

Számok és mátrixok közepei

Temesi Róbert IV. évf.

konzulens: Dr. Petz Dénes, Analízis tanszék

2004. október 25.

1. Bevezetés

A legegyszerűbb számközepek az ókor óta ismertek. Ilyenek például a számtani és geometriai közép. Másokat csak később fedeztek fel, mint például a logaritmikus közepet, és sokukról még ma is csak keveset tudunk, annak ellenére, hogy a gyakorlatban azért találkozhatunk velük. Régóta próbálják a hagyományos kétváltozós közepeket kiterjeszteni többváltozóra, ami több-kevesebb sikerrel megy is. A több változóra való általánosítás az egyik fő kérdés a közepek elméletében. Közepek egy másik fajta általánosítása a mátrixokra való kiterjesztés. Ezzel csak a XX. században kezdtek el foglalkozni. Pozitív mátrixok számtani és harmonikus közepét először csupán a 60-as évek végén vezette be Anderson és Duffin fizikai indíttatás nyomán *Mátrixok soros és párhuzamos összeadása* című 1969-es munkájában. Csupán hat évre rá, 1975-ben ismerték fel, hogy Hilbert terek operátoraira is definiálhatók ezek a közepek, és hogy milyen elméleti fontossága van az operátorelméletben. Még ebben az évben Pusz és Woronowicz bevezette két pozitív operátor geometriai közepét. Nem sokkal később, mintegy öt év múlva operátorok kétváltozós közepeire igen jó karakterizációt talált Kubo és Ando. Minden középnek megfelel egy normalizált operátormonoton valós függvény. Ez igen nagy eredmény. Mivel úgy gondoljuk, hogy a valós függvényeket jól ismerjük, úgy jól ismerjük a kétváltozós közepeket is.

Na de mi a helyzet több változóban? Vagy akárcsak mi a helyzet három változóban? Több mint húsz év telt el a kétváltozós közepek nagy eredménye óta, de még azóta sem tudtak semmi általánosítást mondani három változóban. Kisebb, speciális eredmények is csak mostanában látnak napvilágot, mint például a geometriai közép egy n változós általánosítása 2004-ben. A közepek nagyrészeről viszont semmit sem tudunk.

Ebben a dolgozatban egy tetszőleges kétváltozós közepet kiterjesztünk három változóra bizonyos speciális esetben, amikor az argumentumok rendezhetőek. Ezt a kiterjesztést használja a geometriai középre 2004-es cikkében Ando, Li és Mathias is [1]. Megmutatják, hogy az irodalomban egyedül ez teljesíti a feltételeket, amit egy geometriai középtől elvárhatunk. Ráadásul a kiterjesztés nem igényli a rendezhetőséget. Kutatásunk tárgya ezért az, hogy vajon más közepekre mit ad a kiterjesztés. Ebben a dolgozatban azt vizsgáljuk, hogy speciálisan a logaritmikus közép esetében egyáltalán értelmes-e a kiterjesztés, ha a mátrixok nem rendezettek. Számítógépi kísérlet útján

sejtést fogalmazunk meg, miszerint tetszőleges három pozitív mátrixnak értelmezhető a kiterjesztés segítségével a logaritmikus közepe.

2. Számok közepei

A számok közepeit axiomatikusan tekintve a következőket mondhatjuk. Először is számokon, a valós számokat értjük. A valós számokon van rendezés, és ez fontos, amikor közepekről akarunk beszélni. Sőt leginkább pozitív valós számok közepeivel foglalkozunk, hiszen már a geometriai közép is csak a pozitív valósokon értelmes.

Két szám közepét egy kétváltozós függvénnyel tudjuk megadni:

$$M(x, y) : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$$

Itt M a közép angol elnevezésének a *mean* szónak a kezdőbetűje. Egy középtől a következő tulajdonságokat várjuk el. Azonos számok közepe legyen azonos, egy közép legyen szimmetrikus, különböző számok közepe essen a két szám közé, ezért közép, legyen monoton, és egy közép legyen folytonos, mint kétváltozós függvény. Képlettel:

- (1) $\forall x \in \mathbb{R}_+ \quad M(x, x) = x$
- (2) $\forall x, y \in \mathbb{R}_+ \quad M(x, y) = M(y, x)$
- (3) $\forall x, y \in \mathbb{R}_+ \text{ ha } x < y, \text{ akkor } x < M(x, y) < y$
- (4) ha $x \leq x'$ és $y \leq y'$, akkor $M(x, y) \leq M(x', y')$
- (5) $M(x, y)$ folytonos

Ezen kívül a legtöbb közép még rendelkezik egy extra tulajdonsággal is, mégpedig a homogenitással, azaz:

- (6) $\forall t \in \mathbb{R}_+ \quad M(tx, ty) = t M(x, y)$

Homogén az aritmetikai:

$$A(x, y) = \frac{x + y}{2},$$

geometriai:

$$G(x, y) = \sqrt{xy},$$

harmonikus:

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

és a logaritmikus közép is:

$$L(x, y) = \begin{cases} \frac{x-y}{\log x - \log y} & \text{ha } x \neq y \\ x & \text{ha } x = y \end{cases}.$$

Persze mindegyikről könnyen belátható az is, hogy teljesítik az (1)–(5) tulajdonságokat is.

