Mátrixelmélet alkalmazásai

Egyenlőtlenségek

SZAKDOLGOZAT

Készítette: Oláh Vivien

Matematika BSc - elemző szakirány

Témavezető: Fialowski Alice, egyetemi docens

ELTE TTK, Algebra és Számelmélet Tanszék

Eötvös Loránd Tudományegyetem
Természettudományi Kar
Budapest, 2013

Tartalomjegyzék

Bevezetés

A lineáris algebra egyik leghasznosabb és legtöbb területen alkalmazott fogalmának a mátrixokat tekinthetjük. A tanulmányozásukra és tulajdonságaik leírására szolgáló mátrixelmélet korábban csak a lineáris algebra egy részét képezte, ma már azonban rendkívül fontos, önálló matematikai területté vált. Tanulmányaim során számos alkalommal tapasztaltam, milyen nagy hasznunkra válik e terület eredményeinek alkalmazása, mind a matematika egyéb területein, mind a matematikán kívül.

A matematikai feladatok megoldása során rengetegszer fordul elő, hogy az eredmények meghatározása körülményes, hosszadalmas művelet, vagy valamilyen ismeret hiányában nem kivitelezhető, így csupán becslésekre hagyatkozhatunk. Hasonló helyzet állhat elő abban az esetben is, ha a mátrix egyik legalapvetőbb jellemzőit, a sajátértékeit próbáljuk meghatározni. A mátrixok számtalan becslés alapjául szolgálnak. Dolgozatomban olyan mátrix egyenlőtlenségekkel foglalkozom, melyekben sajátértékek szerepelnek.

Első körben tisztázom az alapvető mátrixelméleti fogalmakat, majd bevezetem a sajátérték fogalmát, illetve röviden tárgyalom Gersgorin elméletét, mely az egyik legismertebb sajátértékekre vonatkozó becslés. Definiálok néhány speciális mátrixot, mely az egyenlőtlenségek csoportosításában valamint a bizonyításokban lesz hasznunkra, majd áttérek a mátrixegyenlőtlenségekre. Elsőként a pozitív definit és szemidefinit mátrixokra vonatkozó egyenlőtlenségekkel foglalkozom, majd néhány normális mátrixszal kapcsolatos egyenlőtlenséget ismertetek. A dolgozat végén pedig szót ejtek a majorizációról és Hermite-féle mátrixokra vonatkozó majorizációs egyenlőtlenségeket mutatok be.

1. fejezet

Mátrixelméleti alapok

A dolgozatban $n \times n$-es, kvadratikus mátrixokkal foglalkozunk, amelyeket $\mathbb{R}$ vagy $\mathbb{C}$ felett értelmezünk, ezek halmazát $\mathbb{R}^{n\times n}$-el és $\mathbb{C}^{n\times n}$-el jelöljük. Továbbá I_n, vagy egyszerűen jelölve I az egységmátrix, melynek minden főátlóbeli eleme 1, az azon kívüli elemek pedig mind nullák. Most vegyük sorra az alapfogalmakat. Ebben a fejezetben az [1]-[4], [6], [9] forrásokat használtam fel.

1.1. Definíció. *Az $A \in \mathbb{R}^{n\times n}$ mátrixból a sorok és oszlopok felcserélésével képzett, A^T-vel jelölt mátrixot az A transzponáltjának nevezzük.*

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}$$

Ha $A = A^T$, szimmetrikus mátrixról beszélünk.

1.2. Definíció. *Az $A \in \mathbb{C}^{n\times n}$ mátrix konjugált transzponáltjának nevezzük az A^* mátrixot, amelyet akkor kapunk, ha minden a_{ij} elem helyére az a_{ji} elem komplex konjugáltját írjuk. A^*-t az A adjungáltjának is nevezzük.*

$$A^* = \begin{pmatrix} \overline{a}_{11} & \overline{a}_{21} & \dots & \overline{a}_{n1} \\ \overline{a}_{12} & \overline{a}_{22} & \dots & \overline{a}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a}_{1n} & \overline{a}_{2n} & \dots & \overline{a}_{nn} \end{pmatrix}$$

Egy $a + bi$ komplex szám konjugáltja $\overline{a+bi} = a - bi$.

1.3. Definíció. *Ha* $A \in \mathbb{C}^{n\times n}$*-hez létezik olyan* $B \in \mathbb{C}^{n\times n}$*, melyre* $AB = BA = I$*, akkor* A *invertálható, és inverze az* $A^{-1} = B$ *mátrix. Az ilyen mátrixokat regulárisnak, míg a nem invertálható mátrixokat szingulárisnak nevezzük.*

Megjegyezzük, hogy ha A invertálható, akkor csak egy inverze van. Ugyanis ha feltesszük, hogy C is inverze, akkor $CA = AC = I$ és ebből:

$$B = BI = B(AC) = (BA)C = IC = C \quad \text{azaz} \quad B = C.$$

1.4. Definíció. *Azt mondjuk, hogy az* A *és* B *mátrixok hasonlóak, ha létezik olyan reguláris* S *mátrix, hogy* $B = S^{-1}AS$*.*

1.5. Definíció. *Egy* A n*-ed rendű mátrix nyoma a főátlójában álló elemek összege. Jelölés:* $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$.

A mátrix egy másik, igen fontos jellemzője a determináns. Definiálásához szükségünk van a permutáció, valamint az inverzió fogalmának tisztázására:

- Az $\mathbb{N}_n = 1, 2, \ldots, n$ számok valamely sorrendben való felírását az $1, 2, \ldots, n$ elemek egy *permutációjának* nevezzük. Két permutáció akkor tekintendő különbözőnek, ha legalább egy elem elhelyezésében eltérnek egymástól. $\mathbb{N}_n$ összes permutációinak halmaz $\mathbb{S}_n$, elemszáma $n!$.

- Legyen $(a_1, a_2, \ldots, a_i, \ldots, a_j, \ldots, a_n)$ az $1, 2, \ldots, n$ elemek egy permutációja. Azt mondjuk, hogy a_i és a_j elempár inverzióban állnak, ha $i < j$ és $a_i > a_j$. Az *inverziószám* az inverzióban álló elempárok száma.

Ezek után a következőképpen tekinthetünk egy mátrix determinánsára:

1.6. Definíció. *Egy* $A = (a_{ij})$ n*-ed rendű mátrix determinánsa:*

$$\det(A) = \sum_{\sigma \in \mathbb{S}_n} (-1)^{I(\sigma)} a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)},$$

ahol $I(\sigma)$ *a* σ *permutáció inverziószáma.*

1.1. Tétel. *Legyenek* A *és* B *kvadratikus mátrixok. Ekkor a két mátrix szorzatának determinánsa egyenlő a determinánsaik szorzatával.*

$$\det(AB) = \det(A)\det(B).$$

Ennek következményeként elmondhatjuk, hogy egy négyzetes A mátrix akkor és csak akkor invertálható, ha $\det(A) \neq 0$.

