1}{n}\mathbf{1}$.

Az $\boldsymbol{A} - \boldsymbol{I}$ mátrixban a sorok összege 0, tehát, ha az első $n-1$ oszlopot hozzáadjuk az utolsó oszlophoz, az utolsó oszlop nullává válik, tehát

$\boldsymbol{A} - \boldsymbol{I}$ determinánsa nulla. Ez azt jelenti, hogy $\lambda = 1$ sajátérték. Az $\boldsymbol{x} = \frac{1}{n}\mathbf{1}$ triviálisan megoldás, triviálisan $\lambda = 1$-hez tartozó sajátvektor. Egyszeres sajátérték esetén viszont csak konstans szorzóban különbözhet bármely más sajátvektor. □

7. Diagonizálható-e az $\begin{pmatrix} 5 & -1 & -2 \\ 4 & 0 & -2 \\ 4 & -1 & -1 \end{pmatrix}$ mátrix? Ha igen, számoljuk is ki (jól használhatjuk az Excelt az 5.7.-ben leírtak szerint)! Ortogonális transzformációval diagonizálható-e?

A karakterisztikus egyenlete – átalakítások után $-(\lambda - 1)^2(\lambda - 2) = 0$, vagyis sajátértékei: 1, 1, 2. A $\lambda = 1$-hez tartozó két lineárisan független sajátvektor (1, 0, 2) és (1, 2, 1) (így választható), a $\lambda = 2$-hez tartozó sajátvektor (1, 1, 1). Mivel létezik három lineárisan független sajátvektor, a mátrix diagonizálható (6.1.4. Tétel).

A $\boldsymbol{C}$ mátrixot a sajátvektorok alkotják, számoljuk ki az inverzét:

$$\boldsymbol{C} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}, \boldsymbol{C}^{-1} = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 1 & 1 \\ 4 & -1 & -2 \end{pmatrix},$$

és a transzformációs képlet alapján

$$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Ortogonális transzformációval nem diagonizálható, mert nem szimmetrikus a mátrix (6.4.3. Tétel).

8. Igaz-e, hogy szimmetrikus mátrix csak ortogonális mátrixszal diagonizálható?

Nem igaz. Az $\boldsymbol{I}$ bármely teljesrangú $\boldsymbol{C}$ mátrixszal diagonizálható, hiszen $\boldsymbol{C}^{-1}\boldsymbol{I}\boldsymbol{C} = \boldsymbol{I}$. A példa túl triviális, de kevésbé triviális példa is adható. Legyen

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{pmatrix},$$

ahol $\boldsymbol{B}$ tetszőleges szimmetrikus mátrix. $\boldsymbol{C}_1$ legyen tetszőleges megfelelő méretű mátrix, $\boldsymbol{C}_2$ pedig $\boldsymbol{B}$-t diagonalizáló mátrix, akkor

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{C}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{C}_2 \end{pmatrix}$$

diagonalizálja $\boldsymbol{A}$-t és nem feltétlenül ortogonális mátrix.

9. $\mathbf{R}^3$-ban adható-e példa arra, hogy diszjunkt, generáló alterek direkt összegként nem állítják elő $\mathbf{R}^3$-at?

Vegyünk fel $\mathbf{R}^3$-ban három diszjunkt alteret, két egydimenziót és egy kétdimenziót. Az egyszerűség kedvéért, a koordináta egységvektorokat $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$-nak nevezve, legyen az egydimenziós alterek bázisa $\boldsymbol{i}$, ill. $\boldsymbol{j}$, a kétdimenzióé $\boldsymbol{k}$ és $\boldsymbol{i} + \boldsymbol{j}$. Bármely $\boldsymbol{x} \in \mathbf{R}^3$ felírható $\boldsymbol{x} = a\boldsymbol{i} + b\boldsymbol{j} + c\boldsymbol{k}$ alakban, ahol a három tag a három altér egy-egy eleme. De α bármely választása mellett felírható $\boldsymbol{x} = (a - \alpha)\boldsymbol{i} + (b - \alpha)\boldsymbol{j} + [c\boldsymbol{k} + \alpha(\boldsymbol{i} + \boldsymbol{j})]$ alakban is, ahol a három tag szintén a három altér egy-egy eleme. A felbontás tehát nem egyértelmű, vagyis nem direkt összegről van szó.