A háromváltozós közepekre is hasonló tulajdonságokat követelhetünk meg, mint a kétváltozósakra. Egy háromváltozós közép egy $M(x, y, z) : \mathbb{R}_+^3 \rightarrow \mathbb{R}_+$ függvény, ami kielégíti az alábbi tulajdonságokat:

$$(1') \quad \forall x \in \mathbb{R}_+ \quad M(x, x, x) = x$$

$$(2') \quad \forall x, y, z \in \mathbb{R}_+ \quad \text{és } x, y, z \text{ minden } P \text{ permutációjára}$$

$$M(P(x), P(y), P(z)) = M(x, y, z)$$

$$(3') \quad \text{ha } x < y < z, \text{ akkor } x < M(x, y, z) < z$$

$$(4') \quad \text{ha } x \leq x', y \leq y', z \leq z', \text{ akkor } M(x, y, z) \leq M(x', y', z')$$

$$(5') \quad M(x, y, z) \text{ folytonos}$$

A legtöbb háromváltozós középére is igaz a homogenitás:

$$(6') \quad \forall t \in \mathbb{R}_+ \quad M(tx, ty, tz) = t M(x, y, z)$$

A háromváltozós aritmetikai:

$$A(x, y, z) = \frac{x + y + z}{3},$$

geometriai:

$$G(x, y, z) = \sqrt[3]{xyz},$$

illetve harmonikus közép:

$$H(x, y, z) = \frac{3}{\frac{1}{x} + \frac{1}{y} + \frac{1}{z}}$$

kielégíti ezeket a tulajdonságokat. A logaritmusos közép háromváltozós kiterjesztésére több lehetőség is adódik:

$$L_1(x, y, z) := \frac{x}{(\log x - \log y)(\log x - \log z)} + \frac{y}{(\log y - \log x)(\log y - \log z)} + \frac{z}{(\log z - \log x)(\log z - \log y)},$$

vagy

$$L_2(x, y, z) = \frac{1}{2} \frac{(x-y)(x-z)(y-z)}{xy(\log x - \log y) + xz(\log z - \log x) + yz(\log y - \log z)}.$$

Ha x, y, z közül valamelyek azonosak, akkor megfelelő határértéket kell venni a képlet értelmezéséhez. Mindkét képlet a kétváltozós logaritmusos közép képletének egy kiterjesztése egyfajta logika szerint, és mindkettő használatos [5], [6]. Ezért nem is tudjuk, hogy melyiket tekintjük az "igazi" kiterjesztésnek. Adódik egy általános módszer kétváltozós közepek háromváltozós kiterjesztésére. A kétváltozós $M(x, y)$ középből megkaphatunk egy háromváltozós $M(x, y, z)$ közepet. Rendezzük növekvő sorozatba az x, y, z , számokat, és jelölje x_0 a legkisebbet, y_0 a középsőt, és z_0 a legnagyobbat, azaz

$$x_0 \leq y_0 \leq z_0.$$

$i = 1, 2, 3, \dots$ esetén legyen

$$x_i = M(x_{i-1}, y_{i-1}), \quad y_i = M(x_{i-1}, z_{i-1}), \quad z_i = M(y_{i-1}, z_{i-1}).$$

(1), (3) és (4) miatt $n = 0, 1, 2, 3, \dots$ -ra fennáll

$$x_{n-1} \leq x_n \leq y_n \leq z_n \leq z_{n-1}.$$

Ebből következik, hogy x_n sorozat monoton növekvő és korlátos, valamint z_n sorozat monoton csökkenő és szintén korlátos. Azaz létezik a határértékük, amit jelöljünk X -szel és Z -vel. Az

$$y_n = M(x_{n-1}, z_{n-1})$$

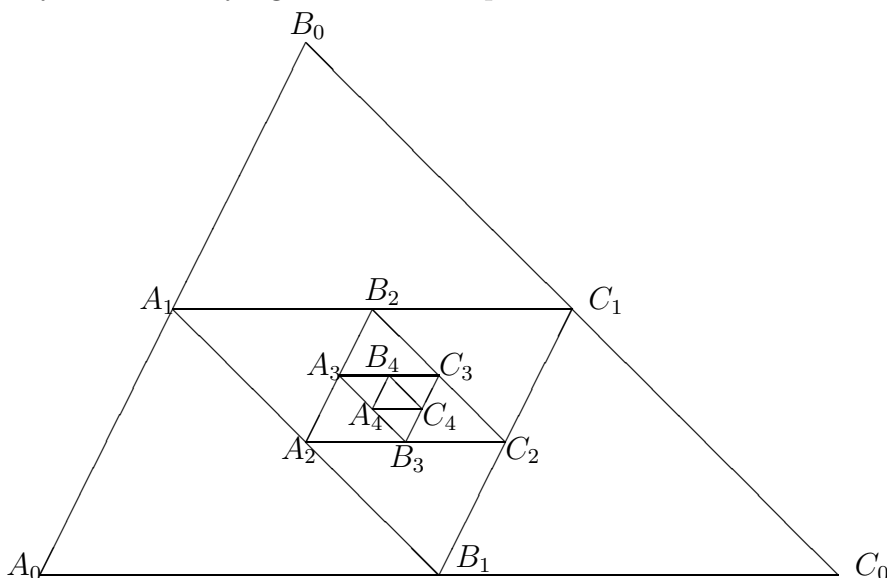
egyenlőségből és (5)-ből következik, hogy y_n is konvergens és határértéke $Y = M(X, Z)$. $X \leq Z$, de $X = Z$ is fennáll, mert ha $X < Z$ lenne, akkor (3) miatt $X < Y < Z$ is teljesülne, de