1.7. Definíció. *Azokat a mátrixokat, melynek elemei mátrixok, blokkmátrixoknak vagy particionált mátrixoknak nevezzük.*

Példa:

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) = \left(\begin{array}{ccc|cc} 1 & 0 & 1 & 0 & 1 \\ 2 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right),$$

ahol a blokkok:

$$A_{11} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, A_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 2 & 3 \end{pmatrix}, A_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_{22} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

1.1. Állítás. *Két azonos módon particionált blokkmátrix esetén az összeadás és a skalárral való szorzás blokkonként elvégezhető.*

Dolgozatomban a blokkmátrixok közül a 2×2-es blokkmátrixok bizonyulnak a leghasznosabbnak, főleg ha 1 vagy 2 blokk a 0 mátrixszal egyenlő. A 2×2-es blokkmátrixok szorzatára a következőképpen tekinthetünk:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

2. fejezet

Sajátértékek és Gersgorin elmélet

A fejezethez felhasznált források a [2],[6],[9],[13],[15].
Jelöljük a komplex elemű $n \times n$-es mátrixok halmazát $\mathbb{M}_n$-el. Legyen $A \in \mathbb{M}_n$, λ pedig komplex skalár. Ekkor az A mátrix sajátértékét a következőképpen definiálhatjuk.

2.1. Definíció. *Azt mondjuk, hogy λ az $A \in \mathbb{M}_n$ mátrix sajátértéke, ha létezik olyan nullától különböző x vektor, melyre $Ax = \lambda x$. Ekkor x az A mátrix λ sajátértékéhez tartozó sajátvektora.*

2.2. Definíció. *Az $A \in \mathbb{M}_n$ mátrixhoz tartozó karakterisztikus polinomnak nevezzük a $P(\lambda) = \det(A - \lambda I)$ polinomot. Ennek gyökei adják az A mátrix sajátértékeit. A λ sajátértékhez tartozó x sajátvektorokat az $(A - \lambda I)x = 0$ egyenletből számíthatjuk ki.*

2.1. Tétel. *Az $A \in \mathbb{M}_n$ mátrixnak a multiplicitást is figyelembe véve pontosan n darab sajátértéke van.*

2.2. Tétel. *Egy $A \in \mathbb{M}_n$ mátrixnak pontosan akkor nincs nulla sajátértéke, ha nemszinguláris.*

2.3. Tétel. *Hasonló mátrixok sajátértékei megegyeznek.*

2.3. Definíció. *Legyen $A \in \mathbb{M}_n$ és legyenek $\mu_1, \ldots, \mu_i$ az A^*A mátrix sajátértékei. Ekkor a nemnegatív $\sigma_i = \sqrt{\mu_i}$ számokat az A mátrix szinguláris értékeinek nevezzük.*

A sajátértékek ismeretében a determinánsra, valamint a nyomra új összefüggéseket írhatunk fel.

2.4. Tétel. *Legyenek $A \in \mathbb{M}_n$ sajátértékei $\lambda_1 \dots \lambda_n$, ekkor*

$$\det(A) = \prod_{i=1}^{n} \lambda_i, \quad \operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i.$$

Ezek az értékek megjelennek a karakterisztikus polinomban: a determináns a konstans tag, a nyom a $(-\lambda)^{n-1}$ együtthatója.

Most pedig nézzük Gersgorin tételét, valamint a tétel bizonyítását.

2.5. Tétel (Gersgorin). *Legyen $A \in \mathbb{M}_n$ valamint*

$$r_i = \sum_{j=1, j\neq i}^{n} |a_{ij}|, \quad i = 1, 2, \dots, n.$$

Ekkor A minden sajátértéke az a_{ii} körüli r_i sugarú körök uniójában helyezkedik el.

Bizonyítás: Legyen λ az A mátrix egy tetszőleges sajátértéke, x pedig a hozzá tartozó sajátvektor. Továbbá tegyük fel, hogy x_p az x vektor legnagyobb abszolút értékű koordinátája, azaz

$$|x_p| \geq |x_i|, \quad i = 1, 2, \dots, n.$$

Ekkor $x_p \neq 0$. Az $Ax = \lambda x$ mátrixegyenletből a p. sort felírva adódik, hogy

$$\sum_{j=1}^{n} a_{pj} x_j = \lambda x_p.$$

Illetve, ha a baloldali összegből kivonom $a_{pp} x_p$-t, akkor

$$\sum_{j=1, j\neq p}^{n} a_{pj} x_j = x_p(\lambda - a_{pp}).$$

Vegyük mindkét oldal abszolút értékét. Ekkor felírhatjuk az alábbiakat:

$$|x_p||\lambda - a_{pp}| = \left| \sum_{j=1, j\neq p}^{n} a_{pj} x_j \right|$$

$$
\begin{aligned}
&\leq \sum_{j=1, j\neq p}^{n} |a_{pj}x_j| \\
&= \sum_{j=1, j\neq p}^{n} |a_{pj}||x_j| \\
&\leq |x_p| \sum_{j=1, j\neq p}^{n} |a_{pj}| \\
&= |x_p| r_p.
\end{aligned}
$$

Az egyenlőtlenséget $|x_p| \geq 0$-val osztva megkapjuk az állítást. Ez A minden sajátértékére megmutatható, tehát tényleg minden sajátérték a körökben helyezkedik el. ∎

3. fejezet

Speciális mátrixok

Ismerkedjünk meg néhány speciális tulajdonságú kvadratikus mátrixtípussal. Ezek fogják a továbbiakban tárgyalt egyenlőtlenségek csoportosítását segíteni, illetve tulajdonságaikat, sajátértékeikre vonatkozó sajátosságaikat felhasználjuk későbbi bizonyítások során. A fejezetben az [1],[3],[5],[6],[9] forrásokat használtam.

3.1. Diagonális és trianguláris mátrix

Elsőként foglalkozzunk a *diagonális* és a *trianguláris* mátrixokkal, melyek különlegességét a bennük szabályosan elhelyezkedő zérus elemek adják. A diagonális mátrix minden főátlón kívüli eleme nulla, azaz minden $i \neq j$-re $a_{ij} = 0$. Jelölés:

$$\mathrm{diag}(a_{11}, \dots, a_{nn}) = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Speciális diagonális mátrix az I egységmátrix. Trianguláris mátrixok esetén beszélhetünk *alsó* és *felső háromszögmátrixokról.* Előbbinél a főátló feletti, utóbbinál a főátló alatti elemek egyenlőek nullával, míg a többi elem helyén nullától eltérő értékek is lehetnek.

Könnyen megmutatható, hogy diagonális mátrixot szorozva diagonális mátrixszal, eredményként szintén diagonális mátrixot kapunk. Valamint két alsó és két felső háromszögmátrix szorzata rendre ugyancsak alsó és felső háromszögmátrixot ad. A későbbi tételek kimondása és bizonyítása során is fogunk használni felső háromszögmátrixokat, amiket mindig T-vel fogunk jelölni.

