Eötvös Loránd Tudományegyetem,
Természettudományi Kar

Mátrix-monoton és mátrix-konvex függvények

BSc Szakdolgozat

Szabó Olivér Dániel
Matematika BSc, Alkalmazott matematikus szakirány

Témavezető: *Frenkel Péter,* adjunktus
Algebra és Számelmélet Tanszék
Eötvös Loránd Tudományegyetem, Természettudományi Kar

Budapest, 2015

Köszönetnyilvánítás

Ezúttal köszönöm témavezetőmnek, Frenkel Péternek, aki rendszeres konzultációk során készségesen segített a téma megértésében, és több bizonyítás levesetéséhez is hozzájárult.

Tartalomjegyzék

Bevezető **1**

1. Előkészületek **2**

1.1. Alapdefiníciók és állítások 2

1.2. Mátrix-monotonitás és mátrix-konvexitás definíciója 3

1.3. Nevezetes függvények vizsgálata 5

2. Kapcsolat mátrix-monotonitás és konvexitás között a pozitív félegyenesen **13**

3. A kapcsolat általánosítása (Löwner- és Krauss-mátrixok) **19**

3.1. Osztott differenciák 19

3.2. A Fréchet-derivált 21

3.3. Löwner-mátrixok 25

3.4. Krauss-mátrixok 30

4. Lieb konkávitás-tétele **36**

Hivatkozások **41**

Bevezető

A szakdolgozat célja a valós függvények tulajdonságainak egy kiterjesztését, a mátrix-monotonitást illetve a mátrix-konvexitást bemutatni és egy jól struktúrált egységként vizsgálni.

A mátrix-monoton függvények elméletét eleinte Löwner vizsgálta, majd Krauss a mátrix-konvex függvények körét tárta fel. Az önadjungált mátrixokon értelmezett függvények nagy szerepet játszanak a kvantum-infomációelméletben.

Már a klasszikus analízisben is a monotonitás és a konvexitás között szoros kapcsolat van, a mátrix függvényeken ezek a kapcsolatok megerősödnek. Ugyan erősebb állításokhoz juthatunk ennek révén, de ennek az az ára, hogy sokkal nehezebb kérdés megállapítani egy függvényről, hogy mátrix-monoton, illetve, hogy mátrix-konvex.

Rajendra Bhatia [1, I,V,IX,X fejezet] könyvében, illetve Frank Hansen [2] cikkében több bizonyítássorozat megtalálható, ezek meghatározzák témánk fő irányát, melyben minden nyílt intervallumra egy erős kapcsolatot mutatunk meg a mátrix-monotonitás és konvexitás között. Ezen felül az úton kitérünk az osztott differenciákra, és kiterjesztjük a deriválás fogalmát ezekre a függvényekre.

Végül kitérünk egy konkrét mátrix egyenlőtlenségre, itt bebizonyítjuk a Lieb [3] konkávitás-tételét. Eredetileg Lieb a Wigner-Yanase-Dyson feltevés bizonyításánál az $f : A \to A^q$ függvény konkávitásából ($0 \le q \le 1$) indult ki, mely szintén nem egy triviális kérdés, ezért volt egyik célunk azt is belátni útközben. Ezt a tételt B. Simon és T. Ando is bebizonyították, és az általuk adott, témánkhoz közelebb eső bizonyítást mutatjuk be.

1. Előkészületek

1.1. Alapdefiníciók és állítások

Ebben az alfejezetben kimondunk néhány gyakran használt definíciót, jelölést, és ezekhez szorosan kapcsolódó állítást. Ezek vagy a téma feldolgozását segítik, vagy gyakran használhatóak további definíciók és bizonyítások kimondásánál.

1.1. Definíció. *$M_n(\mathbb{K})$ Az $n \times n$-es négyzetes mátrixok halmaza $\mathbb{K}$ felett.*

1.2. Definíció. *$U \in M_n(\mathbb{C})$ unitér, ha $U^*U = UU^* = I$, azaz $U^* = U^{-1}$.*

1.3. Definíció. *$A \in M_n(\mathbb{C})$ önadjungált, ha $A^* = A$. Jelölés: $A \in M_n(\mathbb{C})^{sa}$.*

1.4. Állítás. *Legyenek $X, Y \in M_n(\mathbb{C})$ önadjungáltak. Ekkor XY önadjungált akkor, és csak akkor, ha $XY = YX$.*

1.5. Állítás. *Ha $A \in M_n(\mathbb{C})$ önadjungált, akkor $\exists$ U unitér és D valós diagonális mátrixok, hogy $A = U^*DU$.*

1.6. Állítás. *Ha $A \in M_n(\mathbb{C})$ önadjungált, akkor A minden sajátértéke valós.*

1.7. Definíció. *$x, y \in \mathbb{C}^n$ vektorok skalárszorzata $\langle x, y\rangle = \sum\limits_{i=1}^{n} x_i\overline{y_i}$.*

1.8. Definíció. *$X, Y \in M_n(\mathbb{C})$ mátrixok skalárszorzata $\langle X, Y\rangle = tr(X^*Y)$.*

1.9. Definíció. *$A \in M_n(\mathbb{C})$ pozitív definit (jel: $A > 0$), ha önadjungált és*

$$\langle Ax, x\rangle > 0 \;,\; \forall x \in \mathbb{C}^n \backslash \{\underline{0}\}.$$

1.10. Definíció. *$A \in M_n(\mathbb{C})$ pozitív szemidefinit (jel: $A \geq 0$), ha önadjungált és*

$$\langle Ax, x\rangle \geq 0 \;,\; \forall x \in \mathbb{C}^n.$$

1.11. Állítás. *A pozitív definit $\Leftrightarrow$ A sajátértékei pozitívak és A önadjungált.*

1.12. Állítás. *A pozitív szemidefinit $\Leftrightarrow$ A sajátértékei nemnegatívak és A önadjungált.*

1.13. Állítás. *A pozitív szemidefinit $\Rightarrow$ A átlóelemei nemnegatívak.*

1.14. Definíció. *Legyen $A \in M_n(\mathbb{C})$. Ekkor A operátornormája:*

$$\|A\| = \sup\{\|Av\| : \|v\| = 1, v \in \mathbb{C}^n\}$$

1.15. Definíció. *Legyen $A \in M_n(\mathbb{C})$. Ekkor A spektrálsugara*

$$\varrho(A) = max\{|\lambda| \colon \lambda \text{ sajátértéke } A\text{-nak}\}.$$

1.16. Állítás. *Legyen $A \in M_n(\mathbb{C})$. Ekkor $\varrho(A) \leq \|A\|$. Ha A normális ($AA^* = A^*A$), speciálisan ha A önadjungált, akkor itt egyenlőség áll fenn.*

Megjegyzés: Ez nem minden normára igaz, például a Frobenius-normára nem teljesül.

1.17. Állítás. *Legyen $A, B \in M_n(\mathbb{K})$, és B invertálható. Ekkor*

$$\varrho(AB) = \varrho(BA).$$

Bizonyítás:

Vegyük észre, hogy AB és BA hasonlóak:

$$AB \sim BA \Leftrightarrow \exists X \in M_n(\mathbb{K}) : BA = X(AB)X^{-1}.$$

Ebben az esetben az $X = B$ választás jó lesz.

Mivel AB és BA hasonlóak, ezért a karakterisztikus polinomjuk megegyezik, azaz

$$det(BA - \lambda I) = \chi_{BA}(\lambda) = \chi_{AB}(\lambda) = det(AB - \lambda I).$$

Ezzel készen vagyunk, ugyanis ha ugyanaz AB és BA karakterisztikus polinomja, akkor a sajátértékeik is mind megegyeznek, ebből meg következik, hogy a spektrálsugaruk is azonos.

1.2. Mátrix-monotonitás és mátrix-konvexitás definíciója

Az önadjungált mátrixok igen jó tulajdonságokkal bírnak, és a rajtuk értelmezett függvények pedig számos alkalmazáshoz vezetnek. Ennek tükrében ezeket fogjuk vizsgálni a továbbiakban.

1.18. Definíció. *Legyenek $X, Y \in M_n(\mathbb{C})$ önadjungáltak. Ekkor $X \geq Y \Leftrightarrow X - Y \geq 0$, azaz $X - Y$ pozitív szemidefinit.*

1.19. Definíció. *Legyen $A \in M_n(\mathbb{C})$ önadjungált, tehát unitér mátrixszal diagonalizálható ($A = U^*DU$, ahol U unitér és D valós diagonális mátrix). Legyen I intervallum. Ha $f : I \to \mathbb{R}$ függvény, és A sajátértékei I-ben vannak, akkor $f(A) = U^*f(D)U$, ahol $f(D)$ diagonális és $f(D)_{i,i} = f(d_{i,i})$.*

Megjegyzés: $f(A)$ nem függ U választásától.

1.20. Példa. $f(x) = |x|$ *(abszolútérték)*

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

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

1.21. Definíció. *Legyen I intervallum. Az f függvény n-monoton növő ($n \geq 1$) I-n, ha minden olyan $X, Y \in M_n(\mathbb{C})$ önadjungált mátrixra, melyek sajátértékei I-ben vannak,*

$$X \leq Y \Rightarrow f(X) \leq f(Y) \ .$$

f n-monoton csökkenő ha $X \leq Y$ esetén $f(X) \geq f(Y)$.

Mivel f n-monoton csökkenő ha $-f$ n-monoton növő, ezért általában a monoton növő esettel elég foglalkozni.

1.22. Definíció. *Legyen I intervallum. Az f függvény n-konvex ($n \geq 1$) I-n, ha minden olyan $X, Y \in M_n(\mathbb{C})$ önadjungált mátrixra, melyek sajátértékei I-ben vannak*

$$\forall \lambda \in (0,1) \text{ esetén } f(\lambda X + (1-\lambda)Y) \leq \lambda f(X) + (1-\lambda) f(Y) \ .$$

f n-konkáv, ha $-f$ n-konvex.

Megjegyzés: $\lambda X + (1-\lambda)Y$ sajátértékei mind I-ben vannak, tehát f értelmezhető rajta.

1.23. Állítás. *Ha az f függvény folytonos, és $\lambda = \frac{1}{2}$ -re teljesül a konvexitás feltétele (azaz f gyengén konvex), akkor minden $\lambda \in [0,1]$ -re is teljesül.*

Megjegyzés: Ha egy $f : I \to \mathbb{R}$ ($I \subseteq \mathbb{R}$) függvény n-monoton (illetve n-konvex), akkor minden $m \leq n$, $m \in \mathbb{N}^+$ esetén f m-monoton (illetve m-konvex).

1.24. Állítás. *n-monoton illetve n-konvex függvények konvergens sorozatának pontonkénti limesze szintén n-monoton illetve n-konvex.*

1.25. Definíció. *Az f függvény operátor-monoton illetve operátor-konvex ha n-monoton illetve n-konvex minden $n \in \mathbb{N}$ esetén.*

1.26. Állítás. *Operátor-monoton illetve operátor-konvex függvények konvergens sorozatának pontonkénti limesze szintén operátor-monoton illetve operátor-konvex.*

1.3. Nevezetes függvények vizsgálata

Ebben a részben több ismert függvényt fogunk vizsgálni, Rajendra Bhatia [1, V. fejezet] könyvének feldolgozásával eljutunk az A^q függvényig ($0 \leq q \leq 1$), melyet további eredmények eléréséhez például Lieb használt fel [3] -ban elért eredményeihez. Mivel viszonylag egyszerű (vagy annak tűnő) függvényekről van szó, ezért könnyen alkalmazhatóak tételek illetve állítások bizonyításaiban.

1.27. Példa. *Az $f(X) = X^2$ függvény operátor-konvex $\mathbb{R}$-en.*
Bizonyítás:
Legyen $A, B \in M_n(\mathbb{C})$ önadjungált.
Ekkor azt kell belátni hogy $\forall \lambda \in (0,1)$:

$$(\lambda A + (1-\lambda)B)^2 \leq \lambda A^2 + (1-\lambda)B^2,$$

ami ekvivalens azzal, hogy: $\lambda A^2 + (1-\lambda)B^2 - (\lambda A + (1-\lambda)B)^2 \geq 0$. Innen

$$\begin{gathered}
\lambda A^2 + (1-\lambda)B^2 - (\lambda A + (1-\lambda)B)^2 = \\
\lambda A^2 + (1-\lambda)B^2 - [\lambda^2 A^2 + \lambda(1-\lambda)AB + \lambda(1-\lambda)BA + (1-\lambda)^2 B^2] = \\
(\lambda - \lambda^2)A^2 - \lambda(1-\lambda)AB - \lambda(1-\lambda)BA + ((1-\lambda) - (1-\lambda)^2)B^2 = \\
\lambda(1-\lambda)A^2 - \lambda(1-\lambda)AB - \lambda(1-\lambda)BA + (1-\lambda)(1-(1-\lambda))B^2 = \\
\lambda(1-\lambda)[A^2 - AB - BA + B^2] = \\
\lambda(1-\lambda)(A-B)^2 \geq 0,
\end{gathered}$$

viszont az utolsó alak $\forall\ A, B$-re fennáll, tehát az eredeti állítás is igaz.

1.28. Példa. *Az $f(X) = X^2$ függvény nem operátor-monoton növő $[0, \infty)$-n.*
Bizonyítás: ellenpéldával.

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, A - B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$A^2 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}, B^2 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, A^2 - B^2 = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}.$$

Ekkor $A \geq B$ esetén $A^2 \geq B^2$ nem teljesül, ugyanis $det(A^2 - B^2) < 0$.

Amennyiben ez nem elég meggyőző példa (ugyanis B-nek van 0 sajátértéke), akkor [5, p. 35]-ban találhatunk egy példát ugyanerre, amelyben A és B már pozitív definit mátrixok.

Emiatt igen fontos, hogy csak óvatosan vonhatunk le következtetéseket a mátrix függvényekre a skalár függvények köréből.

