A KULLBACK-LEIBLER RELATÍV ENTRÓPIA FÜGGVÉNY ALKALMAZÁSA PÁROS ÖSSZEHASONLÍTÁS MÁTRIX EGY PRIORITÁSVEKTORA MEGHATÁROZÁSÁRA[1]

KOMÁROMI ÉVA
Budapesti Corvinus Egyetem

A dolgozatban a döntéselméletben fontos szerepet játszó páros összehasonlítás mátrix prioritásvektorának meghatározására új megközelítést alkalmazunk. Az A páros összehasonlítás mátrix és a prioritásvektor által definiált B konzisztens mátrix közötti eltérést a Kullback-Leibler relatív entrópia-függvény segítségével mérjük. Ezen eltérés minimalizálása teljesen kitöltött mátrix esetében konvex programozási feladathoz vezet, nem teljesen kitöltött mátrix esetében pedig egy fixpont problémához. Az eltérésfüggvényt minimalizáló prioritásvektor egyben azzal a tulajdonsággal is rendelkezik, hogy az A mátrix elemeinek összege és a B mátrix elemeinek összege közötti különbség éppen az eltérésfüggvény minimumának az n-szerese, ahol n a feladat mérete. Így az eltérésfüggvény minimumának értéke két szempontból is lehet alkalmas az A mátrix inkonzisztenciájának a mérésére.

Kulcsszavak: AHP, páros összehasonlítás mátrix, többszempontú döntések, Kullback-Leibler relatív entrópia, eltérésfüggvények, konvex program.

1 Bevezetés

Az AHP (Analytic Hierarchy Process) a többszempontú döntési problémák kezelésére alkalmas eljárás. Kulcseleme a páronkénti összehasonlítások – másként: páros összehasonlítások – használata. A páros összehasonlítás a döntéshozó véleményének numerikus reprezentációja az egyes döntési alternatívák fontosságáról minden egyes másik döntési alternatívához viszonyítva egy adott kritérium szempontjából: a_{ij} azt mutatja meg, hányszor fontosabb az i-edik alternatíva a j-ediknél. Nyilvánvaló, hogy

$$a_{ij} > 0\,, \quad a_{ii} = 1\,, \quad a_{ij} = \frac{1}{a_{ji}}\,; \quad i,j = 1,\ldots,n\,,$$

ahol n a döntési alternatívák száma. Az AHP alkalmazása során az egyes alternatívákhoz prioritásokat határozunk meg – ezeket az irodalomban gyakran preferencia értékeknek, vagy súlyoknak is nevezik. Ha a döntéshozó a fontosságok megítélésében konzisztens lenne, vagyis ha $a_{ij} = a_{ik}a_{kj}$ lenne minden i,j,k indexhármas esetén, akkor az n alternatívához tartozó prioritásokat

[1]Beérkezett: 2013. március 13. E-mail: `komaromi@uni-corvinus.hu`.

(preferencia értékeket, súlyokat) könnyen megkaphatnánk. De a döntéshozók megítélése rendszerint csökkenően konzisztens az alternatívák számának növekedésével. A kérdés ez: hogyan rendeljünk az egyes alternatívákhoz prioritásokat – az n alternatívához prioritásvektort – ha a megadott $A = \{a_{ij}\} \in R^{n \times n}$ páros összehasonlítás mátrix nem konzisztens? Az alapelképzelés az, hogy olyan $p = (p_1, \ldots, p_n)$ vektor alkalmas a döntéshozó prioritásainak a kifejezésére, amely esetében az A mátrix a $B = \{p_i/p_j\}$ konzisztens mátrixtól lehetőleg kevéssé tér el.

Az AHP-vel jelzett kutatási terület és módszertan Thomas L. Saaty nevéhez fűződik [1980], aki számos cikket, könyvet szentelt a sajátérték módszer (EM) kifejlesztésére és vizsgálatára – közülük egyet emelünk itt ki, amelyben a szerző a módszer legfontosabb tulajdonságait foglalja össze [2003]. E módszer lényege az, hogy a prioritásvektort így kapjuk: $p = \arg\{Ap = \lambda_{\max} p\}$, ahol $\lambda_{\max}$ az A mátrix legnagyobb sajátértéke. Mivel $\lambda_{\max}$ értéke n, ha a mátrix konzisztens, és nagyobb n-nél egyébként, ezért az inkonzisztencia mérőszámául Saaty a $(\lambda_{\max} - n)/(n-1)$ értéket, illetve ennek egy tapasztalati konstanssal vett szorzatát választotta. A módszer tanulmányozására iskola alakult, amelynek számos jeles képviselője közül itt hivatkozunk P. T. Harkerre és két dolgozatára [1987 és 1987a]. A másodikban Harker a nem teljesen kitöltött páros összehasonlítás mátrix esetével foglalkozik. Megjegyezzük, hogy a hiányzó elemek kitöltése érdekében a célprogramozás illetve a fuzzy preferencia relációk alkalmazására is vannak törekvések, ld. M. Fedrizzi és S. Giove dolgozatát [2007].

Hiányos mátrixok meglétekor több féle kérdés vetődik fel. Az egyik az, hogy szükséges-e a döntéshozót tovább kérdezgetni nézetéről az alternatívák fontosságát illetően – pl. ha a döntéshozó csak a mátrix első sorát töltötte ki, akkor ezzel megadta a prioritásait, amelyekkel a hiányzó elemeket már úgy tudjuk kitölteni, hogy konzisztens mátrixot kapjunk. Ez azonban nincs összhangban azzal a tapasztalattal, hogy a döntéshozók döntései általában nem konzisztensek, ezért néhány további elemre is szükség van, hogy mélyebb betekintést kapjunk a döntéshozónak az alternatívák fontosságát illető véleményébe. Ha a mátrix lényegesen több elemét tölti ki a döntéshozó, akkor szembesül a véleményeiből fakadó inkonzisztenciával, így hajlik korábbi véleményei felülvizsgálatára. Ezek azonban olyan ,,puha" kérdések, amelyekkel e dolgozat nem foglalkozik.