3.1. Tétel (Háromszögmátrixok sajátértéke). *A háromszögmátrixok és így a diagonális mátrixok sajátértékei megegyeznek a főátló elemeivel.*

3.1. Definíció. *Ha az A mátrix hasonló egy diagonális mátrixhoz, akkor diagonalizálhatónak nevezzük.*

3.2. Szimmetrikus és Hermite-féle mátrix

Most térjünk át azokra a mátrixokra, melyeket transzponáltjuk, vagy transzponált konjugáltjuk segítségével definiálunk valamilyen formában. Azt már tudjuk, hogy egy A valós elemű kvadratikus mátrix *szimmetrikus*, ha megegyezik a transzponáltjával, azaz

$$A = A^T.$$

Az A mátrixot *ferdén szimmetrikusnak* mondjuk, ha

$$A = -A^T.$$

A komplex mátrixok körében is megfogalmazható egy analóg definíció. Az $A \in \mathbb{M}_n$ mátrixot *Hermite-félének* nevezzük, ha

$$A = A^*,$$

ahol A^* az A mátrix konjugált transzponáltja, vagy adjungáltja ($A^* = \overline{A}^T$). Továbbá azt mondjuk, hogy A *ferdén hermitikus* mátrix, ha

$$A = -A^*.$$

3.2. Definíció. *A $(V, (.,.))$ párt euklideszi térnek hívjuk, ha V egy vektortér $\mathbb{C}$ felett, és $(.,.) : (V \times V) \to \mathbb{C}$ egy adott függvény, ún. skaláris szorzat az alábbi tulajdonságokkal:*

- $(x,x) > 0$ *minden* $0 \neq x \in V$ *esetén.*

- $(\alpha x, y) = \alpha(x,y)$, *minden* $x, y \in V$, $\alpha \in \mathbb{C}$ *esetén.*

- $\overline{(y,x)} = (x,y)$ *minden* $x, y \in V$ *esetén.*

A $\mathbb{C}^n$ oszlopvektorok terében a belső szorzat: $(x,y) = y^*x = \overline{y_1}x_1 + \cdots + \overline{y_n}x_n$.

Megjegyzés: Az A mátrix adjungáltjának definíciója a belső szorzat segítségével: $(x, Ay) = (A^*x, y)$ minden $x, y \in \mathbb{C}^n$. A belső szorzatból adódó norma: $\|x\| = \sqrt{(x,x)}$.

3.2. Tétel. *Legyen* $A \in \mathbb{M}_n$. *Ekkor a következő állítások ekvivalensek:*

- *A hermitikus mátrix.*

- $x^*Ax \in \mathbb{R}$ *minden* $x \in \mathbb{C}^n$ *esetén.*

- $A^2 = A^*A$.

- $\mathrm{tr}(A^2) = \mathrm{tr}(A^*A)$.

3.3. Tétel. *Hermite-féle mátrix sajátértékei valósak.*

3.3. Pozitív (szemi)definit mátrix

Egy Hermite-féle (szimmetrikus) A mátrix *pozitív definit*, ha

$$(Ax, x) > 0 \quad \text{vagy} \quad \sum_{i,j} a_{ij}\overline{x}_i x_j > 0$$

minden nem nulla $x \in \mathbb{C}$ ($x \in \mathbb{R}$) vektorra. Az A mátrix *pozitív szemidefinit* illetve *nemnegatív definit*, ha

$$(Ax, x) \geq 0 \quad \text{vagy} \quad \sum_{i,j} a_{ij}\overline{x}_i x_j \geq 0$$

minden $x \in \mathbb{C}$ ($x \in \mathbb{R}$) vektorra. Az (Ax, x) forma az A mátrix *kvadratikus alakja*, ami szimmetrikus A és valós x vektor esetén

$$\sum_{i,j} a_{ij}\overline{x}_i x_j \quad \text{helyett a} \quad \sum_{i,j} a_{ij} x_i x_j$$

kifejezéssel egyenlő.

3.4. Tétel. *Legyen $A \in \mathbb{M}_n$ hermitikus. Ekkor a következő állítások ekvivalensek:*

- *A pozitív szemidefinit.*
- *A minden sajátértéke nemnegatív.*
- *A determinánsa nemnegatív.*
- *A minden főminorja nemnegatív.*

Megjegyzés: Az A mátrix i-edik főminorja az i-edik minormátrixának determinánsa. Az i-edik minormátrix alatt pedig A-nak azt az $i \times i$-s részmátrixát értjük, amelynek bal felső eleme a_{11}.

3.5. Tétel. *Legyen $A \in \mathbb{M}_n$ hermitikus. Ekkor a következő állítások ekvivalensek:*

- *A pozitív definit.*
- *A minden sajátértéke pozitív.*
- *A determinánsa pozitív.*
- *A minden főminorja pozitív.*

3.6. Tétel. *Egy pozitív szemidefinit mátrix pozitív definit akkor és csak akkor, ha nemszinguláris.*

3.4. Ortogonális és unitér mátrix

A következő két speciális mátrix definíciói szintén hasonlóak. Akárcsak a szimmetrikus és Hermite-féle mátrixok esetében, a különbség abból adódik, hogy valós vagy komplex mátrixokról beszélünk. Az $U \in \mathbb{M}_n$ mátrix *unitér* mátrix, ha

$$U^*U = UU^* = I$$

teljesül, azaz transzponált konjugáltja egyben az inverze is. Valós elemek esetén az utóbbi összefüggés megfelelője az

$$U^TU = UU^T = I$$

kifejezés, és az ezt kielégítő mátrixokat nevezzük *ortogonális* mátrixoknak.

3.7. Tétel. *Legyen $U \in \mathbb{M}_n$. Ekkor a következő állítások ekvivalensek:*

- *U unitér mátrix.*
- *$|\lambda| = 1$ az U mátrix minden λ sajátértékére.*
- *Minden $x \in \mathbb{C}^n$ vektorra $\|Ux\| = \|x\|$.*
- *Minden $x, y \in \mathbb{C}^n$ vektorra $(Ux, Uy) = (x, y)$.*

3.5. Normális mátrix

Minden szimmetrikus, Hermite-féle, unitér és ortogonális mátrix felcserélhető a transzponált konjugáltjával, azaz $A^*A = AA^*$. Ez éppen a *normális* mátrix definíciója, tehát a felsorolt mátrixok mindegyike egyben normális is.