Megjegyzés: Emiatt az $f(X) = X^2$ függvény legfeljebb 1-konvex lehet, ami igaz is.

Megjegyzés: A mátrixokon vett négyzetre emelés megegyezik a sajátértékek négyzetre emelésével. Legyen A önadjungált, ekkor diagonalizálható, így

$$A^2 = (U^*DU)(U^*DU) = U^*D(UU^*)DU = U^*DDU = U^*f(D)U = f(A),$$

és az n-ik hatványra emelés esetére ($n \in \mathbb{N}$) is hasonlóan fennáll ez a kapcsolat.

1.29. Példa. *Az $f(X) = X^3$ függvény nem operátor-monoton növő $[0, \infty)$-n.*

Bizonyítás: ellenpéldával.

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

$$A^3 = \begin{pmatrix} 34 & 14 \\ 14 & 6 \end{pmatrix}, B^3 = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}, A^3 - B^3 = \begin{pmatrix} 30 & 10 \\ 10 & 2 \end{pmatrix}.$$

Ekkor $A \geq B$ esetén $A^3 \geq B^3$ nem teljesül, ugyanis det$(A^3 - B^3) = -40 < 0$.

1.30. Példa. *Az $f(X) = X^3$ függvény nem operátor-konvex $[0, \infty)$-en*

Bizonyítás: ellenpéldával.

Legyen A, B ugyanaz, mint az 1.29 példában. Ekkor

$$\frac{A^3 + B^3}{2} - \left(\frac{A+B}{2}\right)^3 = \begin{pmatrix} 19 & 9 \\ 9 & 5 \end{pmatrix} - \begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix} = \begin{pmatrix} 6 & 1 \\ 1 & 0 \end{pmatrix}.$$

Mivel det $\begin{pmatrix} 6 & 1 \\ 1 & 0 \end{pmatrix} < 0$, ezért nem teljesül, hogy $\frac{A^3+B^3}{2} \geq \left(\frac{A+B}{2}\right)^3$.

A következő lemmával további függvények fontos tulajdonságait tudjuk bizonyítani.

1.31. Lemma. *Legyenek $A, B \in M_n(\mathbb{C})$ önadjungáltak. Ekkor ha $B \geq A$, akkor minden $K \in M_n(\mathbb{C})$ esetén $K^*BK \geq K^*AK$.*

Bizonyítás:

Legyen $u \in \mathbb{C}$. Ekkor minden u-ra fennáll a következő:

$$\langle u, K^*(B-A)Ku \rangle = \langle Ku, (B-A)Ku \rangle \geq 0,$$

mert feltettük, hogy $(B - A)$ pozitív szemidefinit.

1.32. Állítás. *Az $f(X) = X^{-1}$ függvény operátor-monoton csökkenő $(0, \infty)$-en.*

Bizonyítás:

Legyen $0 < A \leq B$. Azt kell belátni, hogy $A^{-1} \geq B^{-1}$.

Tudjuk, hogy

$$B \geq A.$$

Mivel B invertálható, ezért alkalmazzuk *az* 1.31 lemmát $K = B^{-\frac{1}{2}}$ -re, így megkapjuk, hogy

$$B^{-\frac{1}{2}} B B^{-\frac{1}{2}} \geq B^{-\frac{1}{2}} A B^{-\frac{1}{2}},$$

azaz

$$I \geq B^{-\frac{1}{2}} A B^{-\frac{1}{2}}. \tag{1}$$

Állítás: $0 \leq X \leq I \Leftrightarrow X^{-1} \geq I$

Ez azért teljesül, mert $X = U^* D U$ sajátértékei $[0, 1]$-ben vannak, tehát X^{-1} sajátértékei (ezek reciprokai) $[1, \infty)$-ben vannak, így $X^{-1} - I = U^* D^{-1} U - U^* I U = U^*(D^{-1} - I)U$ sajátértékei nemnegatívak.

Visszatérve a bizonyításhoz, (1)-re alkalmazhatjuk az előbbi kisebb állításunkat, így megkapjuk, hogy

$$I \leq (B^{-\frac{1}{2}} A B^{-\frac{1}{2}})^{-1} = B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}}. \tag{2}$$

Mivel B invertálható, ezért újra alkalmazzuk az 1.31 lemmát, szintén $K = B^{-\frac{1}{2}}$ -re, így megkapjuk (2)-ből, hogy

$$B^{-\frac{1}{2}} I B^{-\frac{1}{2}} \leq B^{-\frac{1}{2}} B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} B^{-\frac{1}{2}},$$

azaz

$$B^{-1} \leq A^{-1},$$

tehát az inverzképzés operátor-monoton csökkenő $(0, \infty)$-en.

1.33. Állítás. *Az $f(X) = X^{-1}$ függvény operátor-konvex* $(0, \infty)$*-en.*

Bizonyítás:

Felhasználjuk az inverzképzés folytonosságát a $(0, \infty)$ intervallumon. Így elég csak $\lambda = \frac{1}{2}$-re belátni a konvexitást (azaz a gyenge konvexitást). Legyenek $A, B \in M_n(\mathbb{C})$ önadjungáltak, és $0 < A \leq B$. Azt kell belátni, hogy

$$\frac{A^{-1} + B^{-1}}{2} - \left(\frac{A+B}{2}\right)^{-1} \geq 0.$$

A következő egyenlőséget fogjuk ezért belátni:

$$\frac{A^{-1} + B^{-1}}{2} - \left(\frac{A+B}{2}\right)^{-1} = \frac{(A^{-1} - B^{-1})(A^{-1} + B^{-1})^{-1}(A^{-1} - B^{-1})}{2}. \tag{3}$$

Ha ezt belátjuk, akkor ebből következik az állítás, ugyanis $(A^{-1}+B^{-1})^{-1}$ pozitív definit, $(A^{-1} - B^{-1})$ önadjungált (a monotonitásból következik az is, hogy pozitív szemidefinit, de ez nem szükséges), így az egész jobb oldal pozitív szemidefinit, tehát a bal oldal is, mert megegyeznek.

Szorozzuk meg az (3) egyenletet jobbról $2(A + B)$-vel. Ekkor

$$(2I + A^{-1}B + B^{-1}A) - 4I = (A^{-1} - B^{-1})(A^{-1} + B^{-1})^{-1}(A^{-1}B - B^{-1}A),$$

azaz

$$A^{-1}B + B^{-1}A - 2I = (A^{-1} - B^{-1})(A^{-1} + B^{-1})^{-1}(A^{-1}B - B^{-1}A).$$

A bal oldalt tovább alakítva azt kapjuk, hogy

$$(A^{-1} - B^{-1})(B - A) = (A^{-1} - B^{-1})(A^{-1} + B^{-1})^{-1}(A^{-1}B - B^{-1}A).$$

Innen elég belátni, hogy

$$(B - A) = (A^{-1} + B^{-1})^{-1}(A^{-1}B - B^{-1}A).$$

Szorozzuk meg mindkét oldalt balról $(A^{-1} + B^{-1})$-zel. Ekkor

$$(A^{-1} + B^{-1})(B - A) = A^{-1}B - B^{-1}A.$$

Ez láthatóan minden A, B-re fennáll, tehát az eredeti egyenlőség is.

Megjegyzés: Ebben az esetben is helytálló az $f(X) = X^{-1}$ jelölés az $f(t) = \frac{1}{t}$ -re, ugyanis ha $X = UDU^*$, akkor

$$Xf(X) = UDU^*Uf(D)U^* = UD(U^*U)D^{-1}U^* = UDD^{-1}U^* = UU^* = I,$$

tehát $f(X)$ valóban az X inverze.

1.34. Lemma. *Legyen B pozitív definit és $B \geq A \geq 0$. Ekkor $\|A^{\frac{1}{2}}B^{-\frac{1}{2}}\| \leq 1$.*

Bizonyítás:

Mivel B invertálható, ezért alkalmazzuk *az* 1.31 lemmát $K = B^{-\frac{1}{2}}$ -re, így megkapjuk, hogy

$$B \geq A \Rightarrow B^{-\frac{1}{2}}BB^{-\frac{1}{2}} \geq B^{-\frac{1}{2}}AB^{-\frac{1}{2}},$$

azaz

$$I \geq B^{-\frac{1}{2}}AB^{-\frac{1}{2}}.$$

Ezt a következőképp alakítsuk tovább:

$$I \geq B^{-\frac{1}{2}}AB^{-\frac{1}{2}} = B^{-\frac{1}{2}}A^{\frac{1}{2}}A^{\frac{1}{2}}B^{-\frac{1}{2}} = \left(A^{\frac{1}{2}}B^{-\frac{1}{2}}\right)^* A^{\frac{1}{2}}B^{-\frac{1}{2}}.$$

Felhasználjuk, hogy egy K mátrix kontrakció ($\|K\| \leq 1$) akkor és csak akkor, ha $K^*K \leq I$ ($\Leftrightarrow KK^* \leq I$), így megkapjuk a lemma állítását, azaz $\|A^{\frac{1}{2}}B^{-\frac{1}{2}}\| \leq 1$.

1.35. Állítás. *Az $f(X) = X^{\frac{1}{2}}$ függvény operátor-monoton növő $[0, \infty)$-n.*

Bizonyítás:

Legyen $0 \leq A \leq B$ és tegyük fel, hogy B invertálható ($B > 0$). Ekkor az 1.34 lemma pont azt mondja ki, hogy

$$\|A^{\frac{1}{2}}B^{-\frac{1}{2}}\| \leq 1.$$

Az operátornormát becsülhetjük alulról a spektrálsugárral, így

$$\varrho(A^{\frac{1}{2}}B^{-\frac{1}{2}}) \leq \|A^{\frac{1}{2}}B^{-\frac{1}{2}}\| \leq 1.$$

Felhasználjuk, hogy $A^{\frac{1}{2}}B^{-\frac{1}{2}}$ és $B^{-\frac{1}{4}}A^{\frac{1}{2}}B^{-\frac{1}{4}}$ hasonlóak (és feltevés szerint léteznek), ezért a következőt állíthatjuk:

$$\varrho(A^{\frac{1}{2}}B^{-\frac{1}{2}}) = \varrho(B^{-\frac{1}{4}}A^{\frac{1}{2}}B^{-\frac{1}{4}}).$$

Tehát eddig beláttuk, hogy $\varrho(B^{-\frac{1}{4}}A^{\frac{1}{2}}B^{-\frac{1}{4}}) \leq 1$, ez azt jelenti, hogy

$$\varrho(B^{-\frac{1}{4}}A^{\frac{1}{2}}B^{-\frac{1}{4}}) = \|B^{-\frac{1}{4}}A^{\frac{1}{2}}B^{-\frac{1}{4}}\| \leq 1,$$

ugyanis $B^{-\frac{1}{4}}A^{\frac{1}{2}}B^{-\frac{1}{4}}$ már önadjungált. Ezért

$$B^{-\frac{1}{4}}A^{\frac{1}{2}}B^{-\frac{1}{4}} \leq I. \tag{4}$$

Most az (4) egyenletre alkalmazzuk az 1.31 lemmát $K = B^{\frac{1}{4}}$-del, így

$$B^{\frac{1}{4}}B^{-\frac{1}{4}}A^{\frac{1}{2}}B^{-\frac{1}{4}}B^{\frac{1}{4}} \leq B^{\frac{1}{4}}IB^{\frac{1}{4}},$$

azaz

$$A^{\frac{1}{2}} \leq B^{\frac{1}{2}}.$$

Ezzel megkaptuk, hogy a gyökvonás operátor-monoton növő $(0,\infty)$-n.

Abban az esetben, ha B szinguláris, azt tudjuk, hogy $\varepsilon > 0$-ra

$$A^{\frac{1}{2}} \leq (B + \varepsilon I)^{\frac{1}{2}}.$$

Innen $\varepsilon \to 0$ esetén szintén megkapjuk, hogy $A^{\frac{1}{2}} \leq B^{\frac{1}{2}}$.

1.36. Állítás. *Legyen* $0 \leq q \leq 1$, $q \in \mathbb{R}$. *Az* $f(X) = X^q$ *függvény operátor-monoton növő* $[0,\infty)$*-n.*

Bizonyítás:

Legyen $r = \frac{m}{2^n}$ alakú valós szám, ahol $1 \leq m \leq 2^n$ egész és $n \in \mathbb{N}$. Az ilyen alakú r számokról látjuk be, hogy $f(X) = X^r$ operátor-monoton növő.

Teljes indukcióval n-re fogjuk belátni az állítást.

Indukciós feltevés: Minden $r = \frac{m}{2^j}$ ($j \in \{1,\ldots,n\}$ és $m \in \{1,\ldots,2^n\}$) esetén

$$0 \leq A \leq B \Rightarrow A^r \leq B^r,$$

azaz f operátor-monoton növő.

Ha $n = 1$, ekkor $j = 1$ és $m \in \{1,2\}$. Két eset van, az elsőben $r = \frac{1}{2}$, azaz a gyökvonás monotonitásáról van szó, ezt már beláttuk 1.35 -ben. A második esetben $r = 1$, azaz az identitás függvényt kapjuk, erre triviálisan fennáll.

Tegyük fel, hogy $n = k - 1$ -ig igaz, bizonyítsuk $n = k$-ra.

Tudjuk, hogy $j \in \{1, \ldots, k-1\}$ és $m \in \{1, \ldots, 2^{k-1}\}$ esetén igaz az indukciós állítás $r = \frac{m}{2^j}$-ra.

Legyen $0 < A \leq B$. (ii. esethez szükséges feltenni).