A módszerek másik csoportját olyan ,,távolságminimalizálási" eljárások alkotják, amelyekben a $p = (p_1, \ldots, p_n)$ prioritásvektorra kikötjük, hogy az A mátrix a $B = \{p_i/p_j\}$ konzisztens mátrixhoz a ,,legközelebbi" legyen. A ,,legközelebbi" szó jelentésétől függ az alkalmazott módszer, különböző módszerek rendszerint különböző p vektorokhoz vezetnek. E módszerek közül a legismertebbek:

- $p \in \underset{\sum_i p_i = 1, p > 0}{\arg\min} \sum_i \sum_j (a_{ij} - p_i/p_j)^2$ – legkisebb négyzetek módszere (LSM);

- $p \in \underset{\sum_i p_i=1, p>0}{\arg\min} \sum_i \sum_j (a_{ij}p_j - p_i)^2$ – súlyozott legkisebb négyzetek módszere;
- $p \in \underset{\sum_i p_i=1, p>0}{\arg\min} \sum_i \sum_j (a_{ij} - p_i/p_j)^2 \frac{p_j}{p_i}$ – chi-négyzet módszer;
- $p \in \underset{\sum_i p_i=1, p>0}{\arg\min} \sum_i \sum_j (\ln a_{ij} - \ln p_i + \ln p_j)^2$ – logaritmikus legkisebb négyzetek módszere (LLSM).

S. I. Gass és T. Rapcsák [2004] a fenti két irányzatot összekapcsolta, J. Krovák [1987] numerikusan összehasonlította az EM, az LSM és LLSM módszereket. E. U. Choo és W. C. Wedley e módszerekről szóló összefoglaló cikkére [2004], illetve A. Ishizaka és A. Labibnak az AHP új fejleményeit bemutató cikkére [2011] hívjuk fel itt a figyelmet, az utóbbit ajánlva azon érdeklődő olvasónak, aki az AHP hatalmas irodalmáról áttekintést akar szerezni. Megemlítjük, hogy működik az AHP iskola magyar kutatói csoportja is, mindenekelőtt a SZTAKI Operációkutatás Laboratóriumában és a Corvinus Egyetem Operációkutatás Tanszékén, Rapcsák Tamás és Temesi József kezdeményezésére alakult. Tagjai jelentős publikációkkal járultak hozzá az AHP módszertanának a fejlesztéséhez, közöttük Bozóki S., Fülöp J. és Rónyai L. [2010], Csató L. [2012], Bozóki S., Fülöp J., Poesz A. [2011] cikkei az EM és LLSM módszerek alkalmazásával a nem teljesen kitöltött páros összehasonlítás mátrixok optimális kiegészítéséről közöl eredményeket, Temesi J. cikke [2010] és Farkas A., P. Lancaster, Rózsa P. dolgozata [2003] a páros összehasonlítás mátrixok konzisztenciáját tanulmányozta, Kéri Gerzson munkája [2005] a páros összehasonlítás mátrixokra fogalmaz meg kritériumokat.

A jelen dolgozat a páros összehasonlítás mátrix egy prioritásvektorának meghatározására olyan módszert alkalmaz, amely a „távolságminimalizálási" eljárások csoportjába tartozik. Távolság helyett azonban eltérést kell mondanunk, mert az alkalmazott függvény nem teljesíti a távolságfüggvények minden tulajdonságát. A következő részben bemutatunk két eltérésfüggvénycsaládot, amelyek tagjai alkalmasak lehetnek arra, hogy ezekkel mérjük az A páros összehasonlítás mátrix és a prioritásvektorból képzett $B = \{p_i/p_j\}$ konzisztens mátrix eltérését, közelségét. A harmadik részben az A és B mátrixok eltérését a p_i és $a_{ij}p_j$ értékeknek az információelméletben, statisztikában és számos más területen alkalmazott Kullback-Leibler relatív entrópia-függvény alkalmazásával kapott eltéréseinek összegeként fogjuk fel (erre a továbbiakban A és B Kullback-Leibler eltéréseként hivatkozunk), bemutatjuk az így származtatott (K-L) minimalizálási modellt, amelynek a tulajdonságait tanulmányozzuk: bizonyítjuk a célfüggvény konvexitását, és hogy a modellnek van egyetlen pozitív megoldása. A negyedik részben a célfüggvénynek a modell paramétereitől való függését vizsgáljuk és egyben a konzisztencia növelésének módjára teszünk javaslatot. Az ötödik részben algoritmust javaslunk a feladat megoldására, felhasználva a feladathoz tartozó Kuhn-Tucker stacionárius pont problémát alkotó egyenletrendszert; bemutatjuk, hogy a feladat egyetlen p optimális megoldásához tartozó célfüggvény-érték, amely az A mátrix és az optimális p vektor által meghatározott B mátrix Kullback-Leibler eltérése, egyenlő a két mátrix elemeinek összegei közötti különbség n-ed részével is (n a feladat mérete). A hatodik részben vizsgáljuk a nem teljesen kitöltött páros összehasonlítás mátrix esetén kapott feladatot, amelyet egy fixpont-problémával azonosítunk. Végül a numerikus számításokat illusztráljuk két klasszikus példán. Az összefoglalás és az irodalom zárják a dolgozatot.

2 Entrópiaszerű eltérésfüggvények

Az $x \in R^n$ vektornak az $y \in R^n$ vektortól való eltérését mérő függvénynek a következő két tulajdonsággal kell rendelkeznie: a) nemnegatív legyen, és b) értéke akkor és csak akkor legyen 0, ha $x = y$. Nem követeljük meg a távolságfüggvény további tulajdonságait: nem írjuk elő a szimmetricitást és a háromszög-egyenlőtlenséget.

A két legismertebb eltérésfüggvény-család: a Bregman-divergencia és a Csiszár-féle φ-divergencia.

A *Bregman-divergencia* (ld. Bregman [1967]) arra az észrevételre épül, hogy egy $X_0 \subset R^n$ nyílt halmazon értelmezett tetszőleges n-változós szigorúan konvex, differenciálható f függvényre fennáll, hogy

$$f(x) - f(y) > \nabla f(y)(x-y)\,; \quad \forall x, y \in X_0\,, x \neq y\,.$$

Így $d(x,y) = f(x) - f(y) - \nabla f(y)(x-y)$ eltérésfüggvény, mert $d(x,y) > 0$ és $d(x,y) = 0$ akkor és csak akkor, ha $x = y$.