$$z_n = M(y_{n-1}, z_{n-1})$$

egyenlőségből és (5)-ből $Z = M(Y, Z)$ -t kapnánk, de ez $Y < Z$ -vel együtt ellentmond (3)-nak. Tehát $X = Z$ és (1) miatt $X = Y = Z$ is igaz. Legyen $M(x, y, z)$ ez a közös határérték.

Így $M(x, y, z)$ tényleg közép. (1'), (2') triviálisan igaz, (3') és (4') (3) és (4) segítségével és indukcióval könnyen belátható, (5') pedig abból következik, hogy x_n, y_n és z_n tart $M(x, y, z)$ -hez és x, y, z folytonos függvénye, ami indukcióval következik (5)-ből. $M(x, y, z)$ homogén is, ami $M(x, y)$ homogenitásából következik.

Ez az algoritmus a kétváltozós számtani középéből előállítja a háromváltozósat. Ennek bizonyítására könnyű geometriai interpretációt is találni.



Legyen most $x_0 < y_0 < z_0$, hogy ne legyenek elfajult esetek. x_0, z_0 -t ábrázoljuk az x, y koordináta-rendszer x tengelyén: $A_0 = (x_0, 0), C_0 = (z_0, 0)$. y_0 -t pedig ábrázoljuk az x tengelyen kívül: $B_0 = (y_0, 1)$. Ezek egy háromszöget határoznak meg. Ekkor A_0, B_0, C_0 x tengelyre vett vetülete rendre x_0, y_0, z_0 . Ekkor x_1, y_1, z_1 rendre az $[x_0, y_0], [x_0, z_0]$ és

$[y_0, z_0]$ szakaszok felezőpontjai. Mivel valójában csak a vetületekre vagyunk kíváncsiak, ezért ezeknek megfeleltethetjük rendre pont A_0, B_0, A_0, C_0 és B_0, C_0 oldal szakaszok felezőpontjait. Így kapunk egy újabb háromszöget a kiindulási háromszög súlyvonaláiból. Az így kapott háromszögnek is a súlyvonalait véve folytathatjuk az eljárást. Az n . háromszög csúcsainak a vetülete az x tengelyre pont x_n, y_n, z_n . Viszont geometriából tudjuk, hogy egy háromszögnek és a súlyvonaláiból alkotott háromszögnek azonos a súlypontja. A súlypont vetülete viszont mindig az eredeti három szám háromváltozós aritmetikai közepe. A háromszögek oldalhosszai feleződnek, így a csúcsok tartanak a súlyponthoz, a vetületek pedig az aritmetikai középhez.

Tetszőleges x, y, z esetén a számtani középpel végzett algoritmusból $\log x, \log y, \log z$ behelyettesítésével $\log(\sqrt[3]{xyz})$ adódik. e^x folytonossága miatt a sorozatok tagjaira e^x -et alkalmazva, azt látjuk, hogy az x, y, z -vel kezdődő sorozatok $\sqrt[3]{xyz}$ -hoz tartanak. De ezeknek a sorozatoknak a tagjai pont azok, mint amelyeket x, y, z -ből indulva geometriai közép vételével kapnánk, hiszen $e^{\frac{x+y}{2}} = \sqrt{e^x e^y}$. Ami azt mutatja, hogy az algoritmus a kétváltozós geometriai középből kiindulva a háromváltozósat adja.

Hasonlóan

$$\frac{2}{\frac{1}{x} + \frac{1}{y}} = \frac{1}{\frac{\frac{1}{x} + \frac{1}{y}}{2}}$$

és

$$\frac{3}{\frac{1}{x} + \frac{1}{y} + \frac{1}{z}} = \frac{1}{\frac{\frac{1}{x} + \frac{1}{y} + \frac{1}{z}}{3}}$$

miatt a kétváltozós harmonikus középből is megkapható a háromváltozós harmonikus közepet.