Most két esetre bontsuk a bizonyítást:

(i) Ha $m \in \{1, \ldots, 2^{k-1}\}$, akkor az indukciós feltevésből

$$A^{\frac{m}{2^{k-1}}} \leq B^{\frac{m}{2^{k-1}}}$$

Mivel a gyökvonásról beláttuk, hogy operátor-monoton növő 1.35-ben, ezért

$$A^{\frac{m}{2^k}} = \left(A^{\frac{m}{2^{k-1}}}\right)^{\frac{1}{2}} \leq \left(B^{\frac{m}{2^{k-1}}}\right)^{\frac{1}{2}} = B^{\frac{m}{2^k}}$$

(ii) Ha $m \in \{(2^{k-1}+1), \ldots, 2^k\}$, akkor mivel $0 < A \leq B$, ezért (felhasználva, hogy az inverzképzés operátor-monoton csökkenő 1.32 szerint)

$$A^{-1} \geq B^{-1}.$$

Innen alkalmazzuk az 1.31 lemmát $K = B^{\frac{m}{2^k}}$ -val:

$$B^{\frac{m}{2^k}} A^{-1} B^{\frac{m}{2^k}} \geq B^{\frac{m}{2^k}} B^{-1} B^{\frac{m}{2^k}},$$

azaz

$$B^{\frac{m}{2^k}} A^{-1} B^{\frac{m}{2^k}} \geq B^{\frac{m}{2^{k-1}}-1}. \tag{5}$$

Mivel $\frac{m}{2^{k-1}} - 1 = \frac{m-2^{k-1}}{2^{k-1}}$ és $m > 2^{k-1}$, ezért $(m - 2^{k-1}) \in \{1 \ldots 2^{k-1}\}$.

Emiatt az indukciós feltevéssel ezt alulról tovább becsülhetjük azzal, hogy

$$B^{\frac{m}{2^{k-1}}-1} \geq A^{\frac{m}{2^{k-1}}-1}. \tag{6}$$

Összesítve az (5) és (6)-ban kapott eredményeket, azt kapjuk, hogy

$$B^{\frac{m}{2^k}} A^{-1} B^{\frac{m}{2^k}} \geq A^{\frac{m}{2^{k-1}}-1}. \tag{7}$$

Még egyszer alkalmazzuk az 1.31 lemmát, a (7) egyenlőtlenségre, most $K = A^{-\frac{1}{2}}$ -del:

$$A^{-\frac{1}{2}} B^{\frac{m}{2^k}} A^{-1} B^{\frac{m}{2^k}} A^{-\frac{1}{2}} \geq A^{-\frac{1}{2}} A^{\frac{m}{2^{k-1}}-1} A^{-\frac{1}{2}}. \tag{8}$$

Ezt átalakíthatjuk a következőképpen:

$$\left(A^{-\frac{1}{2}}B^{\frac{m}{2^k}}A^{-\frac{1}{2}}\right)^2 \geq A^{\frac{m}{2^{k-1}}-2}.$$

Felhasználjuk újra, hogy a gyökvonás operátor-monoton növő 1.35 szerint. Így

$$A^{-\frac{1}{2}}B^{\frac{m}{2^k}}A^{-\frac{1}{2}} \geq \left(A^{\frac{m}{2^{k-1}}-2}\right)^{\frac{1}{2}} = A^{\frac{m}{2^k}-1}.$$

Szintén az 1.31 lemma segítségével $K = A^{\frac{1}{2}}$-ra leolvasható, hogy

$$A^{\frac{1}{2}}A^{-\frac{1}{2}}B^{\frac{m}{2^k}}A^{-\frac{1}{2}}A^{\frac{1}{2}} \geq A^{\frac{1}{2}}A^{\frac{m}{2^k}-1}A^{\frac{1}{2}},$$

azaz

$$B^{\frac{m}{2^k}} \geq A^{\frac{m}{2^k}}.$$

Eredményképp beláttuk, hogy $r = \frac{m}{2^n}$ $(m \in \{1, \ldots, 2^n\})$ esetén

$$0 < A \leq B \Rightarrow A^r \leq B^r.$$

Mivel az $\frac{m}{2^n}$ alakú racionális számok sűrűn helyezkednek el $[0,1]$-en, ezért ez minden $q \in [0,1]$-re is teljesül.

Továbbá $0 \leq A \leq B$ esetre az állítás f folytonossága miatt teljesül.

2. Kapcsolat mátrix-monotonitás és konvexitás között a pozitív félegyenesen

Ebben a részben összekapcsoljuk a mátrix-monotonitást a mátrix-konvexitással. Az itt található bizonyítások nagy része Frank Hansen [2] cikkéből származnak. Olyan tételek kerülnek bizonyításra, melyek segítségével egy függvény már ismert tulajdonságából (monotonitás vagy konvexitás) a másikról valamilyen információt nyerünk. Ezek segítségével nem lehet megtudni bármilyen függvény összes tulajdonságát, viszont arra a későbbi részekben mutatunk be eszközöket.

2.1. Állítás. *Legyen I nyílt intervallum, $f : I \to \mathbb{R}$ konkáv/konvex. Ekkor f folytonos.*

2.2. Lemma. *Legyen $D \in M_n(\mathbb{C}), \|.\|$ operátornorma, és $\varepsilon, \lambda > 0$. Ha $\varepsilon\lambda > \|D\|^2$ akkor*

$$\begin{pmatrix} \varepsilon I & D \\ D & \lambda I \end{pmatrix} \geq 0.$$

Bizonyítás. Legyen $(u, v)^\top \in \mathbb{C}^{2n}$.
Azt kell belátni, hogy

$$\varepsilon\langle u, u\rangle + \lambda\langle v, v\rangle + \langle u, Dv\rangle + \langle Du, v\rangle \geq 0. \tag{9}$$

A Cauchy-Schwarz egyenlőtlenséggel becsülhetjük, hogy

$$|\langle u, Dv\rangle| \leq |u|\,|Dv| \leq \|D\|\,|u|\,|v|,$$

azaz ha felbontjuk az abszolútértéket akkor $-\|D\|\,|u|\,|v| \leq \langle u, Dv\rangle$.
Hasonlóan megkapjuk, hogy $-\|D\|\,|u|\,|v| \leq \langle Du, v\rangle$.
Tehát (9)-et becsülve

$$\varepsilon\langle u, u\rangle + \lambda\langle v, v\rangle + \langle u, Dv\rangle + \langle Du, v\rangle \geq \varepsilon|u|^2 + \lambda|v|^2 - 2\|D\|\,|u|\,|v|.$$

Most használjuk fel, hogy $\varepsilon\lambda > \|D\|^2$, így

$$\varepsilon|u|^2 + \lambda|v|^2 - 2\|D\|\,|u|\,|v| \geq \varepsilon|u|^2 + \lambda|v|^2 - 2\sqrt{\varepsilon\lambda}\,|u|\,|v| = \left(\sqrt{\varepsilon}|u| - \sqrt{\lambda}|v|\right)^2 \geq 0.$$

2.3. Tétel. *Legyen $f : (a, \infty) \to \mathbb{R}$ $2n$-monoton növő függvény $(a \in \mathbb{R})$. Ekkor f folytonos és n-konkáv (a, ∞)-en.*

Bizonyítás. Feltehető, hogy $a = 0$. Legyen $X_1, X_2 \in M_n(\mathbb{C})$ pozitív definit és $s \in [0, 1]$. Legyen $V \in M_{2n}(\mathbb{C})$ a következő:

$$V = \begin{pmatrix} \sqrt{s} I_n & -\sqrt{(1-s)} I_n \\ \sqrt{(1-s)} I_n & \sqrt{s} I_n \end{pmatrix},$$

ahol $I_n \in M_n(\mathbb{C})$ egységmátrix. A V mátrix unitér, ugyanis $VV^* = V^*V = I_{2n}$. Innen számoljuk ki, hogy

$$V^* \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} V = \begin{pmatrix} sX_1 + (1-s)X_2 & \sqrt{s}\sqrt{1-s}(X_2 - X_1) \\ \sqrt{s}\sqrt{1-s}(X_2 - X_1) & (1-s)X_1 + sX_2 \end{pmatrix}.$$

Legyen $\lambda I_n \geq (1-s)X_1 + sX_2$. Ekkor adott $\varepsilon > 0$-ra fennáll:

$$\begin{pmatrix} sX_1 + (1-s)X_2 + \varepsilon I_n & 0 \\ 0 & 2\lambda I_n \end{pmatrix} - \begin{pmatrix} sX_1 + (1-s)X_2 & \sqrt{s}\sqrt{1-s}(X_2 - X_1) \\ \sqrt{s}\sqrt{1-s}(X_2 - X_1) & (1-s)X_1 + sX_2 \end{pmatrix} \tag{10}$$

$$\geq \begin{pmatrix} \varepsilon I_n & -\sqrt{s}\sqrt{1-s}(X_2 - X_1) \\ -\sqrt{s}\sqrt{1-s}(X_2 - X_1) & \lambda I_n \end{pmatrix}. \tag{11}$$

Ez azért igaz, mert bal oldalra rendezve azt kapjuk, hogy

$$\begin{pmatrix} 0 & 0 \\ 0 & \lambda I_n - (1-s)X_1 - sX_2 \end{pmatrix} \geq 0,$$

és ez λ definíciójából adódóan teljesül.

A 2.2 lemma szerint (11) pozitív szemidefinit, ha $\varepsilon\lambda \geq ||\sqrt{s}\sqrt{1-s}(X_2 - X_1)||_2^2$.

Elég nagy λ-ra azt kapjuk (10)-ből, hogy

$$\begin{pmatrix} sX_1 + (1-s)X_2 + \varepsilon I_n & 0 \\ 0 & 2\lambda I_n \end{pmatrix} - \begin{pmatrix} sX_1 + (1-s)X_2 & \sqrt{s}\sqrt{1-s}(X_2 - X_1) \\ \sqrt{s}\sqrt{1-s}(X_2 - X_1) & (1-s)X_1 + sX_2 \end{pmatrix} \geq 0,$$

azaz

$$\begin{pmatrix} sX_1 + (1-s)X_2 + \varepsilon I_n & 0 \\ 0 & 2\lambda I_n \end{pmatrix} \geq \begin{pmatrix} sX_1 + (1-s)X_2 & \sqrt{s}\sqrt{1-s}(X_2 - X_1) \\ \sqrt{s}\sqrt{1-s}(X_2 - X_1) & (1-s)X_1 + sX_2 \end{pmatrix}.$$

Mivel f $2n$-monoton növő, ezért alkalmazhatjuk mindkét oldalra (itt használjuk fel hogy f értelmezési tartománya kiterjed a végtelenbe) és továbbra is fennáll

$$\begin{pmatrix} f(sX_1 + (1-s)X_2 + \varepsilon I_n) & 0 \\ 0 & f(2\lambda I_n) \end{pmatrix} \geq f \begin{pmatrix} sX_1 + (1-s)X_2 & \sqrt{s}\sqrt{1-s}(X_2 - X_1) \\ \sqrt{s}\sqrt{1-s}(X_2 - X_1) & (1-s)X_1 + sX_2 \end{pmatrix}. \tag{12}$$

A (12) egyenlőtlenség jobb oldaláról tudjuk a következőket:

$$f\begin{pmatrix} sX_1+(1-s)X_2 & \sqrt{s}\sqrt{1-s}(X_2-X_1) \\ \sqrt{s}\sqrt{1-s}(X_2-X_1) & (1-s)X_1+sX_2 \end{pmatrix} = f\left(V^*\begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}V\right) =$$
$$V^*\begin{pmatrix} f(X_1) & 0 \\ 0 & f(X_2) \end{pmatrix} = \begin{pmatrix} sf(X_1)+(1-s)f(X_2) & \sqrt{s}\sqrt{1-s}(f(X_2)-f(X_1)) \\ \sqrt{s}\sqrt{1-s}(f(X_2)-f(X_1)) & (1-s)f(X_1)+sf(X_2) \end{pmatrix}.$$

Tehát

$$\begin{pmatrix} f(sX_1+(1-s)X_2+\varepsilon I_n) & 0 \\ 0 & f(2\lambda I_n) \end{pmatrix} \geq$$
$$\begin{pmatrix} sf(X_1)+(1-s)f(X_2) & \sqrt{s}\sqrt{1-s}(f(X_2)-f(X_1)) \\ \sqrt{s}\sqrt{1-s}(f(X_2)-f(X_1)) & (1-s)f(X_1)+sf(X_2) \end{pmatrix}.$$

Vegyük észre, hogy itt a bal felső részmátrixokra:

$$f(sX_1+(1-s)X_2+\varepsilon I_n) \geq sf(X_1)+(1-s)f(X_2), \tag{13}$$

ugyanis ha bal oldalra rendeznénk az előzőt, akkor szükséges a pozitív szemidefinitséghez, hogy az átlóelemek nemnegatívak legyenek (ennek blokkosított változata, hogy az átlóelemek pozitív szemidefinit mátrixok).

Innen már csak azt kell belátni, hogy εI_n elhagyásával is fennáll a (13) egyenlőtlenség. Ezt úgy kapjuk meg, hogy belátjuk f-ről, hogy folytonos.

Jelölje f jobboldali határértékét $t \in \mathbb{R}$-ben:

$$f_+(t) = \lim_{\omega\to 0} f(t+\omega), \text{ ahol } \omega > 0,\ t > a.$$

Legyen $t_1, t_2 \in \mathbb{R}$, $t_1, t_2 > 0$. Ekkor f monotonitásából megkapjuk, hogy

$$sf_+(t_1)+(1-s)f_+(t_2) \leq sf(t_1+\varepsilon)+(1-s)f(t_2+\varepsilon) \tag{14}$$

bármely kicsi, rögzített ε-ra.

Feltettük, hogy f $2n$-monoton, tehát a (13) egyenlőtlenség fennáll 1 dimenziós esetben is. Emiatt tovább becsülhetjük a (14) egyenlőtlenséget a következőképpen.