Ha például $f(x) = \sum_{i=1}^n x_i^2$, akkor $d(x,y) = \sum_{i=1}^n (x_i - y_i)^2$ – az euklideszi távolság négyzetét kapjuk. Ha $f(x) = \sum_{i=1}^n x_i \ln x_i,\ x_i > 0\ (i = 1, \ldots, n)$, akkor $d(x,y) = \sum_{i=1}^n (x_i \ln \frac{x_i}{y_i} + y_i - x_i)$, $y > 0$ – ez a Kullback-Leibler relatív entrópia-függvény. Ha $f(x) = -\sum_{i=1}^n \ln x_i,\ x_i > 0\ (i = 1, \ldots, n)$, akkor $d(x,y) = \sum_{i=1}^n (-\ln \frac{x_i}{y_i} + \frac{x_i}{y_i} - 1)$ – ez a Burg entrópia.

További példákat és gazdag irodalomjegyzéket találhat az érdeklődő olvasó pl. Kiwiel dolgozatában [1997].

A Bregman-divergencia nem korlátozódik a szeparábilis függvényekre, a *Csiszár-féle φ-divergencia* (ld. Csiszár Imre [1967]) azonban igen. A szóban forgó $\varphi(t)$ a pozitív félegyenesen értelmezett kétszer differenciálható szigorúan konvex valós függvény az alábbi tulajdonságokkal:

$$\varphi(1) = 0\,; \quad \varphi'(1) = 0\,; \quad \varphi''(1) > 0\,; \quad \lim_{t \to 0} \varphi'(t) = -\infty\,,$$

Az $x \in R^n_{++}$ vektornak az $y \in R^n_{++}$ vektortól való φ-divergencia eltérését ekkor így kapjuk:

$$d(x,y) = \sum_{i=1}^n d_i(x_i, y_i) = \sum_{i=1}^n y_i \varphi(\frac{x_i}{y_i})\,.$$

Könnyű belátni, hogy ekkor $d_i(x_i, y_i)$ folytonosan differenciálható és szigorúan konvex az első változójában, $\lim_{t\to+\infty} d'(t,a) = +\infty$, $\lim_{t\to 0} d'(t,a) = -\infty$, $d'(a,a) = 0$, ha $a > 0$. Ha például $\varphi(t) = t\ln t - t + 1$, akkor a Kullback-Leibler relatív entrópia-függvényt kapjuk. Ha $\varphi(t) = -\ln t + t - 1$, akkor

$$d_i(x_i, y_i) = y_i \ln \frac{y_i}{x_i} + x_i - y_i \ .$$

Ha $\varphi(t) = (\sqrt{t} - 1)^2$, akkor $d_i(x_i, y_i) = (\sqrt{x_i} - \sqrt{y_i})^2$ – ez a Hellinger távolság négyzetének egy konstans-szorosa.

További példákat találhat az érdeklődő olvasó pl. Iusem, Svaiter, Teboulle dolgozatában [1994].

3 Az A és B mátrixok eltérésének mérése a Kullback-Leibler relatív entrópia-függvény segítségével

Minthogy a $B = \{p_i/p_j\}$ mátrix elemei változatlanok, ha a p vektort egy pozitív számmal megszorozzuk, a p vektor normalizálása érdekében bevezetjük a $\sum_{i=1}^n p_i = 1$ feltételt. Ebben és a következő részekben feltesszük, hogy az A mátrix teljesen kitöltött.

Mind a súlyozott legkisebb négyzetek módszere, mind a legkisebb négyzetek módszere a Bregman függvények családjába tartozó eltérésfüggvényt alkalmaz: a p_i és az $a_{ij}p_j$, illetve az a_{ij} és a p_i/p_j $(i = 1,\ldots,n; j = 1,\ldots,n)$ értékek különbségeinek négyzetösszegének minimalizálása formájában keresi a prioritásvektort. E módszerek előnyeiről és hátrányairól az olvasó a hivatkozott dolgozatokból is tájékozódhat.

A jelen dolgozatban egy más megközelítést ajánlunk: A és B eltérését a p_i és az $a_{ij}p_j$ értékeknek a Kullback-Leibler relatív entrópiával mért eltérése összegeként határozzuk meg. Ez azt jelenti, hogy a prioritásvektort olyan $p^0 = (p_1^0, \ldots, p_n^0)$ vektor formájában keressük, amely minimalizálja az eltérések összegét, amelyre

$$p^0 \in \operatorname*{arg\,min}_{\sum_i p_i = 1, p > 0} \sum_i \sum_j (p_i \ln \frac{p_i}{a_{ij}p_j} + a_{ij}p_j - p_i) \ .$$

Miért jó ez a módszer?

- mert a minimumpont egyértelmű;
- mert a minimumpont pozitív és a p^0 vektor formájában kinyilvánított sorrendek nem változnak, ha a vektort megszorozzuk egy pozitív konstanssal;
- mert B prioritásvektora szintén p^0 akkor, ha elemeit p^0 elemei segítségével határozzuk meg;

- mert nem kell az alkalmazáshoz a páros összehasonlítás mátrixnak teljesen kitöltöttnek lennie (csak összefüggőnek);
- mert p^0 azzal a tulajdonsággal rendelkezik, hogy az optimális célfüggvény-érték legnagyobb csökkenését az A mátrix a_{ij} elemének és reciprokának változtatásával, a többi elem változatlanul hagyása mellett akkor érjük el, ha ezt az elemet az optimális p_i^0/p_j^0 értékkel helyettesítjük.

Ezeket az észrevételeket igazoljuk a következő részben, amelyben először az alábbi matematikai programozási feladat tulajdonságait vizsgáljuk:

$$\text{(K-L)}\qquad \begin{aligned} C(A,p) &= \sum_{i=1}^{n}\sum_{j=1}^{n}(p_i \ln \frac{p_i}{a_{ij}p_j} + a_{ij}p_j - p_i) \to \min \\ p \in P &= \{\, p \in R_{+}^{n} : \sum_{j=1}^{n} p_j = 1\,\} \quad p > 0\,, \end{aligned}$$

ahol A az adott páros összehasonlítás mátrix, R_+^n a nemnegatív térnegyed.

A $C(A,p)$ függvényt a továbbiakban az A páros összehasonlítás mátrix és a p elemei segítségével meghatározott konzisztens B mátrix Kullback-Leibler eltérésének nevezzük és az A és B *közelségeként* fogjuk fel. A (K-L) feladat egy optimális p megoldásából származtatott B mátrix pedig e felfogásnak megfelelően a ,,legközelebbi" az A páros összehasonlítás mátrixhoz. Felhívjuk a figyelmet arra, hogy ez a definíció tetszőleges két mátrix, vagy akár a B és A mátrixok – a sorrend fontos! – eltérésének, közelségének mérésére nem alkalmas.