Más példa talán még egyszerűbb. Vegyünk fel $k > 3$ darab vektort, melyek generálják $\mathbf{R}^3$-at, és bármely kettő lineárisan független. Minden egyes vektor generál egy alteret, ezek az alterek diszjunktak. Mivel a vektorok generálják $\mathbf{R}^3$-at, minden $\boldsymbol{x} \in \mathbf{R}^3$ felírható ezen vektorok lineáris kombinációjaként, de a felírás akkor és csak akkor egyértelmű, ha a vektorok lineárisan függetlenek, ami $k > 3$ esetén lehetetlen.

10. Kérdés. Igaz-e, hogy ha $\boldsymbol{v}$ adott sajátvektor, akkor azon $\boldsymbol{x}$ elemek halmaza, melyek gyökere $\boldsymbol{v}$, a $\mathbf{0}$-ral kiegészítve $\boldsymbol{B}$-re invariáns alteret ad?

A kérdés becsapós. Nyilván invariáns halmazt ad, de nem alteret! $\boldsymbol{a}$ legyen $p > 2$ indexű vektor, melynek a gyökere $\boldsymbol{v}$, és $\boldsymbol{v}_1$ legyen $\boldsymbol{v}$-től lineárisan független sajátvektor. Ekkor $\boldsymbol{a}$ is, és $\boldsymbol{a} + \boldsymbol{v}_1$ is eleme a halmaznak, hiszen mindkettő gyökere $\boldsymbol{v}$: $\boldsymbol{B}^{p-1}(\boldsymbol{a} + \boldsymbol{v}_1) = \boldsymbol{B}^{p-1}\boldsymbol{a} = v$. A két vektor különbsége $\boldsymbol{v}_1$ azonban nem eleme a halmaznak.

11. Kérdés. Igaz-e, hogy az $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ mátrix tükröző transzformációval Jordan-alakra hozható? Határozzuk meg a másodlagos sajátvektorait!

A karakterisztikus egyenlet alapján $\lambda = 1$ háromszoros sajátérték. Az $\boldsymbol{A}$ mátrix másodlagos sajátvektorai $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$, hiszen $\boldsymbol{A}\boldsymbol{e}_1 = \boldsymbol{e}_1$, $\boldsymbol{A}\boldsymbol{e}_2 = \boldsymbol{e}_2 + \boldsymbol{e}_1$, $\boldsymbol{A}\boldsymbol{e}_3 = \boldsymbol{e}_3 - \boldsymbol{e}_2$. Ha az $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ bázist átalakítjuk úgy, hogy $\boldsymbol{e}_1$-et és $\boldsymbol{e}_2$-t megtartva $\boldsymbol{e}_3$ helyett $-\boldsymbol{e}_3$-at választunk, akkor $\boldsymbol{A}\boldsymbol{e}_1 = \boldsymbol{e}_1$, $\boldsymbol{A}\boldsymbol{e}_2 = \boldsymbol{e}_2 + \boldsymbol{e}_1$, $\boldsymbol{A}(-\boldsymbol{e}_3) = (-\boldsymbol{e}_3) + \boldsymbol{e}_2$. Ennek a transzformációnak a mátrixa, melyet tükrözéssel értünk el, a $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ Jordan-alakú mátrix.

12. Igaz-e, hogy minden vetítésnek (a $\mathbf{0}$-t és $\boldsymbol{I}$-t itt nem tekintve vetítésnek) pontosan két különböző sajátértéke van? Igaz-e, hogy minden vetítés normája 1?

A vetítés definíciója: $\boldsymbol{A}^2 = \boldsymbol{A}$. A λ sajátértékre és $\boldsymbol{v}$ sajátvektorra $\boldsymbol{Av} = \lambda\boldsymbol{v}$, szorozzuk ezt $\boldsymbol{A}$-val balról: $\boldsymbol{A}^2\boldsymbol{v} = \lambda\boldsymbol{Av}$, vagyis $\boldsymbol{Av} = \lambda^2\boldsymbol{v}$. A két összefüggésből $\lambda = \lambda^2$, vagyis $\lambda = 0$ vagy 1. A magtér elemeire $\lambda = 0$, a vetítés képterének az elemire $\lambda = 1$ teljesül.