Legyen $x_1 = t_1+\varepsilon$, $x_2 = t_2+\varepsilon$. x_1 és x_2-re alkalmazzuk a (13) egyenlőtlenséget

$$sf(t_1+\varepsilon)+(1-s)f(t_2+\varepsilon) \leq f(s(t_1+\varepsilon)+(1-s)(t_2+\varepsilon)+\varepsilon) = f(st_1+(1-s)t_2+2\varepsilon). \tag{15}$$

Összefoglalva a (14) és (15) egyenlőtlenségeket, eddig azt láttuk be, hogy

$$sf_+(t_1) + (1-s)f_+(t_2) \leq f(st_1 + (1-s)t_2 + 2\varepsilon). \tag{16}$$

Tartsunk (16)-ban ε-nal a 0-hoz. Ekkor definíció szerint

$$sf_+(t_1) + (1-s)f_+(t_2) \leq f_+(st_1 + (1-s)t_2), \tag{17}$$

azaz f_+ konkáv, és ebből következik, hogy f_+ folytonos.
Legyen $\delta \in \mathbb{R}$, $\delta > 0$ kicsi. Mivel f monoton növő, ezért

$$f_+(t-\delta) \leq f(t) \leq f_+(t),$$

ha $0 < \delta < t$. Mivel f_+ folytonos, ezért ha $\delta \to 0$, akkor $f(t) = f_+(t)$ minden $t > 0$-ra. Ezzel megkaptuk, hogy f is folytonos, tehát (13) fennáll $\varepsilon \to 0$ esetben is, így

$$f(sX_1 + (1-s)X_2) \geq sf(X_1) + (1-s)f(X_2), \tag{18}$$

azaz f n-konkáv.

2.4. Következmény. *Ha $f(X)$ operátor-monoton növő függvény (a, ∞)-en, akkor f operátor-konkáv (a, ∞)-en.*

Bizonyítás: Tegyük fel hogy f n-monoton növő minden $2n \in \mathbb{N}$ esetén. Ekkor f n-konkáv minden $n \in \mathbb{N}$ esetén az előbbi tétel alapján.

2.5. Tétel. *Ha $f(X)$ nemnegatív függvény, és n-konkáv (a, ∞)-en $(n \in \mathbb{N})$, akkor f n-monoton növő (a, ∞)-en.*

Bizonyítás: Feltehető, hogy $a = 0$. Legyenek $X, Y \in M_n(\mathbb{C})$ pozitív definit mátrixok és $X < Y$, továbbá $\lambda \in (0,1)$ szám. Ekkor

$$\lambda Y = \lambda X + \lambda(Y - X) = \lambda X + (1-\lambda)(\lambda \tfrac{1}{1-\lambda}(Y - X)).$$

Mivel f n-konkáv, ezért definícióból következik, hogy

$$f(\lambda Y) \geq \lambda f(X) + (1-\lambda) f(\lambda \tfrac{1}{1-\lambda}(Y - X)),$$

továbbá, ha fennáll, hogy f nemnegatív, akkor az előbbi állítást bővíthetjük az egyik tag elhagyásával (ugyanis nincs negatív skalárral megszorozva):

$$f(\lambda Y) \geq \lambda f(X) + (1-\lambda) f(\lambda \tfrac{1}{1-\lambda}(Y - X)) \geq \lambda f(X),$$

azaz

$$f(\lambda Y) \geq \lambda f(X).$$

Ha $\lambda \to 1$, akkor megkapjuk, hogy $f(X) \leq f(Y)$, ugyanis f folytonos (mert konkáv).

Bővebben, ha csak azt tudjuk, hogy $X \leq Y$, akkor vegyünk egy $\mu \in (0,1)$ számot, így μX pozitív definit marad. Innen $\mu X < X \leq Y$ miatt

$$f(\mu X) < f(Y)\ .$$

Tehát $\mu \to 1$ esetén megkapjuk a monotonitás feltételét: $f(X) \leq f(Y)$.

2.6. Következmény. *Ha $f(X)$ operátor-konkáv $(0,\infty)$-en és nemnegatív, akkor f operátor-monoton növő $(0,\infty)$-en.*

2.7. Következmény. *Ha $f(X)$ nemnegatív, akkor f operátor-konkáv $(0,\infty)$-en akkor és csak akkor ha operátor-monoton növő $(0,\infty)$-en.*

Bizonyítás: a 2.4 és a 2.6 következmények egyesítése.

2.8. Állítás. *Ha egy $f(X)$ függvény operátor-monoton növő $(0,\infty)$-en, és nemnegatív, folytonos függvény, akkor $g(X) = \frac{1}{f(X)}$ operátor-konvex $(0,\infty)$-en.*

Bizonyítás:

Legyenek $A, B \in M_n(\mathbb{C})$, $A, B > 0$. Tudjuk, hogy f operátor-monoton növő, ezért a 2.7 következmény miatt operátor-konkáv, így $\lambda = \frac{1}{2}$-re azt kapjuk, hogy

$$\frac{f(A) + f(B)}{2} \leq f\left(\frac{A+B}{2}\right). \tag{19}$$

Ezek sajátértékei még mindig $(0,\infty)$-ben vannak, így felhasználjuk, hogy az inverzképzés operátor-monoton csökkenő 1.32 szerint, ezért

$$\left(\frac{f(A) + f(B)}{2}\right)^{-1} \geq \left(f\left(\frac{A+B}{2}\right)\right)^{-1}. \tag{20}$$

Fel fogjuk használni, hogy az inverzképzés konvex $(0,\infty)$-n. Tudjuk, hogy $f(A)$ és $f(B)$ sajátértékei szintén $(0,\infty)$-ben vannak, ezért $\lambda = \frac{1}{2}$ esetén (ezt bizonyítottuk 1.33 -ben) ez azt jelenti, hogy

$$\frac{(f(A))^{-1} + (f(B))^{-1}}{2} \geq \left(\frac{f(A) + f(B)}{2}\right)^{-1}. \tag{21}$$

Ha összeillesztjük az (20) és (21) egyenlőtlenségeket, akkor pont azt kapjuk, amit szeretnénk belátni, azaz

$$\frac{(f(A))^{-1} + (f(B))^{-1}}{2} \geq \left(f\left(\frac{A+B}{2}\right)\right)^{-1},$$

ami ekvivalens azzal, hogy

$$\frac{g(A) + g(B)}{2} \geq g\left(\frac{A+B}{2}\right).$$

Ezzel beláttuk, hogy $\lambda = \frac{1}{2}$ esetén teljesül a konvexitás, így felhasználva, hogy f folytonos, ebből következik, hogy g operátor-konvex $(0,\infty)$-n.

3. A kapcsolat általánosítása (Löwner- és Krauss-mátrixok)

Eddig a pozitív félegyenesen értelmezett függvényekkel dolgoztunk, most már általánosabb állításokról lesz szó. Ezen felül olyan állítások hangzanak el, melyek segítségével jó esetben egy konkrétan megadott függvényről meg tudjuk állapítani a monotonitását, illetve a konvexitását. Itt leginkább Rajendra Bhatia [1] könyve alapjan közelítettük meg a Löwner- és a Krauss mátrixokat.

3.1. Osztott differenciák

3.1. Definíció. *$f(x)$ elsőrendű osztott differenciája az $x_1 \neq x_2$ helyeken*

$$[x_1, x_2]_f = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Ha $x_1 = x_2$, akkor $[x_1, x_1]_f = f'(x_1)$.

3.2. Definíció. *$f(x)$ másodrendű osztott differenciája, ha $x_1 \neq x_3$,*

$$[x_1, x_2, x_3]_f = \frac{[x_2, x_3]_f - [x_1, x_2]_f}{x_3 - x_1}.$$

Ha $x_1 = x_3$, akkor $[x_1, x_2, x_1]_f = \frac{f'(x_1) - [x_2, x_1]_f}{x_1 - x_2}$.
Ha $x_1 = x_2 = x_3$, akkor $[x_1, x_1, x_1]_f = \frac{f''(x_1)}{2}$.

3.3. Állítás. *Az osztott differenciák lineárisak, azaz:*

$$[x_1, \ldots, x_n]_{f+g} = [x_1, \ldots, x_n]_f + [x_1, \ldots, x_n]_g,$$

$$[x_1, \ldots, x_n]_{\lambda f} = \lambda [x_1, \ldots, x_n]_f.$$

3.4. Állítás. *A másodrendű osztott differenciák szimmetrikusak, azaz*

$$[a, b, c]_f = [b, a, c]_f = [c, b, a]_f.$$

Megjegyzés: Az osztott differenciák mind szimmetrikusak, elsőrendű esetben ez triviális. Mivel a másodrendű esetet fogjuk használni, egyszerűbb bizonyítást alkalmazhatunk.

Bizonyítás:

$$[a, b, c]_f = \frac{\frac{f(a)-f(b)}{a-b} - \frac{f(b)-f(c)}{b-c}}{a - c} = \frac{(b-c)(f(a) - f(b)) - (a-b)(f(b) - f(c))}{(a-c)(a-b)(b-c)} =$$

$$= \frac{b(f(a)-f(c))+c(f(b)-f(a))+a(f(c)-f(b))}{(a-c)(a-b)(b-c)}.$$

Az utolsó alakban vegyük észre, hogy bármely két bemenetet megcserélve a számláló és a nevező is előjelet vált, így ugyanezt a kifejezést adja.

3.5. Lemma. *Legyen $\lambda_i, \lambda_j \in \mathbb{R}$, és $f(x) = x^n$. Ekkor*

$$[\lambda_i, \lambda_j]_f = \sum_{s=1}^{n} \lambda_i^{s-1}\lambda_j^{n-s}.$$

Bizonyítás:

$$[\lambda_i, \lambda_j]_f = \frac{f(\lambda_i)-f(\lambda_j)}{\lambda_i-\lambda_j} = \frac{\lambda_i^n-\lambda_j^n}{\lambda_i-\lambda_j} = \lambda_i^{n-1}+\lambda_i^{n-2}\lambda_j+\ldots+\lambda_i\lambda_j^{n-2}+\lambda_j^{n-1} = \sum_{s=1}^{n} \lambda_i^{s-1}\lambda_j^{n-s}.$$

Megjegyzés: Az összeg szimmetrikus, így felcserélhető λ_i és λ_j szerepe. Ezt az osztott differenciák szimmetriájával is lehet indokolni.

3.6. Lemma. *Legyen $\lambda_i, \lambda_j, \lambda_k \in \mathbb{R}$, és $f(x) = x^n$. Ekkor*

$$[\lambda_i, \lambda_j, \lambda_k]_f = \sum_{s=1}^{n}\sum_{r=1}^{n-s} \lambda_i^{r-1}\lambda_j^{s-1}\lambda_k^{n-r-s}.$$

Bizonyítás:

$$[\lambda_i, \lambda_j, \lambda_k]_f = \frac{[\lambda_j, \lambda_k]_f - [\lambda_i, \lambda_j]_f}{\lambda_k-\lambda_i} = \frac{\sum\limits_{s=1}^{n} \lambda_j^{s-1}\lambda_k^{n-s} - \sum\limits_{s=1}^{n} \lambda_i^{s-1}\lambda_j^{n-s}}{\lambda_k-\lambda_i} =$$

$$= \frac{\sum\limits_{s=1}^{n} \lambda_j^{s-1}\lambda_k^{n-s} - \sum\limits_{s=1}^{n} \lambda_i^{n-s}\lambda_j^{s-1}}{\lambda_k-\lambda_i} = \frac{\sum\limits_{s=1}^{n} \lambda_j^{s-1}(\lambda_k^{n-s}-\lambda_i^{n-s})}{\lambda_k-\lambda_i} =$$

$$= \sum_{s=1}^{n} \lambda_j^{s-1}\frac{\lambda_k^{n-s}-\lambda_i^{n-s}}{\lambda_k-\lambda_i} = \sum_{s=1}^{n} \lambda_j^{s-1}\sum_{r=1}^{n-s}\lambda_i^{r-1}\lambda_k^{(n-s)-r} = \sum_{s=1}^{n}\sum_{r=1}^{n-s}\lambda_j^{s-1}\lambda_i^{r-1}\lambda_k^{n-r-s}.$$

Az alakítások során kétszer felhasználjuk a 3.5 lemmát, egyszer a bizonyítás elején, másodszor pedig ott, hogy

$$\frac{\lambda_k^{n-s}-\lambda_i^{n-s}}{\lambda_k-\lambda_i} = [\lambda_k, \lambda_i]_g,$$

ahol $g(x) = x^{n-s}$.

3.2. A Fréchet-derivált

A klasszikus analízisben sok információt nyerünk valós függvényekről a deriválás segítségével. Emiatt érdemes kiterjeszteni ezt a fogalmat a mátrixok körébe.

3.7. Definíció. *Egy $f : M_n(\mathbb{C})^{sa} \to M_n(\mathbb{C})^{sa}$ önadjungált mátrixokon értelmezett függvény Fréchet-deriválható (a továbbiakban: deriválható) egy $A \in M_n(\mathbb{C})$-ben, ha létezik olyan $Df(A) : M_n(\mathbb{C})^{sa} \to M_n(\mathbb{C})^{sa}$ szintén önadjungáltakon értelmezett lineáris leképezés, hogy minden önadjungált $H \in M_n(\mathbb{C})$-re*

$$\lim_{\|H\|\to 0} \frac{\|f(A+H) - f(A) - Df(A)(H)\|}{\|H\|} = 0.$$

Megjegyzés: Ha f valóban deriválható A-ban, akkor

$$Df(A)(H) = \frac{d}{dt} f(A+tH)|_{t=0},$$

ami a H irány menti deriváltja f-nek A-ban.

Megjegyzés: Észrevehető, hogy a klasszikus $f : \mathbb{R} \to \mathbb{R}$ függvényeknél

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

átírható a következő alakba:

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

így a skalár esetben f Fréchet-deriváltja a-ban $f'(a)$.

Erre a deriválásra fennállnak a szokásos deriváltra érvényes kapcsolatok, sőt az egyértelműség is. Ezeket most nem fogjuk belátni, de kimondjuk a jobb átláthatóság végett.