Vizsgáljuk a (K-L) feladatot. Először vegyük észre, hogy a célfüggvény értelmezési tartománya a pozitív p vektorok R_{++}^n nyílt halmaza, ezért vizsgálnunk kell majd azt a kérdést, felveszi-e a minimumát $C(A,p)$ a korlátos, zárt, konvex P halmaz és R_{++}^n metszetén. Az első állításokban $C(A,p)$ tulajdonságait $p \in R_{++}^n$ vektorokra vizsgáljuk.

Vegyük észre, hogy $C(A,p)$ differenciálható a P halmaz relatív belső pontjainak halmazán: a

$$\operatorname{rel\,int} P = \{p \in R_{++}^n : \sum_{j=1}^{n} p_j = 1\} = P \cap R_{++}^n$$

halmazon, hiszen a függvényt alkotó összeg minden tagja differenciálható a $\operatorname{rel\,int} P$ halmazon.

1. Állítás. *Ha $p \in R_{++}^n$, akkor a $C(A,p)$ függvény nemnegatív és értéke 0 akkor és csak akkor, ha $p_i = a_{ij}p_j$ minden i,j indexpár ($i = 1,\dots,n$; $j = 1,\dots,n$) esetén.*

Bizonyítás. Emeljük ki a célfüggvényben p_i-t rögzített i és j indexek esetén (megtehetjük, mert p_i a feltevés szerint pozitív):

$$p_i \ln \frac{p_i}{a_{ij}p_j} + a_{ij}p_j - p_i = p_i(-\ln \frac{a_{ij}p_j}{p_i} + \frac{a_{ij}p_j}{p_i} - 1)\ .$$

Mint ismeretes, $\ln a \leq a - 1$ $\forall a \in R_{++}$ esetén és egyenlőség teljesül akkor és csak akkor, ha $a = 1$. Az állítás fennáll tehát a $C(A,p)$-t alkotó összeg minden tagjára, így magára az összegre is. $\square$

2. Állítás. *$C(A,p)$ szigorúan konvex és differenciálható a* $\operatorname{rel\,int} P$ *halmazon.*

Bizonyítás. Írjuk fel a $C(A,p)$ függvény tagjait az alábbi módon:

$$C(A,p) = \sum_{i=1}^{n} (np_i \ln p_i - p_i \sum_{j=1}^{n} \ln a_{ij} - p_i \sum_{j=1}^{n} \ln p_j + \sum_{j=1}^{n} a_{ij} p_j - np_i) \, .$$

Az $x \ln x$ szigorúan konvex függvény értelmezési tartománya az R_{++} pozitív nyílt félegyenes, a függvényt azonban a szokásoknak megfelelően lezárhatjuk a $\lim_{x \to 0} x \ln x = 0 \ln 0 = 0$ függvényérték-adással. Így a $C(A,p)$-t alkotó fenti szummában az első tag, és a második, negyedik és ötödik tag, amelyek lineárisak, összege:

$$C_1(A,p) = \sum_{i=1}^{n} \Big(np_i \ln p_i - p_i \sum_{j=1}^{n} \ln a_{ij} + \sum_{j=1}^{n} a_{ij} p_j - np_i \Big)$$

szeparábilis zárt szigorúan konvex függvényt alkot R_+^n- n. $C(A,p)$ fennmaradó tagja

$$C_2(A,p) = - \sum_{i=1}^{n} p_i \sum_{j=1}^{n} \ln p_j = \Big(- \sum_{j=1}^{n} p_j \Big) \Big(\sum_{j=1}^{n} \ln p_j \Big) \, ,$$

és $C_2(A,p) = - \sum_{j=1}^{n} \ln p_j$ a $\operatorname{rel\,int} P$ tartományon, vagyis szigorúan konvex. Ezzel az állítást beláttuk. $\square$

A következő állítás itt leírt bizonyítása Fülöp Jánostól származik, aki ezt az eredeti kéziratban szereplő eléggé körülményes gondolatmenet helyett ajánlotta.

3. Állítás. *$C(A,p)$ felveszi a minimumát a* $\operatorname{rel\,int} P$ *halmazon, a minimumpont egyértelmű.*

Bizonyítás. Az egyértelműség $C(A,p)$ szigorúan konvex voltából következik. A létezést kell bizonyítanunk. Vegyük észre, hogy az előző bizonyításban szereplő $C_1(A,p)$ függvény a minimumát a korlátos és zárt P halmazon felveszi, jelölje a minimumértéket L. Válasszunk egy tetszőleges $\hat{p} \in P$, $\hat{p} > 0$ pontot és legyen $K = C(A,\hat{p})$. Az optimális megoldás szempontjából P azon p pontjai érdekesek, amelyekre $K \geq C(A,p)$. Bármely $p \in P$, $p > 0$ esetén $-\ln p_i \geq 0$, mivel $p_i < 1$, $i = 1, \ldots, n$, így

$$L - \ln p_i \leq L - \sum_{j=1}^{n} \ln p_j \leq C(A,p) \, , \quad \forall p > 0, \; p \in P, \; i = 1, \ldots, n \, .$$

Válasszunk tetszőlegesen egy $\check{p} \in P$ pontot, amelyre $\check{p} > 0$ és $K \geq C(A,\check{p})$. Ekkor $L - \ln \check{p}_i \leq K$, azaz

$$\check{p}_i \geq e^{L-K} \, , \quad i = 1, \ldots, n \, .$$

A P halmaznak a $K \geq C(A,p)$ feltételt kielégítő pontjai tehát kompakt halmazt alkotnak, amelyen a folytonos $C(A,p)$ függvény szükségképpen felveszi a minimumát. A minimumpont egyben $C(A,p)$ minimumpontja a P halmazon. Ezzel az állítást beláttuk. □

4 Konzisztencia

A közelség fenti definíciója alapján az A páros összehasonlítás mátrixhoz létezik pontosan egy „legközelebbi" B konzisztens mátrix: a (K-L) feladat egyetlen optimális megoldását alkotó p prioritásvektor elemeiből alkotott hányadosok mátrixa. Mivel az eltérés nemnegatív és zéró akkor és csak akkor, ha a két mátrix egyenlő, ezért a két mátrix eltérése egyben az A mátrix inkonzisztenciájának méréséül is szolgálhat. Ebben a megállapításban Saaty (1987) gondolatmenetét követtük, aki az inkonzisztencia mérésére azért alkalmazta a $C.I. = (\lambda_{\max} - n)/(n-1)$ relatív eltérést (ill. ennek egy konstanssal való szorzatát), mert $\lambda_{\max} > n$. Hasonlóan ehhez a gondolatmenethez, javasoljuk a relatív eltérés és egyben az inkonzisztencia mérésére a $\min C(A,p)$ értéket, amely javaslatot a következő részben leírt 1. Következmény szándékunk szerint meggyőzően alátámasztja.