Más meggondolással, minden vetítés a 4.5. fejezetben elmondottak szerint a $\boldsymbol{B}_0$-lal jelölt kanonikus alakra hozható, amiből látható, hogy sajátértékei 0 és 1.

Az ortogonális vetítés normája 1, mert a Pythagoras-tételből (3.4.5. Tétel) következik, hogy $||\boldsymbol{Ax}|| \leq ||\boldsymbol{x}||$, és a vetítés képterének az elemeire az egyenlőség teljesül.

Ha $\boldsymbol{A}$ nem ortogonális vetítés, akkor vezessük be az alábbi jelöléseket. Legyen $\boldsymbol{A}$ képtere R, magtere M és képezzük a magtér ortogonális kiegészítő alterét $M^\perp$-t. Válasszunk tetszőlegesen egy $\boldsymbol{x} \in M^\perp$, $\boldsymbol{x} \neq 0$ vektort, ennek $\boldsymbol{A}$ szerinti vetülete legyen $\boldsymbol{y} = \boldsymbol{Ax}$. Ha $\boldsymbol{y}$-t ortogonálisan vetítjük $M^\perp$-re, akkor $\boldsymbol{x}$-et kapunk, hiszen $\boldsymbol{Ax} - \boldsymbol{x}$, az $\boldsymbol{A}$ vetítés vetítősugara, mindig M-be tartozik, ezért ortogonális $M^\perp$-re, tehát tényleg ortogonális vetületről van szó. A Pythagoras-tételt (3.4.5. Tétel) alkalmazva alkalmas $\boldsymbol{x}$ vektorra $||\boldsymbol{Ax}|| > ||\boldsymbol{x}||$, hiszen $\boldsymbol{Ax} - \boldsymbol{x} \neq 0$ elérhető, mert az $\boldsymbol{A}$ vetítés a feltétel szerint nem ortogonális. Ebből látható, hogy $||\boldsymbol{A}|| > 1$. Kimondhatjuk tétel formájában is:

Tétel. Az $\boldsymbol{A}$ vetítés akkor és csak akkor ortogonális, ha $||\boldsymbol{A}|| = 1$.

Még tovább léphetünk. A vetítés α szögét is definiálhatjuk a $\sin\alpha = \frac{1}{||\boldsymbol{A}||}$ képlettel. (Rajzoljuk fel három dimenzióban, és láthatjuk, hogy két sík ($M^\perp$ és R) szögét pontosan így definiáltuk.)

13. Megegyeznek-e $\boldsymbol{AB}$ és $\boldsymbol{BA}$ sajátértékei? A sajátértékek multiplicitása is megegyezik? Igaz-e, hogy $||\boldsymbol{A}|| = ||\boldsymbol{A}^*||$?

A nem nulla sajátértékek egyezése könnyen bizonyítható. Legyen $\lambda \neq$ $\neq 0$ $\boldsymbol{AB}$-nek sajátértéke, és $\boldsymbol{v}$ legyen a hozzátartozó sajátvektor, akkor $\boldsymbol{ABv} = \lambda\boldsymbol{v}$. Szorozzuk meg balról mindkét oldalt $\boldsymbol{B}$-vel, akkor $\boldsymbol{BABv} =$ $= \lambda\boldsymbol{Bv}$, vagyis $\boldsymbol{Bv}$ a $\boldsymbol{BA}$-nak sajátvektora ugyanazon λ sajátérték mellett. $\boldsymbol{Bv} = \mathbf{0}$ nem lehet, mert akkor $\boldsymbol{ABv} = \lambda\boldsymbol{v}$ alapján $\boldsymbol{v}$ is nulla lenne.

A további vizsgálat egy kicsit bonyolultabb. Tegyük fel először, hogy $|\boldsymbol{B}| \neq 0$. Tekintsük a $|\boldsymbol{BAB} - \lambda\boldsymbol{B}| = 0$ λ-ra vonatkozó egyenletet. A $\boldsymbol{B}$ kiemelhető jobbról is és balról is, ezért a fenti egyenlet átalakítható $|\boldsymbol{B}| \cdot |\boldsymbol{AB} - \lambda\boldsymbol{I}| = 0$, illetve $|\boldsymbol{BA} - \lambda\boldsymbol{I}| \cdot |\boldsymbol{B}| = 0$ alakba, ami azt jelenti, hogy a $|\boldsymbol{AB} - \lambda\boldsymbol{I}|$ és a $|\boldsymbol{BA} - \lambda\boldsymbol{I}|$ karakterisztikus polinomok azonosak.