3.8. Állítás. *legyenek $f, g : M_n(\mathbb{C})^{sa} \to M_n(\mathbb{C})^{sa}$ deriválható függvények. Ekkor a következő deriválási szabályok teljesülnek.*

Összegfüggvény deriváltja:

$$D(f+g)(A) = Df(A) + Dg(A).$$

Szorzatfüggvény deriváltja (Leibniz-szabály):

$$D(fg)(A) = (Df(A))g(A) + f(A)(Dg(A)).$$

Megjegyzés: A mátrixokon vett szorzás nem kommutatív, ezért oda kell figyelni amikor alkalmazzuk a deriválási szabályokat, leginkább például a Leibniz szabályra.

3.9. Állítás. *Legyenek A , $B \in M_n(\mathbb{C})$ önadjungáltak. Ekkor a következő néhány elemi függvény deriváltja:*

Legyen $f(A) = A^2$, ekkor

$$Df(A)(B) = AB + BA. \tag{22}$$

Legyen $f(A) = A^n$, ekkor

$$Df(A)(B) = \sum_{j+k=n-1} A^j B A^k. \tag{23}$$

Legyen $f(A) = A^{-1}$, ekkor

$$Df(A)(B) = -A^{-1} B A^{-1}. \tag{24}$$

3.10. Definíció. *Legyen az $f : M_n(\mathbb{C})^{sa} \to M_n(\mathbb{C})^{sa}$ önadjungált mátrixokon értelmezett függvény Fréchet-deriválható egy $X \subset M_n(\mathbb{C})^{sa}$ nyílt halmazon, ekkor adott a $Df : X \to L(M_n(\mathbb{C})^{sa}, M_n(\mathbb{C})^{sa})$ lineáris leképezés, ahol $L(Y, Z)$ jelöli az $Y \to Z$ lineáris leképezések halmazát. Ekkor f kétszer deriválható egy $A \in X \subset M_n(\mathbb{C})$ helyen, ha létezik olyan $D^2 f : X \to L(M_n(\mathbb{C})^{sa}, L(M_n(\mathbb{C}), M_n(\mathbb{C})^{sa}))$ bilineáris leképezés, hogy minden önadjungált $H_1, H_2 \in M_n(\mathbb{C})$-re*

$$\lim_{\|H_2\| \to 0} \frac{\|Df(A + H_2)(H_1) - Df(A)(H_1) - D^2 f(A)(H_1, H_2)\|}{\|H_2\|} = 0.$$

Jelölés: Ha $H_1 = H_2 = H$, akkor $D^2 f(A)(H_1, H_2) = D^2 f(A)(H)$.

Megjegyzés: Ha f valóban deriválható A-ban, akkor

$$D^2 f(A)(H) = \frac{d^2}{dt^2} f(A + tH)|_{t=0},$$

ami a H irány menti második deriváltja f-nek A-ban.

3.11. Állítás. *Legyenek A , $B \in M_n(\mathbb{C})$ önadjungáltak. Ekkor az $f(A) = A^n$ függvény másodrendű deriváltja:*

$$D^2 f(A)(B) = 2 \sum_{i+j+k=n-2} A^i B A^j B A^k.$$

A következő állítások kulcsfontosságúak a továbbiak szempontjából. Itt mondjuk ki a deriváltak és a monotonitás illetve konvexitás közti kapcsolatot, mely a klasszikus analízisben ismert állításoknak a kiterjesztése.

3.12. Állítás. *Egy f függvény n-monoton növő (ill. csökkenő) egy I nyílt intervallumon akkor, és csak akkor, ha $\frac{d}{dt}f(X+tH)|_{t=0}$ pozitív (ill. negatív) szemidefinit minden önadjungált X mátrixra, melynek sajátértékei I-ben vannak, és minden $H \geq 0$ önadjungált mátrixra.*

Bizonyítás:
($\Rightarrow$) Tegyük fel, hogy f egy n-monoton növő függvény I-n. Tudjuk, hogy

$$\frac{d}{dt}f(X+tH)|_{t=0} = \lim_{t\to 0}\frac{f(X+tH)-f(X)}{t}.$$

Legyen X tetszőleges. Mivel H pozitív szemidefinit, így

$$X + tH \geq X,$$

tehát feltevés szerint f monotonitása miatt

$$f(X+tH) \geq f(X),$$

ezért ezt bal oldalra rendezve és t-vel osztva pozitív szemidefinit mátrixot kapunk, és $t \to 0$ esetén ez továbbra is pozitív szemidefinit lesz.

($\Leftarrow$) Tegyük fel, hogy $\exists A \leq B$, hogy $f(A)$ nem kisebb-egyenlő, mint $f(B)$. Ezt úgy tudjuk megfogni matematikailag, hogy $\exists \zeta \in \mathbb{C}^n$, melyre

$$\langle f(A)\zeta, \zeta\rangle > \langle f(B)\zeta, \zeta\rangle.$$

Azt kell belátni, hogy van olyan X és $0 \leq H$, hogy

$$\langle f(X+tH)\zeta, \zeta\rangle'|_{t=0} < 0,$$

azaz $\frac{d}{dt}f(X+tH)|_{t=0}$ nem pozitív szemidefinit.
Legyen $H = B - A \geq 0$. Most tegyük fel, hogy $A < B$, és legyen $g : \mathbb{R} \to \mathbb{R}$,

$$g(t) = \langle f(A+tH)\zeta, \zeta\rangle .$$

Alkalmazhatjuk a Lagrange-féle középértéktételt bármely rögzített ζ mellett, ezért $\exists 0 \leq t_0 \leq 1$, hogy

$$g'(t_0) = \frac{g(1) - g(0)}{1 - 0},$$

azaz

$$\langle f(A + tH)\zeta, \zeta\rangle'|_{t=t_0} = \langle f(B)\zeta, \zeta\rangle - \langle f(A)\zeta, \zeta\rangle.$$

Feltevés szerint azt is tudjuk, hogy $\exists\zeta$, hogy $\langle f(B) - f(A)\zeta, \zeta\rangle < 0$. Erre a ζ-ra teljesül:

$$\frac{d}{dt}\langle f(A + tH)\zeta, \zeta\rangle|_{t=t_0} < 0.$$

Tehát válasszunk úgy, hogy $X = A + t_0 H$, mert ekkor a deriváltra fennáll

$$\frac{d}{dt}\langle f(X + tH)\zeta, \zeta\rangle|_{t=0} < 0,$$

és ezt szerettük volna belátni.

3.13. Példa. *Térjünk vissza az 1.28 példában vett $f(X) = X^2$ függvényre. Ekkor $Df(A)(B) = AB + BA$. Ebben a példában lévő $A, B \in M_2(\mathbb{C})$ mátrixokra $det(AB + BA) = -1$ teljesült, tehát f nem 2-monoton növő $[0, \infty)$-n.*

3.14. Állítás. *Egy $f \in C^2(I)$ függvény n-konvex (konkáv) egy I nyílt intervallumon akkor, és csak akkor, ha $\frac{d^2}{dt^2} f(X + tH)|_{t=0}$ pozitív (negatív) szemidefinit minden X önadjungált mátrixra, melynek sajátértékei I-ben vannak, és minden $H \geq 0$ önadjungált mátrixra.*

Megjegyzés: Ez ekvivalens azzal, hogy f n-konvex akkor, és csak akkor, ha f' n-monoton növő.

Bizonyítás:

Ezt az előző bizonyításhoz képest máshogy vizsgáljuk meg. Definíció szerint f n-konvex I-n ha minden $A, B \in M_n(\mathbb{C})$ önadjungált mátrixra, melyek sajátértékei I-ben vannak, és minden $0 \leq \lambda \leq 1$-ra

$$f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda) f(B),$$

Azaz minden $\zeta \in \mathbb{C}^n$ -re

$$\langle f(\lambda A + (1 - \lambda)B)\zeta, \zeta\rangle \leq \lambda\langle f(A)\zeta, \zeta\rangle + (1 - \lambda)\langle f(B)\zeta, \zeta\rangle.$$

Rögzítsük le ζ-t, és legyen

$$K = \{A \in M_n(\mathbb{C}) \text{ önadjungált } | A \text{ sajátértékei } I\text{-ben vannak}\}.$$

Ekkor K egy konvex halmaz, ugyanis $A, B \in K \Rightarrow \lambda A + (1-\lambda)B \in K$ $(0 \leq \lambda \leq 1)$.

Tekintsük azt a $g : K \to \mathbb{R}$ lineáris funkcionált, hogy

$$g(A) = \langle f(A)\zeta, \zeta \rangle.$$

Vegyük észre, hogy f n-konvex akkor és csak akkor, ha g konvex.

Azt is tudjuk valós analízisből, hogy ekvivalensek:

(i) g konvex.

(ii) g minden egyenes K-t metsző részén konvex.

Egy egyenest viszont pont egy $X + tH$ alakkal jellemezhetünk, és ha egy X pontban kicsit eltérünk H irányban, mivel g konvex, ezért itt a második derivált nemnegatív.

Visszafele ugyanez a bizonyítás igaz, ugyanis a felhasznált állítások ekvivalenciát mondtak ki a lépések között.

3.3. Löwner-mátrixok

Az ebben a részben definiált Löwner-mátrixok szorosan kapcsolódnak a Fréchet-deriváthoz, és mégis máshogyan megfogható eredményeket kapunk ezek segítségével.

3.15. Definíció. *Legyen $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Az f függvény és $\lambda_1, \dots, \lambda_n$ által meghatározott Löwner-mátrix:*

$$L_f(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} [\lambda_1, \lambda_1]_f & \dots & [\lambda_1, \lambda_n]_f \\ \vdots & \ddots & \vdots \\ [\lambda_n, \lambda_1]_f & \dots & [\lambda_n, \lambda_n]_f \end{pmatrix}.$$

Jelölés: Legyen $\Lambda = diag(\lambda_1, \dots, \lambda_n)$, ekkor $L_f(\Lambda) = L_f(\lambda_1, \dots, \lambda_n)$.

Jelölés: Legyen A önadjungált, így $A = U^*\Lambda U$. Ekkor $L_f(A) = U^* L_f(\Lambda)U$.

A következő állításban megmutatjuk a Löwner-mátrixok és a Fréchet-derivált közti kapcsolatot. Ennek következtében eljutunk egy olyan eszközhöz, aminek a segítségével el tudjuk dönteni, hogy mikor mátrix-monoton egy függvény. Ezt az állítást későbbi állítások bizonyításában is fel fogjuk használni.

3.16. Állítás. *Legyen $f \in C^1(I)$, azaz egyszer folytonosan differenciálható I-n, ahol I nyílt intervallum. Legyen $\Lambda = diag(\lambda_1, \dots, \lambda_n)$, ahol $\lambda_i \in I$, és H önadjungált. Jelölje $\circ$ a elemenkénti szorzást mátrixoknál. Ekkor*

$$Df(\Lambda)(H) = L_f(\Lambda) \circ H. \tag{25}$$

Bizonyítás:

A valós együtthatós polinomfüggvények körében fogjuk bizonyítani, majd onnan approximációval következik az állítás $C^1(I)$-re. Linearitás miatt ezért elég most $f(X) = X^n$ ($n \in \mathbb{N}^+$) alakú függvényekre bizonyítani.

Felhasználjuk, hogy tetszőleges önadjungált A-ra $f(A) = A^n$ esetén $Df(A)(B) = \sum_{j+k=n-1} A^j B A^k$, így

$$(Df(\Lambda)(H))_{(i,j)} = \left(\sum_{s=1}^{n} \Lambda^{s-1} H \Lambda^{n-s} \right)_{(i,j)} = \sum_{s=1}^{n} \lambda_i^{s-1} h_{ij} \lambda_j^{n-s}.$$

A (25) egyenlet jobb oldalának (i,j)-ik eleme pedig a 3.5 lemma miatt:

$$(L_f(\Lambda) \circ H)_{(i,j)} = [\lambda_i, \lambda_j]_f h_{ij} = \sum_{s=1}^{n} \lambda_i^{s-1} \lambda_j^{n-s} h_{ij}.$$

A következő állításban ezt kiterjesztjük az önadjungált mátrixok körébe, hogy bővebb környezetben alkalmazható legyen.

3.17. Állítás. *Legyen $f \in C^1(I)$, azaz egyszer folytonosan differenciálható I-n, ahol I nyílt intervallum. Legyen A önadjungált, és $A = U^* \Lambda U$, $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$, ahol $\lambda_i \in I$, és H önadjungált. Ekkor*

$$Df(A)(H) = L_f(A) \circ H.$$

Bizonyítás:

Továbbra is a polinomok körében bizonyítjuk az állítást, ebből szintén approximációval következik, hogy $C^1(I)$-n is teljesül.

Azt akarjuk belátni, hogy

$$Df(A)(H) = U^*(L_f(\Lambda) \circ UHU^*)U.$$

Kezdjük a bal oldal vizsgálatával. Felhasználjuk, hogy $f(A) = A^n$ esetén $Df(A)(B) = \sum_{j+k=n-1} A^j B A^k$, így

$$Df(A)(H) = \sum_{s=1}^{n} A^{s-1} H A^{n-s} = \sum_{s=1}^{n} U^* \Lambda^{s-1} U H U^* \Lambda^{n-s} U = U^* \left(\sum_{s=1}^{n} \Lambda^{s-1} U H U^* \Lambda^{n-s} \right) U.$$

Mivel U unitér, így nem lehet szinguláris, ezért elég belátni, hogy

$$\sum_{s=1}^{n} \Lambda^{s-1} UHU^* \Lambda^{n-s} = L_f(\Lambda) \circ UHU^*. \tag{26}$$

Írjuk át a (26) jobb oldalát a következő alakba:

$$(L_f(\Lambda) \circ UHU^*)_{(i,j)} = [\lambda_i, \lambda_j]_f \, (UHU^*)_{(i,j)} = \sum_{s=1}^{n} \lambda_i^{s-1} \lambda_j^{n-s} \, (UHU^*)_{(i,j)} \, .$$

Most vizsgáljuk (26)-ban a bal oldal (i, j)-ik elemét, mivel Λ diagonális, ezért könnyen megkapjuk, hogy

$$\sum_{s=1}^{n} (\Lambda^{s-1} UHU^* \Lambda^{n-s})_{(i,j)} = \sum_{s=1}^{n} \lambda_i^{s-1} (UHU^*)_{(i,j)} \lambda_j^{n-s},$$

Ezzel beláttuk, hogy (26) teljesül, így az eredeti állítást is.