A következő kérdés az, hogyan kellene megváltoztatnunk az A mátrix i_0-adik sorában és j_0-adik oszlopában lévő $a_{i_0j_0}$ elemet, illetve ennek a reciprokát a j_0-adik sorban és i_0-adik oszlopban, hogy az A többi elemének változatlanul hagyása mellett a legnagyobb csökkenést érjük el a célfüggvényértékben.

4. Állítás. *Adott i_0, j_0 indexpár és adott p prioritásvektor esetén az A és B mátrixok Kullback-Leibler eltérése akkor csökken a legnagyobb mértékben, ha az $a_{i_0j_0}$ értékét a p_{i_0}/p_{j_0} hányadossal helyettesítjük; $C(A,p)$, mint $a_{i_0j_0}$ függvénye p_{i_0}/p_{j_0} egy környezetében konvex.*

Bizonyítás. Minimalizálni akarjuk adott p mellett a

$$C(A,p) = \sum_i \sum_j (p_i \ln \frac{p_i}{a_{ij}p_j} + a_{ij}p_j - p_i)$$

függvényt, mint a pozitív $a_{i_0j_0}$-ban egyváltozós függvényt. Írjuk fel a deriváltját:

$$\begin{aligned}
\frac{dC(A,p)}{da_{i_0j_0}} &= \Big(p_{i_0} \ln \frac{p_{i_0}}{a_{i_0j_0}p_{j_0}} + a_{i_0j_0}p_{j_0} - p_{i_0} + p_{j_0} \ln \frac{p_{j_0}}{\frac{1}{a_{i_0j_0}}p_{i_0}} + \frac{1}{a_{i_0j_0}}p_{i_0} - p_{j_0}\Big)' \\
&= \Big(-p_{i_0} \ln a_{i_0j_0} + p_{j_0}a_{i_0j_0} + p_{j_0} \ln a_{i_0j_0} + \frac{p_{i_0}}{a_{i_0j_0}}\Big)' \\
&= -\frac{p_{i_0}}{a_{i_0j_0}} + p_{j_0} + \frac{p_{j_0}}{a_{i_0j_0}} - \frac{p_{i_0}}{a_{i_0j_0}^2} = \Big(1 + \frac{1}{a_{i_0j_0}}\Big)\Big(p_{j_0} - \frac{p_{i_0}}{a_{i_0j_0}}\Big) .
\end{aligned}$$

Mivel $a_{i_0j_0}$ negatív nem lehet, ezért a derivált értéke 0, amint ezt vártuk is, ha: $a_{i_0j_0} = p_{i_0}/p_{j_0}$. Nézzük meg, hogy e pontban konvex-e. Írjuk fel a

$C(A,p)$ függvény $a_{i_0j_0}$ szerinti második deriváltját:

$$\frac{d^2C(A,p)}{da^2_{i_0j_0}} = \frac{p_{i_0}}{a^2_{i_0j_0}} - \frac{p_{j_0}}{a^2_{i_0j_0}} + \frac{2p_{i_0}}{a^3_{i_0j_0}} = \frac{1}{a^2_{i_0j_0}}\left(p_{i_0} - p_{j_0} + \frac{2p_{i_0}}{a_{i_0j_0}}\right) .$$

Ha $p_{i_0} \geq p_{j_0}$, akkor $d^2C(A,p)/da^2_{i_0j_0} > 0$. Ha $p_{i_0} < p_{j_0}$, akkor $d^2C(A,p)/da^2_{i_0j_0} > 0$, ha $0 < a_{i_0j_0} < 2p_{i_0}/(p_{j_0} - p_{i_0})$. Adott tetszőleges prioritásvektor esetén tehát $C(A,p)$ mint $a_{i_0j_0}$ függvénye konvex a pozitív számegyenesen, ha $p_{i_0} \geq p_{j_0}$ és konvex a $(0, \frac{2p_{i_0}}{p_{j_0}-p_{i_0}})$ tartományban, ha $p_{i_0} < p_{j_0}$. A p_{i_0}/p_{j_0} érték eleme mindkét tartománynak:

$$\frac{p_{i_0}}{p_{j_0}} \leq \frac{2p_{i_0}}{p_{j_0} - p_{i_0}} \Leftrightarrow p_{j_0} - p_{i_0} \leq 2p_{j_0} \Leftrightarrow -p_{i_0} \leq p_{j_0} .$$

Ezért $C(A,p)$ konvex e pont egy környezetében. Ezt akartuk belátni. □

Vegyük észre, hogy az $a_{i_0j_0} = p_{i_0}/p_{j_0}$ helyettesítéssel az A és B mátrixok Kullback-Leibler eltérése csökkenésének nagysága maga az eltérés:

$$p_{i_0} \ln \frac{p_{i_0}}{a_{i_0j_0}p_{j_0}} + a_{i_0j_0}p_{j_0} - p_{i_0} + p_{j_0} \ln \frac{p_{j_0}}{\frac{1}{a_{i_0j_0}}p_{i_0}} + \frac{1}{a_{i_0j_0}}p_{i_0} - p_{j_0} .$$

Ha tehát abból a szempontból vizsgáljuk meg az A mátrixot, hogy a konzisztencia növelése érdekében melyik összehasonlítás esetében ajánljuk a döntéshozónak, hogy vizsgálja felül a döntését, akkor A azon elemét kell kiválasztanunk, amelyre e csökkenés a legnagyobb. Mivel ez egyben a legnagyobb eltérés kiválasztását is jelenti, a választás összhangban van Saaty (2003) két javasolt módszerével is: az egyik a legnagyobb eltérés, a másik a legnagyobb csökkenés (Harker, 1987) kiválasztásának felel meg.