A logaritmusos közép viszont nem ez a helyzet. Az algoritmus a kétváltozós logaritmusos közép felhasználásával a két háromváltozós logaritmusos középtől különböző háromváltozós logaritmusos közepet ad. Ehhez azt a tényt kell felhasználni, hogy az algoritmus ugyan azt a határértéket adja függetlenül attól, hogy x_0, y_0, z_0 -ból, vagy $n \geq 1$ -re x_n, y_n, z_n -ből indulunk ki. Így például az is igaz, hogy

$$M(x_0, y_0, z_0) = M(M(x_0, y_0), M(x_0, z_0), M(y_0, z_0))$$

Például átírhatjuk L_1 -et és L_2 -t kettes alapú logaritmusra, ekkor $x_0 = 1, y_0 = 2, z_0 = 4$ esetén a fenti egyenlőség nem teljesül.

3. Mátrixok közepei

Közepekre az élet számtalan területén van szükség. Az n változós számtani közép, átlag néven sok helyen előbukkan. Előfordul például a statisztikában, ahol független azonos eloszlású valószínűségi változók átlaga majdnem biztosan tart a közös várható értékükhöz. Igaz ez akkor is, ha a valószínűségi változók több dimenziós valószínűségi vektor változók. Elsők között azonban mindenki általános iskolában találkozik az átlaggal, amikor ki kell számolnia a tanulmányi átlagát. Az egyetemen használatos súlyozott tanulmányi

átlag viszont szigorú értelemben véve már nem közép, mert nem feltétlenül szimmetrikus, illetve több változóban nem feltétlenül permutáció invariáns. Viszont bele tartozik egy tágabb fogalomkörbe, azt mondhatjuk rá, hogy kapcsolás az angol *connection* szó fordításával. Itt a kapcsolás szó jogos, hiszen például a harmonikus közép fele, ami szintén kapcsolás, megadja a fizikában két párhuzamosan kapcsolt ellenállás eredő ellenállását. Láttuk, hogy például a statisztikában vektorok, de általánosabban akár mátrixok közepe is előkerül, nem csak számoké. A fizikában sincs ez másképp. Tekintsünk egy olyan eszközt, "ellenállást", amelynek n ki- és n bemenete van. Egy ilyen "ellenállás" sarkain mérhető feszültségeknek a bemeneti áramerősségekkel arányos változását egy $n \times n$ -es mátrixszal lehet leírni. Két ilyen áramköri elem például párhuzamos kapcsolását pedig a két megfelelő mátrix harmonikus közepének felével lehet megadni [4]. A mátrixok harmonikus közepe ezen a ponton az egyesített áramköri elemek mátrixát jelenti. Azonban belátható, hogy ez az összefüggés a számok közepeire hasonló mátrixokra vonatkozó axiómákat elégíti ki. Vagyis hasonló értelemben közép, mint a számokra alkalmazható változata. Mátrixok kétváltozós közepeire Kubo és Ando alkotott elméletet [2]. Pontosabban szólva végtelen dimenziós Hilbert tereken ható operátorok kétváltozós közepeit tárgyalták axiomatikusan. Itt azonban operátor helyett mindig mátrixot írunk. Minden itt szereplő mátrix $n \times n$ -es önadjungált, de főleg valós elemű pozitív mátrix lesz. A és B $n \times n$ es önadjungált mátrixok esetén akkor mondjuk, hogy $A \geq B$, ha $A - B \geq 0$, azaz a $A - B$ sajátértékei nemnegatívak. Kubo és Ando először bevezették a kapcsolás fogalmát. $M(A, B)$ kétváltozós mátrixfüggvény akkor kapcsolás, ha monoton, folytonos és kielégíti az ún. transzformátor egyenlőtlenséget:

$$CM(A, B)C \leq M(CAC, CBC),$$

ami a homogenitás egy erősebb változata. Egy kapcsolást akkor neveznek középnek, ha normalizált: $M(I, I) = I$, ahol I az identitás. Nem követelik meg a szimmetrikusságot, de mi ezt is elvárjuk.

Következik, hogy egy középre fennállnak az alábbiak, ha a monotonitásnál szigorú egyenlőtlenséget követelünk meg.

$$(I) \quad M(A, A) = A$$

$$(II) \quad M(A, B) = M(B, A)$$

$$(III) \quad \text{ha } A < B, \text{ akkor } A < M(A, B) < B$$

$$(IV) \quad \text{ha } A < A' \text{ és } B < B', \text{ akkor } M(A, B) < M(A', B')$$

$$(V) \quad M(A, B) \text{ folytonos}$$

$$(VI) \quad CM(A, B)C \leq M(CAC, CBC)$$

Ezeket az axiómákat kielégíti a természetes módon értelmezett aritmetikai és harmonikus közép:

$$A(A, B) = \frac{A + B}{2}, \text{ illetve } H(A, B) = 2(A^{-1} + B^{-1})^{-1}.$$

Elvárható, hogy az így kapott közepek kiterjesztései legyenek a számok megfelelő közepeinek. Ezzel nincs gond, mivel 1×1 -es mátrixok esetén visszakapjuk az eredeti közepet. A harmonikus középénél már nem megfelelő a szokásos tört írásmód, inverzet kell alkalmazni. Még bonyolultabb a helyzet a geometriai középével. A \sqrt{xy} képlet értelmes ugyan mátrixokra is, de nem szimmetrikus, és például