3.18. Következmény. *Legyen $\phi : [a, b] \to \{A \in M_n(\mathbb{C})$ önadjungált $|A$ sajátértékei I-ben vannak $\}$. Tegyük fel, hogy $\phi \in C^1([a, b])$. Legyen $f \in C^1(I)$ és $F(t) = f(\phi(t))$. Ekkor*

$$F(b) - F(a) = \int_a^b L_f(\phi(t)) \circ \phi'(t) dt.$$

Bizonyítás:

A Newton-Leibniz formulából adódik megfelelő feltételek mellett, hogy

$$F(b) - F(a) = \int_a^b F'(t) dt,$$

ugyanis feltettük f-ről és ϕ-ről, hogy folytonosak.

Erre az esetre levezetjük a láncszabályt, mivel annak értelmezése nem triviális. Ezért azt állítjuk, hogy

$$(f(\phi(t)))' = Df(\phi(t))(\phi'(t)). \tag{27}$$

Ekkor kicsi h elmozdulással $[a, b]$-ben

$$f(\phi(t + h)) = f(\phi(t) + \phi'(t)h + o(h)).$$

Mivel f folytonos, ezért az $o(h)$ tag kivihető f-ből, és továbbra is $o(h)$ lesz. Ezért ez tovább egyenlő azzal, hogy

$$f(\phi(t+h)) = f(\phi(t)) + Df(\phi(t))(\phi'(t)h) + o(h).$$

Mivel $Df(\phi(t))$ lineáris, így kivihetjük h-t, mivel nem függ tőle, és azt kapjuk, hogy

$$f(\phi(t+h)) = f(\phi(t)) + Df(\phi(t))(\phi'(t))h + o(h).$$

Ez definíció szerint azt jelenti, hogy $Df(\phi(t))(\phi'(t))$ maga az $f(\phi(t))$ deriváltja. Tehát alkalmazhatjuk a láncszabályt, és azt kapjuk, hogy a kompozíció deriváltja

$$F'(t) = (f(\phi(t)))' = Df(\phi(t))(\phi'(t)). \tag{28}$$

Most alkalmazzuk a (3.17) állítást ($A = \phi(t), H = \phi'(t)$), így megkapjuk azt, hogy

$$Df(\phi(t))(\phi'(t)) = L_f(\phi(t)) \circ \phi'(t). \tag{29}$$

Tehát beláttuk két nagyobb lépésben, hogy

$$F(b) - F(a) = \int_a^b Df(\phi(t)) \circ (\phi'(t)) = \int_a^b L_f(\phi(t)) \circ \phi'(t)dt.$$

3.19. Tétel. *Legyen $f \in C^1(I)$, ahol I nyílt intervallum. Ekkor f operátor-monoton növő (ill. csökkenő) akkor és csak akkor, ha $L_f(A)$ pozitív (ill. negatív) szemidefinit minden olyan $A \in M_n(\mathbb{C})$ önadjungált mátrixra, melynek sajátértékei I-ben vannak.*

Bizonyítás:

($\Rightarrow$) Tegyük fel, hogy f operátor-monoton növő, és A sajátértékei I-ben vannak. Legyen

$$H = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix},$$

ekkor $H \geq 0$ ugyanis

$$\langle Hx, x \rangle = \left\langle \left(\sum_i x_i, \dots, \sum_i x_i \right), x \right\rangle = \left(\sum_i x_i \right)^2 \geq 0.$$

Ezért bármilyen $t \in \mathbb{R}^+$-ra $A + tH \geq A$.
Felhasználjuk, hogy f operátor-monoton növő, így

$$f(A + tH) \geq f(A),$$

tehát $f(A + tH) - f(A) \geq 0$.
Vegyük $t \to 0$-ban a határértéket, ebből azt kapjuk, hogy t-vel osztva ($t \geq 0$)

$$Df(A)(H) = \frac{d}{dt} f(A + tH)|_{t=0} \geq 0,$$

ugyanis feltettük, hogy f egyszer folytonosan differenciálható.
Ha alkalmazzuk a 3.17 állítást, akkor azt kapjuk, hogy

$$Df(A)(H) = L_f(A) \circ H \geq 0,$$

viszont ebben az esetben H választása miatt ez ekvivalens azzal, hogy $L_f(A) \geq 0$.
($\Leftarrow$) Legyenek A, B sajátértékei I-ben, és $A \leq B$. Legyen

$$\phi(t) = (1 - t)A + tB,$$

ahol $0 \leq t \leq 1$, ekkor $\phi(t)$ sajátértékei szintén I-ben vannak, emiatt feltevés szerint minden t-re

$$L_f(\phi(t)) \geq 0. \tag{30}$$

Vegyük észre, hogy $\phi(t)$ t-szerinti deriváltja minden t-re

$$\phi'(t) = B - A \geq 0. \tag{31}$$

Felhasználjuk, hogy két pozitív szemidefinit mátrix pontonkénti szorzata szintén pozitív szemidefinit. Így (30) és (31) kompozíciójáról megkapjuk, hogy

$$L_f(\phi(t)) \circ \phi'(t) \geq 0.$$

Mivel ez fennáll minden $t \in [0, 1]$-re, ezért a 3.18 következmény miatt

$$f(B) - f(A) = F(\phi(1)) - F(\phi(0)) = \int_0^1 L_f(\phi(t)) \circ \phi'(t) dt \geq 0,$$

ezért f operátor-monoton növő.

3.20. Példa. *Az $f(t) = e^t$ függvény nem operátor-monoton növő $\mathbb{R}$-en.*

Bizonyítás: ellenpéldával belátjuk, hogy f nem 2-monoton növő.

Legyen $\lambda_1 = 1, \lambda_2 = 0$, ekkor

$$L_f(\lambda_1, \lambda_2) = \begin{pmatrix} e & \frac{e-1}{1-0} \\ \frac{e-1}{1-0} & 1 \end{pmatrix},$$

és $det(L_f(\lambda_1, \lambda_2)) = -e^2 + 3e + 1 \approx -0.234 < 0$.

Ugyan ez csak azt jelenti, hogy az $[0, 1]$ intervallumon f nem 2-monoton növő, de a példában csak arra szerettünk volna rámutatni, hogy f 1-monoton növő függvény, de itt se következik ebből, hogy 2- sőt n-monoton növő.

3.4. Krauss-mátrixok

A Löwner-mátrixokhoz hasonlóan most a konvexitás, illetve a másodrendű deriváltak között mutatjuk meg a kapcsolatot.

3.21. Definíció. *Legyen $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Az f függvény és $\lambda_1, \ldots, \lambda_n$ által meghatározott Krauss-mátrixai $p = 1, \ldots, n$-re:*

$$K_{p,f}(\lambda_1, \ldots, \lambda_n) = 2 \begin{pmatrix} [\lambda_p, \lambda_1, \lambda_1]_f & \ldots & [\lambda_p, \lambda_1, \lambda_n]_f \\ \vdots & \ddots & \vdots \\ [\lambda_p, \lambda_n, \lambda_1]_f & \ldots & [\lambda_p, \lambda_n, \lambda_n]_f \end{pmatrix}.$$

Jelölés: Legyen $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$, ekkor $K_{p,f}(\Lambda) = K_{p,f}(\lambda_1, \ldots, \lambda_n)$.

Jelölés: Legyen A önadjungált, így $A = U^*\Lambda U$. Ekkor $K_{p,f}(A) = U^* K_{p,f}(\Lambda)U$.

3.22. Állítás. *Legyen $f \in C^2(I)$, azaz kétszer folytonosan differenciálható I-n, ahol I nyílt intervallum. Legyen $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$, ahol $\lambda_i \in I$, és H önadjungált. Továbbá legyen $H_p \in M_n(\mathbb{C})$:*

$$(H_p)_{(i,j)} = (diag(H_{(p,1)}, \ldots, H_{(p,n)}))_{(i,j)} = \begin{cases} H_{(p,i)} & ha\ i = j \\ 0 & különben, \end{cases}$$

tehát ez H p-edik sorát tartalmazza a főátlójában.

Ekkor az teljesül, hogy

$$D^2 f(\Lambda)(H) = 2 \sum_{p=1}^{n} H_p^* K_{p,f}(\Lambda) H_p. \tag{32}$$

Bizonyítás:

A valós együtthatós polinomfüggvények körében fogjuk bizonyítani, majd onnan approximációval következik az állítás $C^2(I)$-re. Linearitás miatt ezért elég most $f(X) = X^n$ ($n \in \mathbb{N}^+$) alakú függvényekre bizonyítani.

Kezdjük a jobb oldal felbontásával:

$$(2(H_p^* K_{p,f}(\Lambda) H_p)_{(i,j)} =$$

$$= \left(2 \begin{pmatrix} \overline{h}_{(p,1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \overline{h}_{(p,n)} \end{pmatrix} \begin{pmatrix} [\lambda_p, \lambda_1, \lambda_1]_f & \cdots & [\lambda_p, \lambda_1, \lambda_n]_f \\ \vdots & \ddots & \vdots \\ [\lambda_p, \lambda_n, \lambda_1]_f & \cdots & [\lambda_p, \lambda_n, \lambda_n]_f \end{pmatrix} \begin{pmatrix} h_{(p,1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_{(p,n)} \end{pmatrix} \right)_{(i,j)} =$$

(Ahol $h_{(i,j)}$ jelöli H-nak az (i,j)-ik elemét)

$$= 2\overline{h}_{(p,i)} [\lambda_p, \lambda_i, \lambda_j]_f h_{(p,j)} =$$

(A 3.6 lemma szerint behelyettesítünk az osztott differenciákba, és azt is felhasználjuk, hogy H önadjungált.)

$$= 2h_{(i,p)} \left(\sum_{s=1}^{n} \sum_{r=1}^{n-s} \lambda_p^{r-1} \lambda_i^{s-1} \lambda_j^{n-r-s} \right) h_{(p,j)}.$$

A másik oldalról viszont felhasználjuk, hogy Λ diagonális, így

$$(D^2 f(\Lambda)(H))_{(i,j)} = \left(2 \sum_{q+r+s=n-2} \Lambda^q H \Lambda^r H \Lambda^s \right)_{(i,j)} =$$

$$2 \sum_{q+r+s=n-2} \left(\sum_{p=1}^{n} \lambda_i^q h_{(i,p)} \lambda_p^r h_{(k,j)} \lambda_j^s \right) = 2 \sum_{p=1}^{n} \left(\sum_{q+r+s=n-2} \lambda_i^q h_{(i,p)} \lambda_k^r h_{(p,j)} \lambda_j^s \right) =$$

$$= 2 \sum_{p=1}^{n} h_{(i,p)} \left(\sum_{q+r+s=n-2} \lambda_i^q \lambda_p^r \lambda_j^s \right) h_{(p,j)},$$

tehát a jobb oldalon lévő mátrixok összege (32)-ben tényleg a bal oldali mátrixot adja.

3.23. Tétel. *Legyen $f \in C^2(I)$, ahol I nyílt intervallum. Ekkor f n-konvex I-n akkor, és csak akkor, ha minden $p \in \{1, \dots, n\}$ -re és minden $\lambda_1, \dots, \lambda_n \in I$ esetén*

$$K_{p,f}(\lambda_1, \dots, \lambda_n) \geq 0.$$

Bizonyítás

($\Leftarrow$) Tegyük fel, hogy $K_{p,f}(\lambda_1, \dots, \lambda_n) \geq 0$ minden p-re és $\lambda_, \dots, \lambda_n$-re.

Legyen $H \in M_n(\mathbb{C})$ tetszőleges önadjungált mátrix, ekkor alkalmazzuk az 1.31 lemmát $K = H_p$-ra (ez még mindig az a diagonális mátrix, melynek főátlójában H-nak a p-ik sora szerepel), így

$$H_p^* K_{p,f}(\Lambda) H_p \geq H_p^* 0 H_p = 0.$$

Ekkor a most bizonyított 3.22 állítás alapján ezeket összegezve $p = 1 \dots n$-re

$$D^2 f(\Lambda)(H) = 2 \sum_{p=1}^{n} H_p^* K_{p,f}(\Lambda) H_p \geq 0.$$

Ez viszont a 3.14 állítás miatt azt jelenti, hogy mivel a második derivált pozitív szemidefint mátrix minden H-ra, ezért f n-konvex.

($\Rightarrow$) Tegyük fel, hogy f n-konvex. Ekkor azt akarjuk belátni, hogy tetszőleges $p \in \{1, \dots, n\}$ és $\lambda_1, \dots, \lambda_n \in I$ esetén $K_{p,f}(\lambda_1, \dots, \lambda_n) \geq 0$, azaz minden $\eta \in \mathbb{C}^n$-re

$$\langle K_{p,f}(\Lambda)\eta, \eta \rangle \geq 0.$$

Rögzítsük le p-t és η-t. Mivel f n-konvex, ezért (a másik irány utolsó lépéséhez hasonlóan)

$$D^2 f(\Lambda)(H) = 2 \sum_{p=1}^{n} H_p^* K_{p,f}(\Lambda) H_p \geq 0.$$

Ebből még nem következik, hogy a Krauss-mátrixok külön-külön pozitív szemidefinitek.