5. Állítás. *Legyen p^0 az A páros összehasonlítás mátrix prioritásvektora – a (K-L) feladat optimális megoldása –, és a B mátrix elemeit a p^0 elemeinek hányadosai alkossák. Akkor a B mátrix prioritásvektora is p^0 .*

Az állítás következik abból, hogy a B mátrixnak önmagától való eltérése 0, miközben minden más pozitív mátrixtól való eltérése pozitív az 1. Állítás szerint.

5 A Kullback-Leibler relatív entrópia minimalizálása: a feladat megoldása

Írjuk fel $C(A,p)$ gradiensét:

$$\frac{\partial}{\partial p_i}C(A,p) = \sum_{j=1}^{n} a_{ji} + \sum_{j=1}^{n} \ln a_{ji} + n \ln p_i - \sum_{j=1}^{n} \ln p_j - \frac{\sum_{j=1}^{n} p_j}{p_i} , \quad i = 1, \ldots, n .$$

Kihasználtuk, hogy $a_{ij} = 1/a_{ji}$, $i = 1, \ldots, n$, $j = 1, \ldots, n$. Írjuk fel a (K-L) feladathoz tartozó stacionárius pont problémát. A problémát és az itt következő állításokat ld. Mangasarian könyvében [1969]. Olyan $p = (p_1, \ldots, p_n, \omega, v_1, \ldots, v_n)$ értékeket keresünk, amelyek kielégítik az alábbi feltételeket:

$$\text{(K-T)} \qquad \begin{aligned} &\sum_{j=1}^{n} p_j = 1\,, \quad p = (p_1, \ldots, p_n) \in R^n_{++} \\ &\frac{\partial C(A,p)}{\partial p_i} - \omega - v_i = 0\,, \quad v_i \geq 0\,, \quad i = 1, \ldots, n\,, \\ &\sum_{j=1}^{n} v_j p_j = 0\,. \end{aligned}$$

Legyen $p^0 = (p_1^0, \ldots, p_n^0)$ a (K-L) feladat optimális megoldása – ez a 3. Állítás értelmében létezik és pozitív. Az optimalitás szükségességére vonatkozó Kuhn-Tucker tétel azt mondja ki, hogy a) mivel R^n_{++} nyílt és a (K-L) feladat célfüggvénye és feltételi függvényei differenciálhatók a $p^0 \in R^n_{++}$ pontban; b) mivel a (K-L) feltételrendszere reguláris: lineáris; ezért léteznek olyan $v_1^0, \ldots, v_n^0 \geq 0$, ω^0 duális változó értékek, hogy $(p_1^0, \ldots, p_n^0, \omega^0, v_1^0, \ldots, v_n^0)$ együttesen megoldását alkotják a (K-T) feladatnak. Minthogy $p^0 \in R^n_{++}$, ezért a $\sum_{j=1}^{n} v_j^0 p_j^0 = 0$ komplementaritási feltétel csak akkor teljesül, ha $v_1^0 = 0$, …, $v_n^0 = 0$.

Az állítás fordítva is fennáll. Az optimalitás elégségességére vonatkozó Kuhn-Tucker tétel azt mondja ki, hogy ha $(p_1^0, \ldots, p_n^0, \omega^0, v_1^0, \ldots, v_n^0)$ megoldása a (K-T) feladatnak, akkor a) mivel R^n_{++} nyílt és konvex; b) mivel a (K-L) feladat célfüggvénye és feltételi függvényei differenciálhatók a $p^0 \in R^n_{++}$ pontban és konvexek az R^n_{++} konvex halmazon, ezért p^0 optimális megoldása a (K-L) feladatnak.

A (K-L) feladat egyetlen pozitív optimális megoldása, és csak az, tehát kielégíti a következő problémát:

$$\text{(K-T)} \qquad \begin{aligned} &\sum_{j=1}^{n} p_j = 1\,, \quad p \in R^n_{++} \\ &\sum_{j=1}^{n} \Big(\ln \frac{p_i}{a_{ij} p_j} - \frac{p_j}{p_i} + a_{ji} \Big) = \omega\,, \quad i = 1, \ldots, n\,. \end{aligned}$$

A (K-T) feladatból kitűnik az ω változó jelentése. Szorozzuk meg az i-edik feltételt p_i-vel, és adjuk össze az egyenleteket. Azt kapjuk, hogy

$$\sum_{i=1}^{n} p_i \sum_{j=1}^{n} \Big(\ln \frac{p_i}{a_{ij} p_j} - \frac{p_j}{p_i} + a_{ji} \Big) = \omega \sum_{i=1}^{n} p_i \Leftrightarrow C(A,p) = \omega\,.$$

ω tehát az adott p vektorhoz tartozó, egyben optimális, célfüggvény-értéke a (K-L) feladatnak, ω nemnegatív és 0 akkor és csak akkor, ha $a_{ij} = p_i/p_j$ minden i, j indexpárra. Méri az A mátrix és a p vektor által meghatározott B mátrix Kullback-Leibler eltérését. De vegyük észre azt is, hogy az A mátrix reciprok volta miatt

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(\ln\frac{p_i}{a_{ij}p_j}-\frac{p_j}{p_i}+a_{ji}\Big)=\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(a_{ji}-\frac{p_j}{p_i}\Big)=n\omega\ ,$$

vagyis ω az átlaga az A mátrix illetve a B mátrix elemei összegeinek különbségének, ilyen értelemben is méri az A mátrix B-től való eltérését. Foglaljuk ezt össze:

1. Következmény. *A (K-T) feladat egyetlen (p^0,ω^0) megoldására fennáll, hogy*

$$\omega^0=\min_{p\in P, p\in R^n_{++}} C(A,p)=C(A,p^0)\,,\quad \omega^0=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(a_{ij}-\frac{p_i^0}{p_j^0}\Big)\ .$$

A stacionárius pont probléma alábbi vizsgálatával és átalakításával az a célunk, hogy a (K-L) feladat megoldására módszert mutassunk be a (K-T) feladat megoldása révén. Az egyenletrendszert alakítsuk át. Először írjuk fel mindkét oldalt logaritmusfüggvény formájában:

$$\text{(K-T)}\qquad \begin{gathered}\sum_{j=1}^{n}p_j=1\,,\quad p\in R^n_{++}\\ \ln\Big(p_i^n\cdot\prod_{j=1}^{n}a_{ji}e^{a_{ji}}\cdot\prod_{j=1}^{n}\frac{1}{p_j}e^{-p_j/p_i}\Big)=\ln e^{\omega}\,,\quad i=1,\dots,n\ .\end{gathered}$$