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} < B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

esetén könnyen látható, hogy (III) sem teljesül. Egy másik mód a kiterjesztésre azon alapul, hogy minden kapcsoláshoz bijektíven hozzárendelhető egy $m : (0, \infty) \rightarrow (0, \infty)$ mátrixmonoton függvény, amelyre $m(x) = M(1, x)$, ahol $1, x \in \mathbb{R}_+$. A számok geometriai közepének $g(x) = \sqrt{x}$ felel meg. Ez mátrixmonoton, és a hozzá tartozó közép

$$G(A, B) = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}.$$

Ez is mátrixközép és kiterjesztése a geometriai közép skalár változatának, csakúgy, mint az aritmetikai és harmonikus közép. Látható, hogy nem magától értetődő, hogy egy megszokott számközepet ki lehet terjeszteni mátrixokra is. A középhez tartozó valós függvényt meg kell vizsgálni, hogy mátrixmonoton-e, ami sokszor nehéz feladat. A kétváltozós logaritmus középénél az

$$l(x) = \frac{x-1}{\log x}$$

függvény mátrixmonotonitását kell eldönteni. Szerencsére ennek a függvénynek létezik egy integrál alakja és pedig

$$l(x) = \int_0^1 x^\alpha d\alpha.$$

Tudjuk, hogy x^α mátrixmonoton pontosan akkor, ha $\alpha \in [0, 1]$, de ilyenek összege is az és ilyenek limesze is az, tehát az integrál is mátrixmonoton. Vagyis a kétváltozós logaritmus középénél valóban mátrixközép. Szimmetrikus pozitív definit A mátrix logaritmusát, illetve más f függvényét is úgy vesszük, hogy a mátrixot

$$A = U^*DU$$

alakba írjuk, ahol U unitér, $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, ahol $\lambda_1, \dots, \lambda_n$ az A mátrix sajátértékei, majd a sajátértékeket helyettesítjük a logaritmus illetve a tetszőleges f függvénybe, és visszaszorzunk az unitérekkel, azaz

$$f(A) = U^*f(D)U = U^* \text{diag}(f(\lambda_1), \dots, f(\lambda_n))U.$$

A mátrixmonoton függvényből az alábbi képlettel kaphatjuk vissza a közepet:

$$M(A, B) = A^{1/2}f(A^{-1/2}BA^{-1/2})A^{1/2}.$$

4. Mátrixközepek három változóban

Léteznek többváltozós mátrix közepek is. Az aritmetikai és harmonikus közép egyszerűen kiterjeszthető tetszőlegesen sok változóra, és ezeknek a kiterjesztéseknek a jogosságát senki sem kérdőjelezi meg, mert olyan sok helyen használtak, ráadásul végtelenül egyszerűek. A többi középpel azonban már három változóban is rengeteg probléma akad. A geometriai középtől három változóban több ésszerű tulajdonságot is elvárhatunk. Ando, Li és Mathias egy 2004-es cikkében 10 ilyen tulajdonságot nevez meg. Talán azt várnánk, hogy több változóra nem is nagyon lehet találni olyan közepeket, amik többé kevésbé kielégítik ezeket a tulajdonságokat, ezzel szemben meglepő módon a valóság az, hogy több ilyen közép is van. Mint a háromváltozós számokra vonatkozó logaritmikus középénél a mátrixok háromváltozós geometriai közepénél is az a helyzet, hogy több jelölt is akad. Ezek közül az egyik az első részben ismertetett végtelen algoritmus mátrixos változatával áll elő páronkénti geometriai mátrixközép vételével. Számok között mindig van rendezés, mátrixok között viszont csak kivételes esetben. Az algoritmus bizonyosan működik mátrixokra is abban az esetben, amikor a három mátrix között van rendezés. Legyen $M(A, B)$ egy tetszőleges mátrixközép. Ha $A \leq B \leq C$, akkor legyen $A_0 = A, B_0 = B, C_0 = C$. Valamint rekurzíve $n = 1, 2, 3, \dots$ -ra legyen

$$A_{n+1} = M(A_n, B_n), \quad B_{n+1} = M(A_n, C_n), \quad C_{n+1} = M(B_n, C_n).$$