Mivel p rögzített, így definiálhatjuk a következőket:

Legyen $\varepsilon > 0$, és válasszuk ζ-nak

$$\zeta_i = \begin{cases} \frac{1}{\varepsilon} & \text{ha } i \neq p \\ 1 & \text{ha } i = p \end{cases},$$

továbbá legyen $H \in M_n(\mathbb{C})$ a következő

$$h_{(i,j)} = \begin{cases} \overline{\eta}_i \eta_j & \text{ha } i = j = p \\ \varepsilon \overline{\eta}_i \eta_j & \text{ha } i = p \text{ vagy } j = p \\ \varepsilon^2 \overline{\eta}_i \eta_j & \text{különben.} \end{cases}$$

Ekkor vegyük észre, hogy

$$H_q \zeta = \begin{pmatrix} h_{(q,1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_{(q,n)} \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_n \end{pmatrix} = \begin{cases} \overline{\eta}_p \eta & \text{ha } q = p \\ \varepsilon \overline{\eta}_q \eta & \text{ha } q \neq p. \end{cases}$$

Ezzel a helyettesítéssel, azt kapjuk, hogy

$$0 \leq \langle D^2 f(\Lambda)(H)\zeta, \zeta \rangle = 2 \sum_{q=1}^{n} \langle H_q^* K_{q,f}(\Lambda) H_q \zeta, \zeta \rangle =$$

$$= 2 \sum_{q=1}^{n} \langle K_{q,f}(\Lambda) H_q \zeta, H_q \zeta \rangle = 2 \langle K_{p,f}(\Lambda) H_p \zeta, H_p \zeta \rangle + 2 \sum_{q \neq p} \langle K_{q,f}(\Lambda) H_q \zeta, H_q \zeta \rangle =$$

$$= 2 \langle K_{p,f}(\Lambda) \overline{\eta}_p \eta, \overline{\eta}_p \eta \rangle + 2 \sum_{q \neq p} \langle K_{q,f}(\Lambda) \varepsilon \overline{\eta}_q \eta, \varepsilon \overline{\eta}_q \eta \rangle =$$

$$= 2 \overline{\eta}_p \eta_p \langle K_{p,f}(\Lambda) \eta, \eta \rangle + 2\varepsilon^2 \sum_{q \neq p} \overline{\eta}_q^2 \langle K_{q,f}(\Lambda) \eta, \eta \rangle.$$

Tehát ha $\varepsilon \to 0$, ezért mivel η rözített, így

$$0 \leq 2 \overline{\eta}_p \eta_p \langle K_{p,f}(\Lambda) \eta, \eta \rangle,$$

ezért ha $\eta_p \neq 0$, akkor

$$0 \leq \langle K_{p,f}(\Lambda) \eta, \eta \rangle.$$

Ha $\eta_p = 0$, akkor legyen $\eta_i' = \begin{cases} \eta_i + \delta & \text{ha } i = p \\ \eta_i & \text{ha } i \neq p. \end{cases}$

Ekkor $\delta \to 0$-ból ($\eta' \to \eta$) megkapjuk, hogy $K_{p,f}(\Lambda)$ pozitív szemidefinit.

3.24. Példa. *Az $f(t) = e^t$ függvény nem operátor-konvex $\mathbb{R}$-en.*

Bizonyítás: ellenpéldával belátjuk, hogy f nem 2-konvex.

Legyen $\lambda_1 = 1, \lambda_2 = 0$, ekkor

$$\frac{1}{2}K_{2,f}(\lambda_1, \lambda_2) = \begin{pmatrix} 1 & e-2 \\ e-2 & \frac{1}{2} \end{pmatrix},$$

itt $det(\frac{1}{2}K_{2,f}(\lambda_1, \lambda_2)) = -e^2 + 4e - \frac{7}{2} \approx -0.016 < 0$, tehát f nem 2-konvex $\mathbb{R}$-en.

3.25. Következmény. *Legyen $f \in C^2(I)$, ahol I nyílt intervallum. Ekkor f operátor-konvex akkor, és csak akkor, ha $g(t) = [t, t_0]_f$ operátor-monoton növő I-n minden rögzített $t_0 \in I$-ra.*

Bizonyítás:

Tegyük fel, hogy g n-monoton növő I-n minden t_0-ra. Ez azt jelenti, hogy $f'(t)$ n-monoton növő minden $t \in I$-re (ugyanis minden t_0 egy kis környezetében f monoton növő), tehát 3.14 miatt f n-konvex I-n.

Tegyük fel, hogy f $(n+1)$-konvex. Most azt fogjuk belátni, hogy ekkor g n-monoton növő. Legyen $\Lambda_{n+1} = diag(\lambda_1, \ldots, \lambda_{n+1}) \in M_{n+1}(\mathbb{C})$. Ekkor az f-hez tartozó Krauss-mátrixok minden $\lambda_1, \ldots, \lambda_{n+1}$-re pozitív szemidefinitek, ahol λ_i minden i-re befutja az egész I intervallumot.

Legyen $\zeta \in \mathbb{C}^n$ tetszőleges, rögzített.

Legyen $\eta \in \mathbb{C}^{n+1} = (1, \ldots, 1, 0)$ és legyen

$$H = \begin{pmatrix} 0 & \cdots & \cdots & \overline{\zeta}_1 \\ 0 & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & \overline{\zeta}_n \\ \zeta_1 & \cdots & \zeta_n & 0 \end{pmatrix}.$$

Mivel mindegyik Krauss-mátrix pozitív szemidefinit, ezért

$$0 \leq 2\sum_{p=1}^{n+1} \langle H_p^* K_{p,f}(\Lambda_{n+1}) H_p \eta, \eta \rangle = 2\sum_{p=1}^{n+1} \langle K_{p,f}(\Lambda_{n+1}) H_p \eta, H_p \eta \rangle.$$

(Továbbra is $H_p = diag(H$ p-ik sora).) Ezekre a H és η paraméterekre

$$H_p \eta = \begin{cases} 0 \text{ ha } p \leq n \\ (\zeta_1, \ldots, \zeta_n, 0) = \widehat{\zeta} \text{ ha } p = n+1, \end{cases}$$

ezért az előzőt továbbalakíthatjuk, hogy

$$0 \leq \langle K_{n+1,f}(\Lambda_{n+1})\widehat{\zeta}, \widehat{\zeta}\rangle.$$

Vizsgáljuk meg az $(n+1)$-ik Krauss-mátrix (i,j)-ik elemét (az átláthatóság végett ennek a felét vizsgáljuk), ha $i, j \leq n$:

$$\frac{1}{2}(K_{n+1,f}(\Lambda_{n+1}))_{(i,j)} = [\lambda_{n+1}, \lambda_i, \lambda_j]_f = [\lambda_i, \lambda_{n+1}, \lambda_j]_f =$$

$$= \frac{[\lambda_{n+1}, \lambda_i]_f - [\lambda_{n+1}, \lambda_j]_f}{\lambda_i - \lambda_j} = \frac{g(\lambda_i) - g(\lambda_j)}{\lambda_i - \lambda_j} = [\lambda_i, \lambda_j]_g,$$

ha t_0-nak választottuk λ_{n+1}-et g definíciójában. Tehát írhatjuk, hogy

$$0 \leq \langle K_{n+1,f}(\Lambda_{n+1})\widehat{\zeta}, \widehat{\zeta}\rangle = \langle 2L_g(\Lambda_n)\zeta, \zeta\rangle,$$

ahol $\Lambda_n = diag(\lambda_1, \ldots, \lambda_n)$. Mivel $\zeta \in \mathbb{C}^n$ tetszőleges volt, ezért $L_g(\Lambda_n)$ pozitív szemidefinit, tehát a 3.19 tétel miatt g n-monoton növő I-n.

Ezzel az eredménnyel általánosítottuk az eddigi kapcsolatot a mátrix-monoton és mátrix-konvex függvények között, amit az előző fejezetben láttunk be a $(0, \infty)$ intervallumon, és bővebb körben alkalmazható állításokhoz jutottunk a Löwner- és a Krauss-mátrixok segítségével.

4. Lieb konkávitás-tétele

Zárásként levezetjük Lieb konkávitás-tételét, amit elsőként Lieb bizonyított be, [3]-ben megtalálható az eredeti bizonyítása, ahol az $f(A) = A^q$ $(0 \leq q \leq 1)$ függvény konkávitásából indult ki. Mi azonban a Rajendra Bhatia [1, IX]-ben található, B. Simon és T. Ando bizonyítását mutatjuk itt be Lieb tételének egy gyengébb változatára.

4.1. Lemma. *Legyenek az $R_1, R_2, S_1, S_2, T_1, T_2 \in M_n(\mathbb{C})$ pozitív szemidefinit mátrixok (azaz $R_i, S_i, T_i \geq 0$, ahol $i \in \{1, 2\}$) olyanok, hogy*

- $R_1 R_2 = R_2 R_1$,
- $S_1 S_2 = S_2 S_1$,
- $T_1 T_2 = T_2 T_1$,
- $R_1 \geq S_1 + T_1$,
- $R_2 \geq S_2 + T_2$.

Ekkor minden $0 \leq t \leq 1$ esetén

$$R_1^t R_1^{1-t} \geq S_1^t S_2^{1-t} + T_1^t T_2^{1-t}.$$

Bizonyítás:
Legyen $E = \{t \in [0, 1] \mid R_1^t R_1^{1-t} \geq S_1^t S_2^{1-t} + T_1^t T_2^{1-t}\}$.
Ekkor a feltevések szerint $0 \in E$ és $1 \in E$.
Először azt látjuk be, hogy $\frac{1}{2} \in E$. Legyen $x, y \in \mathbb{C}^n$ tetszőleges. Ekkor

$$|\langle x, (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) y\rangle| \leq |\langle x, S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} y\rangle| + |\langle x, T_1^{\frac{1}{2}} T_2^{\frac{1}{2}} y\rangle| =$$

Felhasználjuk, hogy S_1 és T_1 önadjungált, majd ezt tovább becsüljük felülről a Cauchy-Schwarz egyenlőtlenséggel $(|\langle a, b\rangle| \leq \|a\|\|b\|)$, így azt kapjuk, hogy

$$= |\langle S_1^{\frac{1}{2}} x, S_2^{\frac{1}{2}} y\rangle| + |\langle T_1^{\frac{1}{2}} x, T_2^{\frac{1}{2}} y\rangle| \leq \|S_1^{\frac{1}{2}} x\|\|S_2^{\frac{1}{2}} y\| + \|T_1^{\frac{1}{2}} x\|\|T_2^{\frac{1}{2}} y\| \leq$$

Most alkalmazzuk a Cauchy-Schwarz egyenlőtlenség egy másik alakját $(|a_1 b_1 + a_2 b_2|^2 \leq (a_1^2 + a_2^2)(b_1^2 + b_2^2))$, így azt kapjuk, hogy

$$\leq \left(\|S_1^{\frac{1}{2}} x\|^2 + \|T_1^{\frac{1}{2}} x\|^2\right)^{\frac{1}{2}} \left(\|S_2^{\frac{1}{2}} y\|^2 + \|T_2^{\frac{1}{2}} y\|^2\right)^{\frac{1}{2}}. \tag{33}$$

Vegyük észre, hogy mivel S_1 önadjungált, ezért

$$\|S_1^{\frac{1}{2}}x\|^2 = \langle S_1^{\frac{1}{2}}x, S_1^{\frac{1}{2}}x\rangle = \langle S_1x, x\rangle.$$

Ezt az átalakítást alkalmazva négyszer, (33) megegyezik a következővel:

$$= (\langle S_1x, x\rangle + \langle T_1x, x\rangle)^{\frac{1}{2}} (\langle S_2y, y\rangle + \langle T_2y, y\rangle)^{\frac{1}{2}} = \langle (S_1+T_1)x, x\rangle^{\frac{1}{2}} \langle (S_2+T_2)y, y\rangle^{\frac{1}{2}} \leq$$

Feltevés szerint ($\langle R_i\zeta, \zeta\rangle \geq \langle (S_i+T_i)\zeta, \zeta\rangle \; \forall\zeta \in \mathbb{C}^n$) ezt felülről becsülhetjük azzal, hogy

$$\leq \langle R_1x, x\rangle^{\frac{1}{2}} \langle R_2y, y\rangle^{\frac{1}{2}} = (\langle R_1x, x\rangle \langle R_2y, y\rangle)^{\frac{1}{2}}.$$

Tehát eddig beláttuk, hogy tetszőleges x, y-ra

$$|\langle x, (S_1^{\frac{1}{2}}S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}}T_2^{\frac{1}{2}})y\rangle| \leq (\langle R_1x, x\rangle \langle R_2y, y\rangle)^{\frac{1}{2}}. \tag{34}$$

Ezzel belátjuk, hogy minden u és v egységvektorra (felhasználva, hogy R_1 önadjungált)

$$|\langle u, R_1^{-\frac{1}{2}}(S_1^{\frac{1}{2}}S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}}T_2^{\frac{1}{2}})R_2^{-\frac{1}{2}}v\rangle| = |\langle R_1^{-\frac{1}{2}}u, (S_1^{\frac{1}{2}}S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}}T_2^{\frac{1}{2}})R_2^{-\frac{1}{2}}v\rangle| \leq$$

(itt használjuk az (34) egyenlőtlenséget, ahol $x = R_1^{-\frac{1}{2}}u$ és $y = R_2^{-\frac{1}{2}}v$)

$$\leq \left(\langle R_1R_1^{-\frac{1}{2}}u, R_1^{-\frac{1}{2}}u\rangle \langle R_2R_2^{-\frac{1}{2}}v, R_2^{-\frac{1}{2}}v\rangle\right)^{\frac{1}{2}} = \left(\langle R_1^{\frac{1}{2}}u, R_1^{-\frac{1}{2}}u\rangle \langle R_2^{\frac{1}{2}}v, R_2^{-\frac{1}{2}}v\rangle\right)^{\frac{1}{2}} =$$

$$= (\langle u, u\rangle \langle v, v\rangle)^{\frac{1}{2}} = \|u\|\|v\| = 1.$$

Ez most azt jelenti, hogy

$$\|R_1^{-\frac{1}{2}}(S_1^{\frac{1}{2}}S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}}T_2^{\frac{1}{2}})R_2^{-\frac{1}{2}}\| \leq 1.$$