3.8. Tétel. *Minden normális háromszögmátrix diagonális.*

3.9. Tétel. *Legyen $A \in \mathbb{M}_n$, $\lambda_1, \ldots, \lambda_n$ sajátértékekkel. Ekkor a következő állítások ekvivalensek:*

- *A normális mátrix.*

- $\mathrm{tr}(A^*A) = \sum_{i=1}^n |\lambda_i|^2$.
- *A szinguláris értékei* $|\lambda_1|, \ldots, |\lambda_n|$.
- $(A^*A)^2 = (A^*)^2A^2$.
- $(Ax, Ay) = (A^*x, A^*y)$ *minden* $x, y \in \mathbb{C}^n$*-re.*

3.6. Duplán sztochasztikus és permutáció mátrix

Egy négyzetes mátrixot *duplán sztochasztikusnak* nevezünk, ha minden eleme nemnegatív és minden sorában illetve oszlopában az elemek összege 1. Ez ekvivalens azzal, hogy egy nemnegatív elemekből álló A mátrix duplán sztochasztikus, ha

$$e^T A = e^T \quad \text{és} \quad Ae = e, \quad \text{ahol} \quad e = (1, 1, \ldots, 1)^T.$$

Duplán sztochasztikus mátrixok a *permutáció mátrixok* is, amelyeknek minden sorában és oszlopában kizárólag $0, 1$ elemek állnak, méghozzá úgy, hogy egy sorban és oszlopban pontosan egy darab 1-es áll.

3.10. Tétel. *Duplán sztochasztikus mátrixok szorzata is duplán sztochasztikus mátrix.*

3.11. Tétel (Birkhoff). *Egy A mátrix duplán sztochasztikus akkor és csak akkor, ha felírható permutáció mátrixok konvex kombinációjaként:*

$$A = t_1P_1 + t_2P_2 + \cdots + t_mP_m,$$

ahol $P_1, P_2, \ldots, P_m$ mátrixok permutáció mátrixok, valamint $t_1, t_2, \ldots, t_m$ nemnegatív számok és összegük 1.

4. fejezet

Egyenlőtlenségek pozitív (szemi)definit mátrixokra

Elsőként a Hadamard-egyenlőtlenséggel foglalkozunk. Ez ugyan csak közvetve vonatkozik a mátrixok sajátértékeire, mivel tudjuk, hogy egy mátrix determinánsa felírható sajátértékeinek szorzataként. Azonban mégis kiemelném, mivel motivációt szolgáltatott számos további mátrixegyenlőtlenséghez, amelyeket később, a 6. fejezetben tárgyalok. Ehhez a részhez az [1] valamint [10] forrásokat használtam fel.

4.1. Tétel (Hadamard-egyenlőtlenség). *Ha $A \in \mathbb{M}_n$ pozitív szemidefinit mátrix, akkor*

$$\det(A) \leq \prod_{i=1}^{n} a_{ii}.$$

Egyenlőség akkor és csak akkor teljesül, ha A diagonális mátrix.

Bizonyítás: Tegyük fel, hogy A minden diagonális elemére $a_{ii} > 0$ és legyen $D = \operatorname{diag}(a_{11}^{-1/2}, \dots, a_{nn}^{-1/2})$. Vegyük a $B = DAD$ mátrixot. Ekkor B pozitív szemidefinit és a diagonálisában 1-esek állnak. Legyenek $\lambda_1, \dots, \lambda_n$ a B sajátértékei. Felhasználjuk a számtani és mértani közép közötti egyenlőtlenséget, amely kimondja, hogy bármely nemnegatív $a_1, \dots, a_n \in \mathbb{R}$ valós számok esetén $\sqrt[n]{a_1 \cdot \dots \cdot a_n} \leq \frac{a_1 + \dots + a_n}{n}$. Ekkor:

$$n = \operatorname{tr}(B) = \sum_{i=1}^{n} \lambda_i \geq n \left(\prod_{i=1}^{n} \lambda_i \right)^{1/n} = n(\det(B))^{1/n}.$$

Ebből látszik, hogy $\det(B) \leq 1$. Így

$$\det(A) = \det(D^{-1}BD^{-1}) = \prod_{i=1}^{n} a_{ii} \det(B) \leq \prod_{I=1}^{n} a_{ii}.$$

Egyenlőség akkor és csak akkor áll fenn, ha B sajátértékei azonosak és $\det(B) = 1$. Tehát B az egységmátrix, A pedig diagonális. ∎

A következőkben Kantorovich és Wielandt tételeit tárgyaljuk, melyeknek első publikálásukat követően, számos általánosítása született és több új bizonyítása is napvilágot látott. Mi az eredeti egyenlőtlenségeket fogjuk átvenni és egyszerű bizonyításokat adunk.

4.2. Tétel (Kantorovich). *Legyen $A \in \mathbb{M}_n$ Hermite-féle pozitív definit mátrix és λ_1,λ_n az A-nak legnagyobb valamint legkisebb sajátértéke. Ekkor*

$$(x^*Ax)(x^*A^{-1}x) \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}(x^*x)^2$$

minden $x \in \mathbb{C}^n$ esetén.

Bizonyítás: Feltehetjük, hogy $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ és x egységvektor. Az egyenlőtlenség bal oldalán szereplő kvadratikus formákat összegalakban átírva, valamint felhasználva, hogy x egységvektor, így $(x^*x) = 1$, az alábbi egyenlőtlenséget kapjuk:

$$\sum_{j=1}^{n} \lambda_j x_j^2 \sum_{j=1}^{n} \frac{1}{\lambda_j} x_j^2 \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}.$$

Ezt kell tehát belátnunk. A bizonyítás alapötlete, hogy az alábbi nyilvánvaló egyenlőtlenséget vesszük kiindulási pontnak:

$$\sum_{i=1}^{n} x_i^2(\lambda_1 - \lambda_i) \geq \sum_{i=1}^{n} \frac{\lambda_n}{\lambda_i} x_i^2(\lambda_1 - \lambda_i).$$

Ebből

$$\lambda_1 - \sum_{i=1}^{n} \lambda_i x_i^2 \geq \lambda_1\lambda_n \sum_{i=1}^{n} \frac{1}{\lambda_i} x_i^2 - \lambda_n.$$

Átrendezve

$$\lambda_1 + \lambda_n \geq \lambda_1\lambda_n \left(\sum_{i=1}^{n} \frac{1}{\lambda_i} x_i^2 \right) + \sum_{i=1}^{n} \lambda_i x_i^2.$$

Az utolsó lépéshez felhasználjuk a számtani és mértani közép közötti egyenlőtlenséget. Így tehát

$$\lambda_1 + \lambda_n \geq 2\sqrt{\lambda_1\lambda_n}\sqrt{\left(\sum_{i=1}^{n}\frac{1}{\lambda_i}x_i^2\right)\left(\sum_{i=1}^{n}\lambda_i x_i^2\right)}$$

$$\frac{\lambda_1 + \lambda_n}{2\sqrt{\lambda_1\lambda_n}} \geq \sqrt{\left(\sum_{i=1}^{n}\frac{1}{\lambda_i}x_i^2\right)\left(\sum_{i=1}^{n}\lambda_i x_i^2\right)}.$$

Az így kapott egyenlőtlenséget négyzetre emelve megkapjuk az eredeti állítást.