A három mátrixsorozat ebben az esetben is konvergens, ami ugyan úgy látható be, mint a skalár esetben, az analóg mátrixokról szóló tételek felhasználásával. Az így kapott közös határérték $M(A, B, C)$ háromváltozós mátrixközép, ami hasonlóan bizonyítható, mint számokra. Ezzel az algoritmussal a kétváltozós aritmetikai mátrixközépből megkapjuk a háromváltozósat, és emiatt a kétváltozós harmonikus mátrixközépből megkapjuk a háromváltozósat, mint a skalár esetben. Sőt a konvergencia akkor is működik, amikor a három mátrix nincs rendezve, és ezt nem nehezebb belátni, mint a számok esetében. A geometriai középpel már kicsit más a helyzet. Ennél a nem rendezett esetben a konvergenciát csak 2004-ben látták be [1]. Azóta született rá egyszerűbb megoldás is [3], felhasználva, hogy A -hoz létezik $\lambda > 0$ és $\mu > 0$, hogy

$$A \leq \lambda B \leq \mu C$$

és a geometriai közép azon speciális tulajdonságát, hogy

$$G(\lambda A, \mu B) = \sqrt{\lambda \mu} G(A, B).$$

Más közepekre még nem tudjuk, hogy nem rendezett esetben van-e konvergencia. Az itt tárgyalt negyedik közép esetében, a logaritmikus középénél szintén nem tudjuk, hogy a nem rendezett esetben konvergál-e az iteráció. Végeztünk számítógépes kísérletet erre vonatkozólag. Ez persze nem bizonyítás, mivel numerikus eljárásról van szó, de legalább sejtés megfogalmazására alkalmas. A program octave-ban készült, ami a Matlab szabad változata. Az octave jól alkalmazható numerikus mátrixműveletek végzésére, ezért ebben íródott a program. A program a kísérletek során Linux operációs rendszeren futott, és a futtató shell bash volt. Az octave egyik előnye, hogy a programot tartalmazó

fájl parancssorból is futtatható. Az algoritmust természetesen véges sok lépésig lehetett csak elvégezni. Fontos információt nyújtanak a kiinduló mátrixok, melyek véletlen szimmetrikus pozitív definit mátrixok voltak, és szintén fontos információt nyújtanak az iteráció megszakításakor tapasztalt eltérések a kapott mátrixok között. Ezeket az adatokat célszerű volt fájlba menteni, hogy később is lehessen azokat elemezni. Ugyanakkor a kísérlet futása során kényelmes volt látni, hogy éppen hol tart a program futása, ami viszont nem egy fájlban, hanem a képernyőn kellett, hogy megjelenjen. Tehát a program outputját két helyre is kellett irányítani, ami nem látszott megoldhatóan csak az octave segítségével. Ezért készült egy shell szkript, ami felügyelte az octave program futását. A mellékletekben közöljük az octave programot kommentekkel kiegészítve, valamint a shell szkriptet minimális kommentezéssel. A futtatás során a program sztenderd normális eloszlás szerint készített három diagonális mátrixot, majd ezeket megszorozta három sztenderd normális eloszlású elemeket tartalmazó szimmetrikus mátrixszal, majd ha rendezhetőek voltak újra próbálkozott, amíg nem rendezhetőek lettek. Ezekkel 100 lépést végzett az iterációból páronkénti logaritmikus mátrixközép vételével. A kapott mátrixok páronkénti 2-es normabeli eltéréseinek maximumát vette. Az egész eljárást százszor megismételte a shell szkript, és a száz alkalommal kapott maximális 2-es normabeli eltérések maximumát vette, és megjelölte azt a helyet, ahol az a három mátrix volt található, amlyekkel indulva ez először előfordult.

Így azt kaptuk, hogy a maximális eltérés a fenti paraméterekkel 3×3 -as mátrixokkal 10^{-4} nagyságrendű volt és először a 92. kísérletnél fordult elő. Átlagosan azonban a maximális eltérés 10^{-7} nagyságrendű volt. Ebből arra a sejtésre következtetünk, hogy az algoritmus konvergenciát ad a logaritmikus mátrix középére is a nem rendezett esetben.

Ellenőrzés képpen a kísérletet elvégeztük a geometriai középére is hasonló paraméterekkel rendezett és nem rendezett 3×3 -as mátrixokkal, valamint elvégeztük a kísérletet a logaritmikus középpel 3×3 -as mátrixokkal a rendezett esetben is, és papírforma szerint konvergenciát tapasztaltunk numerikus szinten.

5. Konklúzió, tervek

Egy érdekes és izgalmas kérdés kutatása közben vagyunk éppen. Egyfelől azt tapasztaljuk, hogy két kitüntetett közép van, amire sok tulajdonság automatikusan teljesül, ezek az aritmetikai és harmonikus közép. Van egy harmadik közép is, ami rokon velük, ez a geometriai közép, amely ugyan kis késéssel, de igazodik hozzájuk, vagyis ami könnyen belátható az első kettőre, az idővel belátható a harmadikra is. Másfelől van a többi közép, mint például a logaritmikus közép, amelyek nehezebben megismerhetőek. Ugyanakkor a számítógépes szimuláció egyelőre affelé hajlik, hogy ezekre a közepekre is lehetnek igazak azok a tulajdonságok, ami az első háromra. Nagy kérdés, hogy mitől olyan jól kezelhető az első három közép, mi teszi őket kiemelkedővé. Részleges válasz két változóra már van. Ando és Kubo elmélete arra is rámutat például hogy a közepek között az aritmetikai a legnagyobb, és a harmonikus a legkisebb. Azaz kielégítenek egyfajta extremalitási tulajdonságot. De ez csak a kétváltozós eset, szeretnénk valamit tudni több változóban is. A vizsgálódás további lépcsői között szerepelhet a fenti iteráció számítógépes szimulációja nagyobb dimenziós mátrixokra, más eloszlásokra. Vagy más közepekre, ami magában

foglalja függvények mátrixmonotonitásának vizsgálatát is. Talán bizonyos közepekre sikerül majd bebizonyítani a nem rendezett esetben a konvergenciát hagyományos és korrekt matematikai módszerrel.