Felhasználjuk, hogy $0 \leq A, B$ mátrixokra (ha az egyik invertálható), akkor $\varrho(AB) = \varrho(BA)$, és ha X normális, akkor $\varrho(X) = \|X\|$. Ebből, továbbá R_1 és R_2 kommutativitásából azt kapjuk, hogy

$$1 \geq \|(R_1^{-\frac{1}{2}}(S_1^{\frac{1}{2}}S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}}T_2^{\frac{1}{2}})R_2^{-\frac{1}{4}})R_2^{-\frac{1}{4}}\| \geq \varrho((R_1^{-\frac{1}{2}}(S_1^{\frac{1}{2}}S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}}T_2^{\frac{1}{2}})R_2^{-\frac{1}{4}})R_2^{-\frac{1}{4}}) =$$

$$= \varrho(R_2^{-\frac{1}{4}} R_1^{-\frac{1}{2}} (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) R_2^{-\frac{1}{4}}) = \varrho(R_1^{-\frac{1}{2}} R_2^{-\frac{1}{4}} (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) R_2^{-\frac{1}{4}}) =$$

$$= \varrho(R_1^{-\frac{1}{4}} (R_1^{-\frac{1}{4}} R_2^{-\frac{1}{4}} (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) R_2^{-\frac{1}{4}})) = \varrho(R_1^{-\frac{1}{4}} R_2^{-\frac{1}{4}} (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) R_2^{-\frac{1}{4}} R_1^{-\frac{1}{4}}).$$

Vegyük észre, hogy $R_1^{-\frac{1}{4}} R_2^{-\frac{1}{4}} (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) R_2^{-\frac{1}{4}} R_1^{-\frac{1}{4}}$ önadjungált. Ez abból következik, hogy egyrészt a zárójelen belül mindkét tag önadjungált, mert felcserélhető önadjungált mátrixok szorzata. Másrészt pedig balról és jobbról meg van szorozva egy K, illetve K^* mátrixszal (azaz ha A önadjungált, akkor $(KAK^*)^* = (K^*)^* A^* K^* = KAK^*$). Mivel önadjungált, ezért

$$\|R_1^{-\frac{1}{4}} R_2^{-\frac{1}{4}} (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) R_2^{-\frac{1}{4}} R_1^{-\frac{1}{4}}\| = \varrho(R_1^{-\frac{1}{4}} R_2^{-\frac{1}{4}} (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) R_2^{-\frac{1}{4}} R_1^{-\frac{1}{4}}) \le 1. \quad (35)$$

Ez már megfelelő alakban van, ezért most alakítsuk ezt vissza úgy, hogy

$$R_1^{-\frac{1}{4}} R_2^{-\frac{1}{4}} (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) R_2^{-\frac{1}{4}} R_1^{-\frac{1}{4}} \le I. \quad (36)$$

Alkalmazzuk az (1.31) lemmát először $K = R_1^{\frac{1}{4}}$-re, majd mégegyszer $K = R_2^{\frac{1}{4}}$-re, így

$$R_2^{\frac{1}{4}} R_1^{\frac{1}{4}} R_1^{-\frac{1}{4}} R_2^{-\frac{1}{4}} (S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}}) R_2^{-\frac{1}{4}} R_1^{-\frac{1}{4}} R_1^{\frac{1}{4}} R_2^{\frac{1}{4}} \le R_2^{\frac{1}{4}} R_1^{\frac{1}{4}} I R_1^{\frac{1}{4}} R_2^{\frac{1}{4}},$$

azaz (R_1 és R_2 kommutativitását újra felhasználva)

$$S_1^{\frac{1}{2}} S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}} T_2^{\frac{1}{2}} \le R_2^{\frac{1}{4}} R_1^{\frac{1}{2}} R_2^{\frac{1}{4}} = R_1^{\frac{1}{2}} R_2^{\frac{1}{2}}.$$

Ezzel beláttuk, hogy $\frac{1}{2} \in E$. Most azt látjuk be, hogy E konvex halmaz, azaz ha $\lambda, \mu \in E$, akkor $\frac{\lambda+\mu}{2} \in E$.

Legyen $\lambda, \mu \in E$, ekkor az teljesül, hogy

- $R_1^{\lambda} R_2^{1-\lambda} \ge S_1^{\lambda} S_2^{1-\lambda} + T_1^{\lambda} T_2^{1-\lambda}$,
- $R_1^{\mu} R_2^{1-\mu} \ge S_1^{\mu} S_2^{1-\mu} + T_1^{\mu} T_2^{1-\mu}$.

Vegyük észre, hogy ezek alakja a következő

- $\widehat{R}_1 \ge \widehat{S}_1 + \widehat{T}_1$,

- $\widehat{R}_2 \geq \widehat{S}_2 + \widehat{T}_2$,

és a kommutativitás is teljesül, ugyanis (csak az egyiket írjuk le, a másik kettő ugyanúgy levezethető)

$$\widehat{R}_1\widehat{R}_2 = R_1^\lambda R_2^{1-\lambda} R_1^\mu R_2^{1-\mu} = R_1^\mu R_2^{1-\mu} R_1^\lambda R_2^{1-\lambda} = \widehat{R}_2\widehat{R}_1.$$

Ezért alkalmazhatjuk ezt a lemmát $t = \frac{1}{2}$-re (erre már beláttuk), így

$$(R_1^\lambda R_2^{1-\lambda})^{\frac{1}{2}}(R_1^\mu R_2^{1-\mu})^{\frac{1}{2}} \geq (S_1^\lambda S_2^{1-\lambda})^{\frac{1}{2}}(S_1^\mu S_2^{1-\mu})^{\frac{1}{2}} + (T_1^\lambda T_2^{1-\lambda})^{\frac{1}{2}}(T_1^\mu T_2^{1-\mu})^{\frac{1}{2}}.$$

Kisebb átalakítással azt kapjuk, hogy

$$R_1^{\frac{\lambda+\mu}{2}} R_2^{1-\frac{\lambda+\mu}{2}} \geq S_1^{\frac{\lambda+\mu}{2}} S_2^{1-\frac{\lambda+\mu}{2}} + T_1^{\frac{\lambda+\mu}{2}} T_2^{1-\frac{\lambda+\mu}{2}},$$

tehát $\frac{\lambda+\mu}{2} \in E$, ezért a lemma igaz minden $t \in [0,1]$-re.

4.2. Definíció. *Legyen f egy kétváltozós $\mathbb{K}^2 \to \mathbb{R}$ függvény. Ekkor f kölcsönösen n-konkáv A, B-ben I-n, ha minden $0 \leq \lambda \leq 1$ esetén tetszőleges $A_1, A_2, B_1, B_2 \in M_n(\mathbb{C})$ önadjungált mátrixokra, melyek sajátértékei I-ben vannak,*

$$f(\lambda A_1 + (1-\lambda)A_2, \lambda B + (1-\lambda)B_2) \geq \lambda f(A_1, B_1) + (1-\lambda) f(A_2, B_2).$$

Megjegyzés: f kölcsönösen n-konvex I-n, ha $-f$ kölcsönösen n-konkáv.

Megjegyzés: Ha f kölcsönösen n-konkáv A, B-ben I-n, akkor rögzített A vagy B mellett egy változóban is n-konkáv.

4.3. Tétel. *Legyen $K \in M_n(\mathbb{C})$, és legyen $0 \leq t \leq 1$ valós szám. Ekkor a következő $f : \{X \in M_n(\mathbb{C}) | X^* = X, X \geq 0\}^2 \to \mathbb{R}$,*

$$f(A, B) = tr(K^* A^t K B^{1-t})$$

folytonos függvény kölcsönösen konkáv A, B-ben.

Bizonyítás:

Legyenek $0 \leq A_1, A_2, B_1, B_2 \in M_n(\mathbb{C})$ önadjungáltak, továbbá legyen $A = A_1 + A_2$ és $B = B_1 + B_2$. Jelölje $\widetilde{A}, \widetilde{B} : M_n(\mathbb{C}) \to M_n(\mathbb{C})$, $\widetilde{A}(K) = AK$ a balról szorzás-, és $\widetilde{B}(K) = KB$ a jobbról szorzás operátorát. Hasonlóan definiálható $\widetilde{A}_1(K), \widetilde{A}_2(K)$, illetve $\widetilde{B}_1(K), \widetilde{B}_2(K)$.

Most belátjuk, hogy $\widetilde{A}$, illetve $\widetilde{B}$ pozitív operátorok (hasonlóan belátható, hogy $\widetilde{A}_1, \widetilde{A}_2, \widetilde{B}_1, \widetilde{B}_2$ is pozitívak).

Legyen V_i azon K mátrixok halmaza, amelyben csak az i-edik oszlop lehet nullától különböző. Ekkor V_i invariáns altere $\widetilde{A}$-nak és rajta $\widetilde{A}$ mátrixa (a kézenfekvő bázisban) A.

Emiatt az egész $M_n(\mathbb{C})$ téren $\widetilde{A}$-t reprezentálja az az $n^2 \times n^2$ méretű blokkdiagonális mátrix, melynek mindegyik diagonális blokkja A. Viszont A-ról tudjuk, hogy pozitív szemidefinit, tehát $\widetilde{A}$ mátrixa is pozitív szemidefinit.

Hasonlóan kapjuk, hogy a jobbról szorzásnál K sorai szolgáltatják az n dimenziós invariáns altereket. Ezeken az altereken vett jobbról szorzás mátrixa viszont $B^\top$, ezért az egész $M_n(\mathbb{C})$ téren $\widetilde{B}$ mátrix blokkdiagonális elemei mindegyike $B^\top$. Mivel B pozitív szemidefinit, ezért $B^\top$ is az, tehát $\widetilde{B}$ mátrixa is pozitív szemidefinit.

Ekkor az előző 4.1 lemma alapján $0 \leq t \leq 1$-re

$$\widetilde{A}^t \widetilde{B}^{1-t}(K) \geq \widetilde{A}_1^t \widetilde{B}_1^{1-t}(K) + \widetilde{A}_2^t \widetilde{B}_2^{1-t}(K).$$

Ez azt jelenti, hogy minden $K \in M_n(\mathbb{C})$-re

$$\langle K, \widetilde{A}^t \widetilde{B}^{1-t}(K)\rangle \geq \langle K, \widetilde{A}_1^t \widetilde{B}_1^{1-t}(K) + \widetilde{A}_2^t \widetilde{B}_2^{1-t}(K)\rangle,$$

ahol a skalárszozat itt a mátrixok körében értelmezett (definíció szerint $\langle X, Y\rangle = tr(X^*Y)$). Így a balról-jobbról való szorzásokat elvégezve azt kapjuk, hogy

$$tr(K^* A^t K B^{1-t}) \geq tr(K^* A_1^t K B_1^{1-t}) + tr(K^* A_2^t K B_2^{1-t}).$$

Ha leosztunk kettővel, akkor azt kapjuk, hogy

$$\frac{1}{2} tr(K^* A^t K B^{1-t}) \geq \frac{1}{2} tr(K^* A_1^t K B_1^{1-t}) + \frac{1}{2} tr(K^* A_2^t K B_2^{1-t}). \tag{37}$$

Végül vegyük észre, hogy

$$\frac{1}{2} tr(K^* A^t K B^{1-t}) = tr\left(\frac{1}{2} K^* A^t K B^{1-t}\right) = tr\left(\frac{1}{2^{t+(1-t)}} K^* A^t K B^{1-t}\right) =$$

$$= tr\left(K^* \left(\frac{A}{2}\right)^t K \left(\frac{B}{2}\right)^{1-t}\right) = tr\left(K^* \left(\frac{A_1 + A_1}{2}\right)^t K \left(\frac{B_1 + B_2}{2}\right)^{1-t}\right).$$

Így azt kapjuk (37)-ből, hogy

$$f\left(\frac{A_1 + A_2}{2}, \frac{B_1 + B_2}{2}\right) \geq \frac{1}{2} f(A_1, B_1) + \frac{1}{2} f(A_2, B_2). \tag{38}$$

Ez pont a gyenge kölcsönös konkávitás feltétele, tehát f folytonosságából következik, hogy f kölcsönösen konkáv A, B -ben.

Hivatkozások

[1] Rajendra Bhatia, Matrix Analysis, Springer Science and Business Media, 1997.

[2] Frank Hansen, The fast track to Löwner's theorem, 2013,
`http://arxiv.org/abs/1112.0098`

[3] Elliott H. Lieb, Convex Trace Functions and the Wigner-Yanase-Dyson Conjecture, Academic Press, Inc, 1973,
`http://www.sciencedirect.com/science/article/pii/000187087390011X`

[4] Julius Bendat and Seymour Sherman, Monotone and convex operator functions, 1954,
`http://www.ams.org/journals/tran/1955-079-01/S0002-9947-1955-0082655-4/`

[5] Eduard Jorswieck and Holger Boche, Majorization and Matrix Monotone Functions in Wireless Communications, Now Publishers Inc, 2007

[6] Fumio Hiai and Denes Petz, Introduction to Matrix Analysis and Applications, Springer Science and Business Media, 2014

[7] Jordan Bell, Fréchet derivatives and Gâteaux derivatives, Department of Mathematics, University of Toronto, 2014,
`http://individual.utoronto.ca/jordanbell/notes/frechetderivatives.pdf`

[8] The Waterloo Fractal Coding and Analysis Group, Additional notes on Fréchet derivatives, AMATH 731: Applied Functional Analysis, 2014, `http://links.uwaterloo.ca/amath731docs/frechet.pdf`

[9] Tsuyoshi Ando, Trace-Inequalities and Matrix-Convex Functions, Hokkaido University, Japan, 2009,
`http://www.fixedpointtheoryandapplications.com/content/pdf/1687-1812-2010-241908.pdf`