Figyelembe véve, hogy a logaritmus szigorúan növekvő függvény, ezért az i-edik egyenletben a két oldal egyenlő, ha az argumentumok egyenlők:

$$\prod_{j=1}^{n}a_{ji}e^{a_{ji}}=e^{\omega}\cdot\prod_{j=1}^{n}p_j\cdot\frac{1}{p_i^n}e^{1/p_i}\,,\quad i=1,\dots,n\ .$$

Osszuk el az első egyenlet két oldalát a második, ..., n-edik egyenlet megfelelő oldalaival! Az osztás elvégezhető, mert a szorzatok minden tényezője pozitív. Az alábbi ekvivalens feladathoz jutunk:

$$\text{(K-T)}\qquad \begin{gathered}\sum_{j}p_j=1\,,\quad p\in R^n_{++}\,,\\ \frac{1}{p_i^n}e^{1/p_i}=\frac{\prod_j a_{ji}e^{a_{ji}}}{\prod_j a_{j1}e^{a_{j1}}}\cdot\frac{1}{p_1^n}e^{1/p_1}\,,\quad i=2,\dots,n\ .\end{gathered}$$

Az $i=1$ eset azonosság. Könnyen látható, hogy

- $f(x)=\frac{1}{x^n}e^{1/x}$ a pozitív félegyenesen értelmezett pozitív, szigorúan csökkenő, differenciálható konvex függvény, $\lim_{x\to 0}f(x)=+\infty$, $\lim_{x\to+\infty}f(x)=0$;

- $f^{-1}(y)$ a pozitív félegyenesen értelmezett pozitív, szigorúan csökkenő, differenciálható konvex függvény, $\lim_{y\to 0} f^{-1}(y) = +\infty$, $\lim_{y\to+\infty} f^{-1}(y) = 0$;
- $p_i = f^{-1}(C_i f(p_1))$, ahol $C_i = \left(\prod_{j=1}^n a_{ji} e^{a_{ji}}\right) / \left(\prod_{j=1}^n a_{j1} e^{a_{j1}}\right)$, $i = 2, \ldots, n$;
- $f^{-1}(C_i f(p_1))$ a pozitív félegyenesen értelmezett pozitív, szigorúan növekvő, differenciálható függvény, $\lim_{p_1\to 0} f^{-1}(C_i f(p_1)) = 0$, $\lim_{p_1\to+\infty} f^{-1}(C_i f(p_1)) = +\infty$, $(f^{-1})'(C_i f(p_1)) > 0$ $(p_1 > 0)$.

A (K-T) feladat így a

$$g(p_1) = \sum_{i=1}^n f^{-1}(C_i f(p_1)) = 1$$

egyenlet formájában írható fel, ahol $C_1 = 1$. A fentiek szerint $g(p_1)$ a pozitív félegyenesen értelmezett pozitív, szigorúan növekvő, differenciálható függvény, $\lim_{p_1\to 0} g(p_1) = 0$, $\lim_{p_1\to+\infty} g(p_1) = +\infty$, $g'(p_1) > 0$, $(p_1 > 0)$.

A g függvény jó tulajdonságai miatt az egyenlet megoldására alkalmazhatók ismert numerikus módszerek: a biztosan konvergens intervallumfelező módszer és a húrmódszer, valamint a jó kezdőpontok választása esetén hatékonyabb szelő módszer, akár egyszerű Excel programként is. A szelő módszerrel való megoldás menetét bemutatjuk:

a. Kiindulásként válasszunk két pozitív p_1^1, p_1^2 értéket, amelyekre $g(p_1^1) \leq 1$, $g(p_1^2) \geq 1$. Ha valamelyikre egyenlőség teljesül, akkor az eljárás véget ér. Különben: legyen $k = 3$, folytassuk az eljárást a b. lépéssel.

b. Határozzuk meg p_1^k-t a következő módon:

$$p_1^k = p_1^{k-1} - \frac{g(p_1^{k-1}) - 1}{g(p_1^{k-1}) - g(p_1^{k-2})}(p_1^{k-1} - p_1^{k-2}) .$$

c. Ha $g(p_1^k) = 1$, akkor az eljárás véget ér. Különben: legyen $k = k + 1$, folytassuk az eljárást a b. lépéssel.

(A szóban forgó egyenletek kielégítését a szokásoknak megfelelő „elég nagy pontossággal” követeljük meg. Hogy alkalmanként milyen pontosságot célszerű elérni, erre itt nem térünk ki. Az egyenlet megoldására az Excel solvere is használható természetesen, de előfordul, hogy egy egyszerű solver nem talál megoldást, mert a kiszámítandó prioritások 0-hoz közeli értékek és ez kilendítheti a közelítő eljárást a menetéből.)

6 A (K-L) feladat nem teljesen kitöltött páros összehasonlítás mátrix esetén

A Kullback-Leibler relatív entrópia eltérésfüggvény ez esetben kevesebb tagot tartalmaz. Legyenek az A mátrix megadott (kitöltött) elemeihez tartozó index-halmazok soronként a következők: $\Gamma_i = \{\, j : a_{ij} > 0, (i,j) \text{ kitöltött}\,\}$. Nyilvánvaló, a páros összehasonlítás mátrix pozitív reciprok volta miatt, hogy $j \in \Gamma_i$ akkor és csak akkor, ha $i \in \Gamma_j$. Ekkor a (K-L) feladat, az eltérésfüggvény gradiense és a feladathoz tartozó Kuhn-Tucker stacionárius feladat az alábbi lesz:

$$\text{(K-L-1)} \qquad \begin{gathered} C(A,p) = \sum_{i=1}^{n} \sum_{j\in\Gamma_i} \Big(p_i \ln \frac{p_i}{a_{ij}p_j} + a_{ij}p_j - p_i \Big) \to \min \\ p \in P = \{\, p \in R_+^n : \sum_{j=1}^{n} p_j = 1 \,\}, \quad p \in R_{++}^n \,. \end{gathered}$$

$$\text{(K-T-1)} \qquad \begin{gathered} \frac{\partial C(A,p)}{\partial p_i} = \sum_{j\in\Gamma_i} \Big(\ln \frac{p_i}{a_{ij}p_j} - \frac{p_j}{p_i} + a_{ji} \Big), \quad i = 1,\ldots,n\,, \\ \sum_{i=1}^{n} p_i = 1\,, \quad p \in R_{++}^n\,, \\ \sum_{j\in\Gamma_i} \Big(\ln \frac{p_i}{a_{ij}p_j} - \frac{p_j}{p_i} + a_{ji} \Big) = \omega\,, \quad i = 1,\ldots,n\,. \end{gathered}$$