■

4.1. Lemma. *Legyen* $\mathbf{A} = \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix}$ *nemnulla* 2×2*-es pozitív szemidefinit mátrix,* α *és* β *sajátértékekkel, melyekre* $\alpha > \beta$*. Ekkor*

$$|b|^2 \leq \left(\frac{\alpha - \beta}{\alpha + \beta}\right)^2 ac.$$

Egyenlőség akkor és csak akkor áll fenn, ha $a = c$ *vagy* $\beta = 0$*, azaz* $\mathbf{A}$ *szinguláris.*

Bizonyítás: Oldjuk meg a $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ egyenletet, melyből megkapjuk a sajátértékeket:

$$\alpha,\beta = \frac{(a+b) \pm \sqrt{(a - c^2 + 4|b|^2}}{2}.$$

Számítsuk ki $\frac{\alpha-\beta}{\alpha+\beta}$ értékét és emeljünk négyzetre. Így a tételben kimondott egyenlőtlenséggel ekvivalens formulát kapunk,

$$(a-c)^2(ac - |b|^2) \geq 0.$$

Az állítás ebből azonnal következik. ■

4.3. Tétel (Wielandt). *Legyen* $A \in \mathbb{M}_n$ *Hertmite-féle, pozitív definit mátrix és* λ_1,λ_n *az* A*-nak legnagyobb valamint legkisebb sajátértéke. Ekkor bármely ortogonális* $x, y \in \mathbb{C}^n$ *esetén*

$$|x^*Ay|^2 \geq \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 (x^*Ax)(y^*Ay).$$

Bizonyítás: Legyen $M = (x,y)^* A(x,y)$. M pozitív szemidefinit és

$$M = \begin{pmatrix} x^*Ax & x^*Ay \\ y^*Ax & y^*Ay \end{pmatrix}.$$

Tudjuk, hogy $\lambda_n I \leq A \leq \lambda_1 I$. Feltehetjük, hogy x és y ortonormált vektorok, így $(x,y)^*(x,y) = I_2$. Az egyenlőtlenséget szorozzuk balról $(x,y)^*$-al, jobbról (x,y)-al. Ekkor $\lambda_n I \leq M \leq \lambda_1 I$. Legyenek α és β a 2×2-es M mátrix sajátértékei, $\alpha \geq \beta$, $\alpha > 0$. Ekkor $\lambda_n \leq \beta \leq \alpha \leq \lambda_1$. Alkalmazzuk a lemmát:

$$|x^*Ay|^2 \geq \left(\frac{\alpha - \beta}{\alpha + \beta}\right)^2 (x^*Ax)(y^*Ay).$$

Felhasználjuk a tényt, hogy $\frac{\alpha-\beta}{\alpha+\beta} \leq \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}$. Ezzel beláttuk az állítást. ■

5. fejezet

Egyenlőtlenségek normális mátrixokra

Ez a fejezet az [1],[2],[8],[13] források felhasználásával készült.
Elsőként Schur tételét tárgyaljuk, mely a mátrixok sajátértékei és diagonális elemei közötti összefüggésre mutat rá. Ehhez szükségünk lesz a Schur-féle felbontásra.

5.1. Tétel (Schur-féle felbontás). *Legyenek $\lambda_1, \ldots, \lambda_n$ az $A \in \mathbb{M}_n$ mátrix sajátértékei. Ekkor létezik egy $U \in \mathbb{M}_n$ unitér mátrix, hogy $T = U^*AU$ egy felsőháromszög mátrix:*

$$U^*AU = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

A pontosan akkor normális, ha T diagonális. Az $A = UTU^*$ alakot nevezzük az A mátrix Schur-felbontásának.

5.2. Tétel (Schur). *Legyen $A = (a_{ij})$ $n \times n$-es komplex mátrix $\lambda_1, \ldots, \lambda_n$ sajátértékekkel. Ekkor*

$$\sum_{i=1}^{n} |\lambda_i|^2 \leq \sum_{i,j=1}^{n} |a_{i,j}|^2.$$

Egyenlőség akkor és csak akkor áll fenn, ha A normális mátrix.

Bizonyítás: Legyen $A = U^*TU$ az A mátrix Schur-felbontása, ahol U unitér és T felsőháromszög mátrix. Ekkor $A^*A = U^*T^*TU$, mivel $A^* = U^*T^*U$ és az unitér mátrix definíciójából adódóan $U^*U = I$. A mátrixszorzás definíciójából tudjuk, hogy

$$\mathrm{tr}(A^*A) = \sum_{i,j=1}^{n} |a_{ij}|^2$$

és

$$\mathrm{tr}(T^*T) = \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i<j} |t_{ij}|^2.$$

Ebből az egyenlőtlenség azonnal következik. Vegyük észre, hogy egyenlőség esetén $t_{ij} = 0$, $i < j$, azaz T diagonális. A Schur-féle felbontás szerint pedig A pontosan akkor normális, ha T diagonális. ■

A következő két egyenlőtlenséghez szükségünk lesz néhány további fogalom definiálására.

5.3. Tétel (Spektrál-felbontás). *Legyen A egy $n \times n$-es komplex normális mátrix $\lambda_1, \ldots, \lambda_n$ sajátértékekkel. Ekkor létezik olyan U unitér mátrix, hogy*

$$U^*AU = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Az $A = U\mathrm{diag}(\lambda_1, \ldots, \lambda_n)U^*$ alakot nevezzük A spektrál-felbontásánaknak.

5.1. Definíció. *Egy $\|\cdot\| : \mathbb{M}_n \to \mathbb{R}$ függvényt mátrixnormának nevezünk, ha minden $A, B \in \mathbb{M}_n$ mátrixra és $c \in \mathbb{C}$ számra teljesülnek az alábbi feltételek:*

- $\|A\| \geq 0$
- $\|A\| = 0$ *akkor és csak akkor, ha* $A = 0$
- $\|cA\| \leq |c|\|A\|$ *és*
- $\|A + B\| \leq \|A\| + \|B\|$.

Példa:

- Oszlopnorma: $\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|$.
- Sornorma: $\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|$.

- Schatten-p-norma: $\|A\|_p = (\sum_{i=1}^n \sigma_i^p)^{(1\backslash p)}$, ahol $\sigma_1, \dots, \sigma_n$ az A mátrix szinguláris értékei.

5.2. Definíció. *Legyen A $n \times n$-es komplex mátrix, $\sigma_1, \dots, \sigma_n$ pedig az A mátrix szinguláris értékei. Ekkor A Frobenius-normáját a következőképpen definiáljuk.*

$$\|A\|_F = \left(\sum_{i,j=1}^n |a_{ij}|^2\right)^{(1\backslash 2)} = (\mathrm{tr}(A^*A))^{(1\backslash 2)} = \left(\sum_{i=1}^n \sigma_i^2\right)^{(1\backslash 2)}$$

ahol $\sigma_1, \dots, \sigma_n$ az A szinguláris értékei. Az $\|A\|_F$ helyett az $\|A\|_2$ jelölést is szokás használni.