6. Melléklet A

Az octave program:

```
#! /usr/bin/octave -qf
#A fenti sor azt teszi lehetővé,
#hogya a program parancssorból futtatható legyen,
#de nem minden géptermi gépen működik ezzel a path-szal.

#Ez a program három szimmetrikus pozitív definit mátrixból, készít
#másikat, páronkénti logaritmikusan mátrixközép vételével.
#Ezt iterálja, és megnézi az eredményül kapott mátrixok normabeli eltérését.

#itt lehet választani, hogy milyen mátrixokkal kezdjük az iterációt
#kiindulomxvalasztas nevű változó, értelemszerű jelentéssel, lehet:
#"eloreadott", "veletlen" és "veletlenrendezett"

    kiindulomxvalasztas = "veletlen";

#itt lehet megadni három mátrixot kezdőértéknek
#kiindulomxvalasztas="eloreadott" esetén
A = [[50,3];[3,50]];
B = [[78,5];[5,78]];
C = [[100,1];[1,100]];

#itt lehet megadni a használt véletlen mátrixok méretét, és az iterációk számát
#n akkor számít, ha kiindulomxvalasztas="veletlen" vagy "veletlenrendezett"
n = 3; iteracioszam = 50;

#az alábbi függvény
#egy véletlen szimmetrikus nxn-es mátrixot készít
#N(0,1) eloszlás szerint
function kapottmx = szimvmlmx ( n );
    #egy teljesen véletlen mátrix, amit szimmetrikussá teszünk majd
    A = randn( n, n );
    #rand parancs E(0,1) eloszlás szerint csinálná

    #a főátló alatti rész transzponáltját
    #a főátlós alsóháromszög mátrixhoz adja
    A = tril( A , -1 )' + tril( A );
    kapottmx = A;
```

```

endfunction

#az alábbi függvény
# egy véletlen szimmetrikus pozitív definit nxn-es mátrixot készít
#N(0,1) eloszlás szerint
function kapottmx = posszimvmlmx ( n );
    #egy véletlen nemnegatív diagonális
    D=diag( abs(randn( 1, n )) );
    #bázistranszformációt hajt végre D-n
    A = szimvmlmx( n );
    kapottmx = A * D * A';
endfunction

#az alábbi függvény
#készít három szimmetrikus pozitív definit nxn-es mátrixot egy nagy mátrixban
#egy más mellé rendezve, kiindulomxvalasztas változó értéke szerint
function [retA,retB,retC] = posszimvmlmxek ( n , kiindulomxvalasztas );
    if (strcmp(kiindulomxvalasztas,"előreadott"));
        #az előreadott mátrixok
        retA=A;retB=B;retC=C;
    elseif (strcmp(kiindulomxvalasztas,"véletlen"));
        #három véletlen mátrix, hogy egyik kettő sem összehasonlítható
        do;
            retA=posszimvmlmx(n);
            retB=posszimvmlmx(n);
            retC=posszimvmlmx(n);
        until(rem(sum(eig(retA)<eig(retB)),n)<>0&&
            rem(sum(eig(retA)<eig(retC)),n)<>0&&
            rem(sum(eig(retB)<eig(retC)),n)<>0);
        #kiindulomxvalasztas = "véletlenrendezett"
    else
        #egy véletlen pozitív n-es sorvektor
        d1=abs(randn(1,n));
        #egy d1 diagonálisú és egy szimmetrikus véletlen mátrix
        D1=diag(d1);U1=szimvmlmx( n );
        #egy d1-nél elemenként nagyobb sorvektor
        d2=d1+abs(randn(1,n));
        #egy d2 diagonálisú és egy szimmetrikus véletlen mátrix
        D2=diag(d2);U2=szimvmlmx( n );
        #egy d3-nál is elemenként nagyobb sorvektor
        d3=d2+abs(randn(1,n));
        #egy d3 diagonálisú és egy szimmetrikus véletlen mátrix
        D3=diag(d3);U3=szimvmlmx( n );

#a rendezett mátrixok