Ha $|\boldsymbol{B}| = 0$, akkor „perturbálással” bizonyítható az azonosság. $\boldsymbol{B}$ helyett vegyük a $\boldsymbol{B}_n = \boldsymbol{B} - \varepsilon_n \boldsymbol{I}$ mátrixot, ahol $\varepsilon_n \to 0$, és az ε_n sorozat elkerüli $\boldsymbol{B}$ sajátértékeit. Akkor $|\boldsymbol{B}_n| \neq 0$, tehát az előzőekben bizonyított állítás érvényes: $\cdot|\boldsymbol{A}\boldsymbol{B}_n - \lambda\boldsymbol{I}| = |\boldsymbol{B}_n\boldsymbol{A} - \lambda\boldsymbol{I}|$. Rögzített λ mellett mindkét kifejezés ε_n folytonos függvénye (polinomja), tehát $\varepsilon_n \to 0$ esetén a bal oldal $|\boldsymbol{A}\boldsymbol{B} - \lambda\boldsymbol{I}|$-hez, a jobb oldal $|\boldsymbol{B}\boldsymbol{A} - \lambda\boldsymbol{I}|$-hez tart, tehát $\boldsymbol{A}\boldsymbol{B}$ és $\boldsymbol{B}\boldsymbol{A}$ karakterisztikus polinomja azonos. Ez az összes sajátérték és a multiplicitások egyezését jelenti.

Mivel $||\boldsymbol{A}||$ az $\boldsymbol{A}^*\boldsymbol{A}$, az $||\boldsymbol{A}^*||$ az $\boldsymbol{A}\boldsymbol{A}^*$ legnagyobb sajátértékének a négyzetgyöke, és tudjuk, hogy $\boldsymbol{A}^*\boldsymbol{A}$ és $\boldsymbol{A}\boldsymbol{A}^*$ sajátértékei megegyeznek, a normák is egyenlők.

14. Tudunk-e egyszerű példát adni arra, hogy $\boldsymbol{A}$ minden sajátértéke abszolút értékben egynél kisebb, és $||\boldsymbol{A}|| > 1$?

A 12. Kérdés során láttuk, hogy az $\boldsymbol{A}$ nem ortogonális vetítés normája mindig 1-nél nagyobb. Válasszunk egy c számot úgy, hogy $1 < c < ||\boldsymbol{A}||$. Az $\boldsymbol{A}$ sajátértékei 0, vagy 1 lehetnek. Képezzük a $\frac{1}{c}\boldsymbol{A}$ mátrixot, a sajátértékei 0 és $\frac{1}{c}$, tehát 1-nél kisebbek, ugyanakkor $||\frac{1}{c}\boldsymbol{A}|| = \frac{1}{c}||\boldsymbol{A}|| > 1$.

Még egyszerűbb talán a következő. Legyen $\boldsymbol{C}$ olyan mátrix, melyben minden elem nulla, kivéve a_{1n}-et, legyen $a_{1n} = K$. Az $\frac{1}{2}\boldsymbol{I} + \boldsymbol{C}$ olyan mátrix, melyben minden sajárérték $\frac{1}{2}$, ugyanakkor az $\boldsymbol{x} = (0, 0, \ldots, 0, 1)$ vektorra alkalmazva $||(\frac{1}{2}\boldsymbol{I} + \boldsymbol{C})\boldsymbol{x}|| > K$, tehát $||(\frac{1}{2}\boldsymbol{I} + \boldsymbol{C})|| > K$.

15. Igaz-e, hogy minden négyzetes $\boldsymbol{A}$ mátrixra $\sin^2\boldsymbol{A} + \cos^2\boldsymbol{A} = \boldsymbol{I}$?

Mivel a $\sin x$ és a $\cos x$ hatványsora a teljes komplex síkon konvergens, $\sin\boldsymbol{A}$ és $\cos\boldsymbol{A}$ mindig értelmezhető.