5.3. Definíció. *Egy $\|\cdot\| : \mathbb{M}_n \to \mathbb{R}$ mátrixnormát unitér-invariánsnak mondunk, ha tetszőleges $U, V \in \mathbb{M}_n$ unitér mátrixokkal*

$$\|UAV\| = \|A\|.$$

Könnyen belátható, hogy a Frobenius-norma unitér invariáns, azaz ha $A \in \mathbb{M}_n$ és $U \in \mathbb{M}_n$, ahol U unitér mátrix, akkor $\|AU\|_2 = \|UA\|_2 = \|A\|_2$. Hiszen

$$\|AU\|_2^2 = \mathrm{tr}((AU)^*AU) = \mathrm{tr}(A^*(UU^*)A) = \mathrm{tr}(A^*A) = \|A\|_2^2.$$

Az $\|UA\|_2$ eset hasonlóképpen bizonyítható.

5.4. Definíció. *Két azonos méretű A, B mátrix Hadamard-szorzata az az $A \circ B$-vel jelölt mátrix, melynek elemeit az A és B mátrixok megfelelő elemeinek szorzataként kapjuk:*

$$A \circ B = (a_{ij}b_{ij}).$$

5.4. Tétel (Hoffman-Wielandt). *Legyenek A és B $n \times n$-es normális mátrixok $\lambda_1, \dots, \lambda_n$ és $\mu_1, \dots, \mu_n$ megfelelő sajátértékekkel. Ekkor létezik az $\{1, 2, \dots, n\}$ halmazon olyan p permutáció, melyre*

$$\left(\sum_{i=1}^n |\lambda_i - \mu_{p(i)}|^2\right)^{1/2} \leq \|A - B\|_2.$$

Bizonyítás: Írjuk fel A-t és B-t a spektrál-felbontás szerint

$$A = U^*\mathrm{diag}(\lambda_1, \ldots, \lambda_n)U \quad \text{és} \quad B = V^*\mathrm{diag}(\mu_1, \ldots, \mu_n)V$$

alakban, ahol U és V unitér mátrixok. Használjuk az $E = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $F = \mathrm{diag}(\mu_1, \ldots, \mu_n)$ valamint $W = (w_{ij}) = UV^*$ jelöléseket. Ekkor

$$\|A - B\|_2^2 = \|U^*EU - V^*FV\|_2^2$$

$$\|A - B\|_2^2 = \|U^*(EUV^* - UV^*F)V\|_2^2.$$

Mivel a Frobenius-norma unitér-invariáns,

$$\|A - B\|_2^2 = \|EW - WF\|_2^2.$$

Elvégezve az $EW - WF$ mátrixműveletet olyan mátrixot kapunk melynek i-edik sorában és j-edik oszlopában $(\lambda_i - \mu_j)w_{ij}$ szerepel. Így tehát a Frobenius norma definíciója alapján felírhatjuk, hogy

$$\|A - B\|_2^2 = \sum_{i,j=1}^{n} |\lambda_i - \mu_j|^2 |w_{ij}|^2.$$

Legyenek $G = (|\lambda_i - \mu_j|^2)$ és $S = (|w_{ij}|^2)$. Ekkor

$$\|A - B\|_2^2 = \sum_{i,j=1}^{n} |\lambda_i - \mu_j|^2 |w_{ij}|^2 = e^T(G \circ S)e,$$

ahol $G \circ S$ a G és S mátrixok Hadamard-szorzata, e pedig n-hosszú 1-esekből álló oszlopvektor.

S duplán sztochasztikus mátrix. Birkhoff tétele szerint S felírható permutáció mátrixok konvex kombinációjaként:

$$S = \sum_{i=1}^{m} t_i P_i.$$

Tegyük fel, hogy az $e^T(G \circ P_i)e$, $i = 1, \ldots, m$ értékek között $i = k$-ra kapjuk a legkisebb értéket, azaz $e^T(G \circ P_k)e$ mind közül a legkisebb. Tekintsük P_k-t az $\{1, 2, \ldots, n\}$ halmazon vett p permutációnak. Ekkor

$$\|A - B\|_2^2 = e^T(G \circ S)e = \sum_{i=1}^{m} t_i e^T(G \circ P_i)e$$

$$\geq \sum_{i=1}^{m} t_i e^T (G \circ P_k) e$$

$$= e^T (G \circ P_k) e = \sum_{i=1}^{n} |\lambda_i - \mu_{p(i)}|^2.$$

■

A Hoffman-Wielandt tételben mindkét mátrix, A illetve B is normálisak voltak. A következőkben megmutatjuk, miként módosul az állítás, ha csak az egyik mátrix normális, míg a másik tetszőleges. Ehhez bevezetünk egy új jelölést, valamint kimondunk egy lemmát. Ha A kvadratikus mátrix, akkor felírhatjuk $A = U_A + D_A + L_A$ alakban, ahol U_A,D_A,L_A rendre az A főátló feletti elemeit, diagonális elemeit valamint főátló alatti elemeit jelentik.

Példa: $A = \begin{pmatrix} 1 & 2 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 5 & 0 \end{pmatrix}$

5.1. Lemma. *Legyen A $n \times n$-es normális mátrix. Ekkor*

$$\|U_A\|_2 \leq \sqrt{n-1}\|L_A\|_2, \quad \|L_A\|_2 \leq \sqrt{n-1}\|U_A\|_2.$$

5.5. Tétel (Sun). *Legyen $A \in \mathbb{M}_n$ $n \times n$-es normális mátrix $\lambda_1, \ldots, \lambda_n$ sajátértékekkel és $B \in \mathbb{M}_n$ $n \times n$-es mátrix $\mu_1, \ldots, \mu_n$ sajátértékekkel. Ekkor létezik az $\{1, 2, \ldots, n\}$ halmazon olyan p permutáció, mellyel*

$$\left(\sum_{i=1}^{n} |\lambda_i - \mu_{p(i)}|^2 \right)^{1/2} \leq \sqrt{n}\|A - B\|_2.$$

Bizonyítás: A Schur-felbontás szerint létezik olyan U unitér mátrix, hogy U^*BU felső háromszögmátrix. Tegyük fel, hogy B már felső háromszögmátrix. Így $D_B = \mathrm{diag}(\mu_1, \ldots, \mu_n)$. Legyen $C = A - B$. Ekkor

$$A - D_B = C + U_B, \quad U_B = U_A - U_C, \quad L_A = L_C.$$

Mivel ebben az esetben A és D_B is normális mátrixok, a Hoffman-Wielandt tétel szerint létezik olyan p permutáció az $\{1, 2, \ldots, n\}$ halmazon, hogy

$$\left(\sum_{i=1}^{n} |\lambda_i - \mu_{p(i)}|^2 \right)^{1/2} \leq \|A - D_B\|_2 = \|C + U_B\|_2.$$

Most alkalmazzuk a lemmát.