```

```

        retA=U1*D1*U1;
        retB=U2*D2*U2;
        retC=U3*D3*U3;
    endif;
endfunction;

#az alábbi függvény
#kiszámolja egy négyzetes szimmetrikus mátrix (x-1)/log(x) függvényét,
#ami a kétváltozós logaritmus közephez tartozó egyváltozós függvény
function matrix = logkozfje( A );
    #u-ba teszi az A-t diagonalizáló unitérré,
    #és a d mátrixba a sajátértékekből álló diagonális mátrixot
    [u,d] = eig(A);
    #A mérete-szer csinálja
    for i = 1:size(d)(1,1);
        #d i-edik főátlóbeli elemének veszi (x-1)/log(x) függvényét
        d(i,i) = ( d(i,i)-1 )/( log(d(i,i)) );
    endfor;
    #a visszaadandó mátrixot készíti el; u' az u transzponáltja
    matrix = u*d*u';
endfunction

#itt megkapom a három kívánt mátrixot
[A,B,C] = posszimvlnxek( n , kiindulomxvalasztas );

#ékezetes sztringre konvertálja kiindulomxvalasztas változó értékét
if (strcmp(kiindulomxvalasztas,"előreadott"))
    kiindulomxvalasztas = " előreadott";
elseif (strcmp(kiindulomxvalasztas,"véletlen"))
    kiindulomxvalasztas = " véletlen";
else
    kiindulomxvalasztas = " véletlen rendezett";
endif
#kiírja a három kiinduló mátrixot
printf('\nA kiinduló%s mátrixok: \n\n',kiindulomxvalasztas);
A, B, C

#ez végzi az iterálást iterációs számszor
for i = 1:iteracioszam;

    #logm a mátrixgyökvonás
    #a, b, c jelenti a megfelelő logaritmus közepet

    a=sqrtm(A) * logkozfje( sqrtm(inv(A))* B *sqrtm(inv(A)) ) * sqrtm(A);

```

```

b=sqrtm(A) * logkozfje( sqrtm(inv(A))* C *sqrtm(inv(A)) ) * sqrtm(A);

c=sqrtm(B) * logkozfje( sqrtm(inv(B))* C *sqrtm(inv(B)) ) * sqrtm(B);

#az eredeti mátrixokat lecserélem a logaritmusos közép mátrixokra
A = a; B = b; C = c;
endfor;

#itt veszem az iteráció végén kapott mátrixok
#eltéréseik kettesnormájának maximumát
printf('\nAz eltérések 2-es normában %d lépés után:\n',iteracioszam);
printf('eltérés = %g',max( [ norm( B - A), norm( C - A ), norm( C - B ) ] ));
printf('\n\n');

```

7. Melléklet B

A shell szkript:

```

# Ez a shell program torli eredm fajlt,
# majd lefuttatja $1-szer $2-t es az eredmenyt beleirja
# $3-ba vagy eredm-be, azutan megjeleniti $3 fil-et less-ben
rm -f eredm;
if [[ $1 == "" ]];
then echo hasznalat: ismetel.sh [hanszor] mit [hova]; exit;
fi;
if [[ $1 != "" && $2 == "" ]]
then hanszor=1;
mit=$1;
fi;
if [[ $2 != "" ]]
then hanszor=$1;
mit=$2
fi;
if [[ $3 == "" ]];
then hova=eredm;
else hova=$3;
fi;
echo $hanszor-szer végezve $mit-t kaptuk: > $hova;
for iteracio in $(seq $hanszor);
do echo -e "\n$iteracio. kísérlet:" >> $hova;
echo -en "\rFolyamatban $iteracio. kísérlet";
$mit >> $hova;
done;
echo -en "\r";

```

```

awk 'BEGIN {max=-1000; szamolo=0; elofordulas=0}                                #
/^elt/ {                                                                           #
    split($3,elteres,"e");                                                         #
    if(elteres[2]!=""){                                                             #
        exponent=elteres[2]                                                         #
    }else{                                                                           #
        if(log(elteres[1])/log(10)>=0){                                             #
            exponent=int(log(elteres[1])/log(10))                                   #
        }else{                                                                       #
            exponent=int(log(elteres[1])/log(10))-1                                 #
        };                                                                            #
    };                                                                               #
    szamolo++;                                                                       #
    if(exponent>max){                                                                #
        max=exponent; elofordulas=szamolo                                           #
    }                                                                                 #
}                                                                                   #
END {printf("\nA maximális eltérés az összes kísérlet során                       #
10^%d nagyságrendű,\nés %d. kísérletnél fordult elő                               #
eloször",max,elofordulas)                                                         #
}                                                                                   #
' $hova>>$hova;                                                                     #
less $hova;

```

Hivatkozások

- [1] T.Ando C-K. Li and R. Mathias: *Geometric mean*, Linear Algebra Appl. **385**(2004), 305–334.
- [2] F. Kubo, T. Ando: *Means of Positive Linear Operators*, Math. Ann. **246**(1980), 205–224.
- [3] D. Petz: *Means of positive numbers and matrices*, előkészületben (2004).
- [4] T. Ando, F. Kubo: *Inequalities Among Operator Symmetric Function Means*, Signal processing, scattering and operator theory, and numerical methods, Amsterdam (1989), 535–542.
- [5] P.W. Michor, D. Petz and A. Andai: *On the curvature of a certain Riemannian space of matrices*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **3**(2000), 199–212.
- [6] A. O. Pittenger: *The logarithmic mean in n variables*, Amer. Math. Monthly **92**(1985), 99–104.