Állítsuk elő $\sin^2 x$ és $\cos^2 x$ hatványsorát! Mivel $\sin^2 x = \frac{1}{2} - \frac{1}{2}\cos 2x$, $\sin^2 x$ hatványsora

$$\sin^2 x = \frac{1}{2} - \frac{1}{2}\sum_{k=0}^{\infty} \frac{2^{2k}}{(2k)!} x^{2k},$$

és hasonlóképpen

$$\cos^2 x = \frac{1}{2} + \frac{1}{2}\sum_{k=0}^{\infty} \frac{2^{2k}}{(2k)!} x^{2k}.$$

Ebből

$$\sin^2 \boldsymbol{A} = \frac{1}{2}\boldsymbol{I} - \frac{1}{2}\sum_{k=0}^{\infty} \frac{2^{2k}}{(2k)!} \boldsymbol{A}^{2k} = \frac{1}{2}\boldsymbol{I} - \boldsymbol{C}$$

és
$$\cos^2 \boldsymbol{A} = \frac{1}{2}\boldsymbol{I} + \frac{1}{2}\sum_{k=0}^{\infty} \frac{2^{2k}}{(2k)!}\boldsymbol{A}^{2k} = \frac{1}{2}\boldsymbol{I} + \boldsymbol{C},$$
vagyis $\sin^2\boldsymbol{A} + \cos^2\boldsymbol{A} = \boldsymbol{I}$.

16. Mutassuk meg, hogy $\boldsymbol{A}^*\boldsymbol{A}$ pozitív definit vagy pozitív szemidefinit mátrix aszerint, hogy $\boldsymbol{A}$ teljesrangú mátrix vagy nem. Igaz-e, hogy minden pozitív definit vagy pozitív szemidefinit szimmetrikus mátrix $\boldsymbol{A}^*\boldsymbol{A}$ alakban írható fel?

$\boldsymbol{A}^*\boldsymbol{A}$ nyilván szimmetrikus mátrix, mert transzponáltja önmaga. Mivel $\langle \boldsymbol{x},\, \boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x}\rangle = \langle \boldsymbol{A}\boldsymbol{x},\, \boldsymbol{A}\boldsymbol{x}\rangle = ||\boldsymbol{A}\boldsymbol{x}||^2 \geq 0$, a pozitív szemidefinit vagy definit tulajdonság igazolva van. $|\boldsymbol{A}^*\boldsymbol{A}| = |\boldsymbol{A}^*|\cdot|\boldsymbol{A}| = |\boldsymbol{A}|^2$, tehát ha $|\boldsymbol{A}| = 0$, akkor $|\boldsymbol{A}^*\boldsymbol{A}| = 0$ szintén, és a kvadratikus alak csak pozitív szemidefinit lehet. Ha $|\boldsymbol{A}| \neq 0$, akkor $|\boldsymbol{A}^*\boldsymbol{A}| \neq 0$, tehát pozitív definit alakról van szó.

Ha $\boldsymbol{B}$ pozitív definit vagy pozitív szemidefinit, akkor a $\boldsymbol{C}$ ortogonális transzformációval olyan tiszta négyzetes alakra transzformálható, ahol minden együttható (sajátérték) pozitív, vagy nulla. Jelöljük $\boldsymbol{D}$-vel azt a diagonális mátrixot, melynek főátlóját a sajátértékek négyzetgyökei képezik, és legyen $\boldsymbol{A} = \boldsymbol{D}\boldsymbol{C}^*$. Ilyen választás mellett $\boldsymbol{A}^* = \boldsymbol{C}\boldsymbol{D}$ és
$$\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{C}\boldsymbol{D}\boldsymbol{D}\boldsymbol{C}^* = \boldsymbol{C}\boldsymbol{D}^2\boldsymbol{C}^* = \boldsymbol{B}.$$

17. Melyek a $\boldsymbol{D}$ mátrix $\boldsymbol{D} = \boldsymbol{F}^*\boldsymbol{F}$ faktorizációjához tartozó tényezők, ha
$$\boldsymbol{D} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 8 & 10 \\ 2 & 10 & 16 \end{pmatrix}?$$

A $\boldsymbol{D}$-hez tartozó kvadratikus alak
$$q(\boldsymbol{x}) = x^2 + 4xy + 7xz + 8y^2 + 20xz + 16z^2 =$$
$$= (x + 2y + 3z)^2 + 4y2 + 8xz + 7z^2 =$$
$$= (x + 2y + 3z)^2 + (2x + 2z)^2 + (\sqrt{3}z)^2.$$
Ebből leolvasható, hogy
$$\boldsymbol{F} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & \sqrt{3} \end{pmatrix},$$

és ellenőrizhető, hogy $\boldsymbol{D} = \boldsymbol{F}^*\boldsymbol{F}$.