$$\|C + U_B\|_2^2 = \|C + U_A - U_C\|_2^2 =$$

$$= \|L_C + D_C + U_A\|_2^2$$

$$= \|L_C\|_2^2 + \|D_C\|_2^2 + \|U_A\|_2^2$$

$$\leq \|L_C\|_2^2 + \|D_C\|_2^2 + (n-1)\|L_A\|_2^2$$

$$= \|L_C\|_2^2 + \|D_C\|_2^2 + (n-1)\|L_C\|_2^2$$

$$\leq n\|C\|_2^2 = n\|A - B\|_2^2.$$

Mindkét oldalból négyzetgyököt vonva megkapjuk a bizonyítandó állítást. ■

6. fejezet

Majorizáció és mátrix egyenlőtlenségek

Ebben a fejezetben az [1],[7],[10],[11],[12],[13] és [14] forrásokat használtam fel.
Az első majorizációval kapcsolatos mátrixelméleti eredmény Schur nevéhez köthető. 1923-ban napvilágot látott tétele a hermitikus mátrixok diagonális elemeinek és sajátértékeinek összehasonlításán alapszik. Motivációját a pozitív szemidefinit mátrixok determinánsára, valamint diagonális elemeinek szorzatára vonatkozó Hadamard-egyenlőtlenség adta. Felfedezése óta számos majorziációs összefüggés született a mátrixelméletben. Ezek közül csupán néhány klasszikus eredményt tárgyalunk.

Vegyünk egy $x \in \mathbb{R}^n$, $x = (x_1, x_2, \dots, x_n)$ vektort. Jelöljük $x^{\downarrow}$-al azt a vektort, amelynek elemeit rendre az x nemnövekvő sorba rendezett elemei adják. Tehát

$$x^{\downarrow} = (x_1^{\downarrow}, x_2^{\downarrow}, \dots, x_n^{\downarrow}) \quad \text{és} \quad x_1^{\downarrow} \geq x_2^{\downarrow} \geq \cdots \geq x_n^{\downarrow}.$$

Ennek segítségével definiáljuk két valós vektor között a majorizációt.

6.1. Definíció. *Legyenek $x, y \in \mathbb{R}^n$. Ekkor azt mondjuk, hogy y gyengén majorálja x-et és $x \prec_w y$-al jelöljük, ha*

$$\sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow}$$

minden $k = 1, 2, \ldots, n$ *esetén. Ha ezenfelül teljesül az is, hogy*

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i,$$

akkor azt mondjuk, hogy y *majorálja* x*-et és* $x \prec y$*-al jelöljük.*

Példa:

- $x = (5, 1, 2), y = (0, 8, 0) \quad \Rightarrow \quad x^{\downarrow} = (5, 2, 1), y^{\downarrow} = (8, 0, 0)$
  Ekkor y majorálja x-et, $x \prec y$.

- $x = (5, 1, 2), y = (0, 9, 0) \quad \Rightarrow \quad x^{\downarrow} = (5, 2, 1), y^{\downarrow} = (9, 0, 0)$
  Ekkor y gyengén majorálja x-et, $x \prec_w y$.

A definícióból nyilvánvalóan következik, hogy ha $x \prec y$, akkor $x \prec_w y$. Továbbá azt is egyértelműen megállapíthatjuk, hogy a majorizáció szempontjából egy vektor elemeinek eredeti sorrendje nem számít. Tehát ha egy tetszőleges z vektor ugyanazon elemekből áll mint x, és tudjuk hogy $x \prec y$, akkor $z \prec y$ is igaz.

A későbbiekben kimondott tételek kizárólag Hermite-féle mátrixokra vonatkoznak, melyekről tudjuk, hogy diagonális elemeik, szinguláris- és sajátértékeik valósak. Így ha bevezetjük az alábbi jelöléseket $A \in \mathbb{M}_n$-re :

$$d(A) = (a_{11}, a_{22}, \ldots, a_{nn}),$$

$$\lambda(A) = (\lambda_1, \lambda_2, \ldots, \lambda_n),$$

$$\sigma(A) = (\sigma_1, \sigma_2, \ldots, \sigma_n),$$

a majorizáció értelmezhetővé válik az A mátrix a_{ii}, λ_i és σ_i értékei között.

6.1. Állítás. *Legyenek* $x, y \in \mathbb{R}^n$*. Ekkor* $x \prec y$ *akkor és csak akkor, ha létezik olyan duplán sztochasztikus* $P \in \mathbb{M}_n$ *mátrix, melyre* $x = Py$*. Valamint* $x \prec_w y$ *akkor és csak akkor, ha létezik olyan duplán sztochasztikus* $P \in \mathbb{M}_n$ *mátrix, melyre* $x = Py$*.*

6.1. Tétel (Schur). *Legyen $A \in \mathbb{M}_n$ Hermite-féle mátrix, $a_{11}, \ldots, a_{nn}$ diagonális elemekkel és $\lambda_1, \ldots, \lambda_n$ sajátértékekkel. Ekkor*

$$d(A) \prec \lambda(A).$$

Bizonyítás: A spektrál-felbontás tétele miatt, valamint mivel tudjuk, hogy minden Hermite-féle mátrix normális, létezik olyan U unitér mátrix, hogy $A = U\mathrm{diag}(\lambda_1, \ldots, \lambda_n)U^*$. Az A diagonális elemeire a következőt írhatjuk fel:

$$a_{ii} = \sum_{j=1}^{n} u_{ij}\bar{u}_{ij}\lambda_j = \sum_{j=1}^{n} p_{ij}\lambda_j, \quad i = 1, 2, \ldots, n,$$

ahol $p_{ij} = u_{ij}\bar{u}_{ij}$. Mivel U unitér mátrix, $P = (p_{ij})$ duplán sztochasztikus. Ez nagyon könnyen megmutatható, mindössze az unitér és duplán sztochasztikus mátrixok definícióit kell meggondolni. Ezek után felírhatjuk, hogy

$$(a_{11}, a_{22}, \ldots, a_{nn}) = (\lambda_1, \lambda_2, \ldots, \lambda_n)P.$$

6.1 állítás miatt megkapjuk az eredeti állítást. ■

6.2. Tétel (Fan). *Legyen $A \in \mathbb{M}_n$ Hermite-féle mátrix. Jelöljünk S_k-val k darab tetszőleges ortonormált $x_1, x_2, \ldots, x_k \in \mathbb{C}^n$ vektort. Ekkor*

$$\sum_{i=1}^{k} \lambda_i = \max_{S_k} \sum_{i=1}^{k} x_i^* A x_i, \quad 1 \leq k \leq n.$$

6.3. Tétel (Fan). *Legyenek $A, B \in \mathbb{M}_n$ Hermite-féle mátrixok. Ekkor*

$$\lambda(A + B) \prec \lambda(A) + \lambda(B).$$

Bizonyítás: Használjuk az S_k jelölést 6.2-ben leírtak szerint. Ekkor 6.2 alapján látható, hogy

$$\lambda(A + B) \prec_w \lambda(A) + \lambda(B)$$

minden $k \leq n$-re ekvivalens a következővel:

$$\max_{S_k} \sum_{i=1}^{k} x_i^*(A + B)x_i \leq \max_{S_k} \sum_{i=1}^{k} x_i^* A x_i + \max_{S_k} \sum_{i=1}^{k} x_i^* B x_i.$$

Tudjuk továbbá, hogy a mátrix nyoma egyenlő a sajátértékeinek összegével, valamint hogy

$$\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B).$$

Így $\lambda(A) + \lambda(B)$ valóban majorálja $\lambda(A+B)$-t. ∎

6.4. Tétel (Thompson). *Legyenek $A, B \in \mathbb{M}_n$ Hermite-féle mátrixok, $C = A + B$. Ha az A, B, C mártixok sajátértékei rendre $\alpha_1 \geq \cdots \geq \alpha_n$, $\beta_1 \geq \cdots \geq \beta_n$, $\gamma_1 \geq \cdots \geq \gamma_n$, akkor minden $1 \leq i_1 < \cdots < i_k \leq n$ sorozatra*

$$\sum_{t=1}^{k} \alpha_{i_t} + \sum_{t=1}^{k} \beta_{n-k+t} \leq \sum_{t=1}^{k} \gamma_{i_t} \leq \sum_{t=1}^{k} \alpha_{i_t} + \sum_{t=1}^{k} \beta_t.$$

6.5. Tétel (Lidskii). *Legyenek $A, B \in \mathbb{M}_n$ Hermite-féle mátrixok. Ekkor*

$$\lambda(A) - \lambda(B) \prec \lambda(A - B).$$

Bizonyítás: Írjuk fel az A mátrixot $A = B + (A - B)$ alakban. Erre alkalmazva Thompson tételét:

$$\sum_{t=1}^{k} \gamma_{i_t} \leq \sum_{t=1}^{k} \alpha_{i_t} + \sum_{j=1}^{k} \beta_j \quad k = 1, \ldots, n,$$

ahol $\alpha_1 \geq \cdots \geq \alpha_n$, $\beta_1 \geq \cdots \geq \beta_n$ és $\gamma_1 \geq \cdots \geq \gamma_n$ rendre a B, $A - B$ és A mátrixok sajátértékei. Ebből

$$\max_{1 \leq i_1 < \cdots < i_k \leq n} \sum_{t=1}^{k} (\gamma_{i_t} - \alpha_{i_t}) \leq \sum_{j=1}^{k} \beta_j,$$

és így

$$\lambda(A) - \lambda(B) \prec_w \lambda(A - B).$$

Egyenlőség $k = n$ esetén áll fenn, ebből következik a majorizáció. ∎

Végül nézzünk egy hermitikus blokkmátrixra vonatkozó majorizációs egyenlőtlenséget. Ehhez jelöljük el az A és B mátrixok *direkt összegét* az alábbi módon:

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

6.6. Tétel. *Egy $A \in \mathbb{M}_n$ mátrix hermitikus akkor és csak akkor, ha létezik U unitér mátrix, hogy*

$$A = U^* \mathrm{diag}(\lambda_1, \ldots, \lambda_n) U,$$

ahol λ_i-k valós számok, és egyben A sajátértékei is.

Ez a spektrál-felbontás hermitikus mátrixokra vonatkozó esete.

6.7. Tétel (Fan). *Ha A és $A_{11} \oplus A_{22}$ mátrixok elemei M_n-nek, hermitikusak*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad A_{11} \oplus A_{22} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix},$$

ahol A_{11} $l \times l$-es, A_{22} $m \times m$-es, $l + m = n$, akkor

$$\lambda(A_{11} \oplus A_{22}) \prec \lambda(A).$$

Bizonyítás: Legyenek $\alpha_1, \ldots, \alpha_n$ és $\beta_1, \ldots, \beta_n$ rendre az A_{11} és A_{22} mátrixok sajátértékei. A 6.6 tétel alapján létezik egy $l \times l$-es U_1 és egy $m \times m$-es U_2 unitér mátrix, hogy

$$U_1 A_{11} U_1^* = D_\alpha, \quad U_1 A_{22} U_1^* = D_\beta,$$

ahol D_α és D_β az α_i és β_i sajátértékeket tartalmazó diagonális mátrixok. Vegyük az $U = U_1 \oplus U_2$ $n \times n$-es mátrixot. Ekkor láthatjuk, hogy

$$U(A_{11} \oplus A_{22})U^* = \begin{pmatrix} D_\alpha & 0 \\ 0 & D_\beta \end{pmatrix}, \quad UAU^* = \begin{pmatrix} D_\alpha & M \\ M^* & D_\beta \end{pmatrix},$$

ahol $M = U_1 A_{12} U_2^*$. Mivel U unitér mátrix, $\lambda(A_{11} \oplus A_{22}) = \lambda(U(A_{11} \oplus A_{22})U^*)$ és $\lambda(A) = \lambda(UAU^*)$, és így 6.1 tétel szerint

$$\lambda(A_{11} \oplus A_{22}) = \lambda(U(A_{11} \oplus A_{22})U^*) = \lambda(A_{11} \oplus A_{22}) \prec \lambda(A) = \lambda(UAU^*) = \lambda(A).$$

■

Irodalomjegyzék

[1] Fuzhen Zhang: *Matrix Theory (Basic Results and Techniques)*, Springer, 2011

[2] V.V.Praszolov: *Lineáris algebra*, Typotex Kiadó, 2005

[3] Rózsa Pál: *Lineáris algebra és alkalmazásai*, Műszaki Könyvkiadó, 1976

[4] Peter D. Lax: *Lineáris algebra és alkalmazásai*, Akadémia Kiadó, 2008

[5] Miroslav Fiedler: *Special matrices and their applications in numerical mathematics*, Martinus Nijhoff Publishers, 1986

[6] Wettl Ferenc: *Lineáris algebra*, Typotex Kiadó, 2011

[7] Xingzhi Zhan: *Matrix inequalities*, Springer, 2002

[8] Massoud Malek jegyzete: *Inequalities concerning eigenvalues*, California State University, East Bay, http://www.mcs.csueastbay.edu/ malek/Class/Eigenvalue.pdf

[9] Faragó István, Horváth Róbert: *Numerikus módszerek*, Typotex kiadó, 2011

[10] Minghua Lin: *Angles, Majorization, Wielandt Inequality and Applications*, Ph.D. Thesis, 2013

[11] Dennis S. Bernstein: *Matrix mathematics (theory, facts and formulas)*, Princeton University Press, 2009

[12] Albert W. Marshall, Ingram Olkin, Barry C. Arnold: *Inequalities: Theory of Majorization and Its Applications*, Springer, 2011

[13] Denis Serre: *Matrices: Theory and Applications*, Springer, 2002

[14] http://en.wikipedia.org/wiki/Majorization

[15] http://www.uni-miskolc.hu/evml/database/downloads/linalg/segedletek /num23.pdf

Köszönetnyilvánítás

Ezúton szeretném megköszönni Fialowski Alice Tanárnőnek a támogatást, a téma feldolgozásában való segítséget, valamint hasznos ötleteit és tanácsait, melyekkel hozzájárult szakdolgozatom elkészüléséhez.