18. Igaz-e, hogy egy vetítés akkor és csak akkor ortogonális, ha mátrixa pozitív szemidefinit?

Jelöljük az ortogonális vetítést $\boldsymbol{P}$-vel, akkor $\langle \boldsymbol{x},\ \boldsymbol{P}\boldsymbol{x}\rangle = \langle \boldsymbol{x},\ \boldsymbol{P}^2\boldsymbol{x}\rangle = $ $= \langle \boldsymbol{P}^*\boldsymbol{x},\ \boldsymbol{P}\boldsymbol{x}\rangle = ||\boldsymbol{P}\boldsymbol{x}||^2 \geq 0$.

Másrészt, ha $\boldsymbol{P}$ a vetítés, és $\boldsymbol{P}_0$ jelöli ugyanarra a képtérre történő ortogonális vetítést, akkor válasszunk olyan $\boldsymbol{x}$-et, melyre $\boldsymbol{P}\boldsymbol{x} \neq \boldsymbol{P}_0\boldsymbol{x}$ és legyen $\boldsymbol{y} = \boldsymbol{x} - \frac{1}{2}(\boldsymbol{P} + \boldsymbol{P}_0)\boldsymbol{x}$. Ekkor

$$0 \leq \langle \boldsymbol{y}, \boldsymbol{P}\boldsymbol{y}\rangle = \langle \boldsymbol{y}, \frac{1}{2}(\boldsymbol{P} - \boldsymbol{P}_0)\boldsymbol{x}\rangle = -\langle \boldsymbol{y}, \frac{1}{2}(\boldsymbol{P}_0 - \boldsymbol{P})\boldsymbol{x}\rangle = -\langle \boldsymbol{y}, \boldsymbol{P}_0\boldsymbol{y}\rangle \leq 0,$$

tehát $\langle \boldsymbol{y}, \boldsymbol{P}_0\boldsymbol{y}\rangle = 0$. $\boldsymbol{P}_0$ ortogonalitása miatt $\langle \boldsymbol{x} - \boldsymbol{P}_0\boldsymbol{x}, \boldsymbol{P}_0\boldsymbol{y}\rangle = 0$, a kettő különbségeként, felhasználva, hogy $\boldsymbol{y} - (\boldsymbol{x} - \boldsymbol{P}_0\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{P}_0 - \boldsymbol{P})\boldsymbol{x} = \boldsymbol{P}_0\boldsymbol{y}$, kapjuk, hogy $||\boldsymbol{P}_0\boldsymbol{y}||^2 = \boldsymbol{0}$, $\boldsymbol{P}_0\boldsymbol{y} = \boldsymbol{0}$, de ekkor $\boldsymbol{P}\boldsymbol{x} = \boldsymbol{P}_0\boldsymbol{x}$.

Ha $\boldsymbol{P} = \boldsymbol{I}$, akkor nem igaz az állítás, mert $\boldsymbol{P}$ pozitív definit.

Irodalom

Dancs István, Puskás Csaba: *Vektorterek*, Aula, 2001.

Rózsa Pál: *Bevezetés a mátrixelméletbe*, Typotex, 2009.

Rózsa Pál, Hegedűs Csaba: *Mátrixelmélet*, MTA Matematikai Kut. Int. kiadványa, 1974.

Hajós György: *Bevezetés a geometriába*, Tankönyvkiadó, 1962.

Halmos, Paul R.: *Véges dimenziós vektorterek*, Typotex

Bártfai Pál: *Gazdasági matematika*, Harsányi J. Főiskola tankönyve, 2008.

Bártfai Pál: *Gazdasági matematika – Rutin feladatok*, Harsányi J. Főiskola tankönyve, 2006.

Gantmacher, F. R.: *The theory of matrices*, Chelsea Scientific Books.

Tárgymutató