(A fenti szummákban a megfelelő tag 0 értékű, ha Γ_i üres.) A nem teljesen kitöltött A mátrixot *összefüggőnek* nevezzük, ha bárhogyan is választjuk az $\{1,2,\ldots,n\}$ indexhalmaz egy I valódi részhalmazát, az $\bigcup_{i\in I}\{j : j \in \Gamma_i\}$ halmaz bővebb, mint az I halmaz: $\exists k \in \bigcup_{i\in I}\{j : j \in \Gamma_i\}$, hogy k nem eleme az I halmaznak. Ez másként azt jelenti, hogy A nem összefüggő, ha a sorokat és oszlopokat (az alternatívákat) összehangolt módon átrendezhetjük úgy, hogy az eredményül kapott mátrix dekomponált: a kitöltött elemek első csoportja a mátrix bal felső k_1 sorában és oszlopában helyezkedik el, a maradék pedig a k_1+1-edik, ..., n-edik sorban és oszlopban, $k_1 < n$. A kitöltött és nem teljesen kitöltött páros összehasonlítás mátrix gráf reprezentációja segítségével az összefüggőséget is jól lehet szemléltetni, erről ld. például Bozóki-Fülöp-Rónyai cikkét [2010].

Nem teljesen kitöltött A mátrix esetében a (K-L-1) feladat konvexitására nem támaszkodhatunk. Támaszkodhatunk azonban arra az észrevételre, hogy $C(A,p)$ értéke nem változik, ha adott $p > 0$ prioritásvektor mellett a mátrix hiányzó elemeit kitöltjük a $\frac{p_i}{p_j}$ értékekkel. Jelölje $A(p)$ azt a –teljesen kitöltött– mátrixot, amelynek elemei a következők:

$$a_{ij}(p) = \begin{cases} a_{ij}, & \text{ha } j \in \Gamma_i \\ \frac{p_i}{p_j} & \text{különben.} \end{cases}$$

Világos, hogy $C(A,p) = C(A(p),p)$. A (K-L) problémát tehát nem teljesen kitöltött A páros összehasonlítás mátrix esetén a következőképpen írjuk fel:

$$\text{(K-L-1)} \qquad \begin{gathered} C(A,p) = C(A(p),p) = \sum_{i=1}^{n} \sum_{j\in\Gamma_i} \Big(p_i \ln \frac{p_i}{a_{ij}p_j} + a_{ij}p_j - p_i \Big) \to \min \\ p \in P\,, \quad p \in R_{++}^n\,. \end{gathered}$$

Vizsgáljuk a célfüggvény értékét. Legyen adott a $p^1 \in R^n_{++}$ és legyen $p^2 = \arg\min_{p\in P} C(A(p^1), p)$. A 3. Állítás értelmében p^2 létezik és pozitív. Nyilvánvaló, hogy $C(A(p^1), p) > C(A(p^1), p^2)$ minden p^2-től különböző $p > 0$ vektor esetén, p^1-t is beleértve, ha $p^1 \neq p^2$, vagyis $C(A(p^1), p^1) > C(A(p^1), p^2)$. Vegyük észre másfelől, hogy $C(A(p^1), p^1) = \min_{p\in P} C(A(p), p^1)$, a 4. Állítás szerint. Ha tehát a (K-L-1) feladatnak van $\hat{p} > 0$ optimális megoldása, azaz

$$C(A(\hat{p}), \hat{p}) = \min_{q\in P} \min_{p\in P} C(A(p), q) \ ,$$

akkor a $\hat{p}$ pontra fennáll, hogy

$$(*) \qquad \begin{gathered} C(A(p), \hat{p}) \geq C(A(\hat{p}), \hat{p}) \quad \text{és} \quad C(A(\hat{p}), q) > C(A(\hat{p}), \hat{p}) \\ \text{minden } p \in P, q \in P, q \neq \hat{p} \text{ esetén} \ . \end{gathered}$$

A (*) tulajdonság tehát maga után vonja, hogy $\hat{p}$ fixpontja az alábbi leképezésnek:

$$f(p) : P \to P, \quad f(p) = \arg\min_{q\in P} C(A(p), q) \ .$$

Azt, hogy ha $\hat{p}$ a (*) tulajdonsággal rendelkezik, akkor optimális megoldása lenne a (K-L-1) feladatnak, vagyis hogy $\hat{p}$ nem csak lokális, hanem globális minimumpont lenne, csak akkor állíthatnánk, ha a feladat konvexitását is belátnánk – erre ebben a dolgozatban nem kerül sor.

A továbbiakban nem teljesen kitöltött páros összehasonlítás mátrix esetén olyan prioritásvektor keresünk – és ennek birtokában a feladatunkat megoldottnak tekintjük –, amely a (*) tulajdonsággal rendelkezik. Ilyen prioritásvektor létezésére az alábbiakban konstruktív bizonyítást adunk.

6. Állítás. *Tegyük fel, hogy az A nem teljesen kitöltött páros összehasonlítás mátrix összefüggő. Ekkor a (K-L-1) feladatnak van olyan pozitív megoldása, amelyre a (*) tulajdonság teljesül.*

Bizonyítás. Induljunk ki tetszőleges pozitív $p^0 \in P$ vektorból. Hozzuk létre a p^k sorozatot, ennek a $k = 1, 2, \ldots$ elemét az alábbi módon:

Vizsgáljuk meg először, hogy p^{k-1} rendelkezik-e a (*) tulajdonsággal. *Ha igen*, készen vagyunk, az eljárás befejeződött. *Ha nem*, legyen $p^k = \arg\min_{p\in P} C(A(p^{k-1}), p)$. A 3. Állítás értelmében p^k létezik és pozitív. Az eljárást folytatjuk. Ha az eljárás véges lépésben befejeződik, akkor az állításban szereplő pontot létrehoztuk, a bizonyítás befejeződött. Ha nem, akkor egy végtelen sorozatot kapunk. Mivel a sorozat minden tagja a P korlátos zárt halmazból származik, ezért kiválasztható a $\{p^k\}_{k=0,1,\ldots}$ végtelen sorozatból egy konvergens $\{p^{k_j}\}_{j=0,1,\ldots}$ részsorozat. Jelölje e sorozat határértékét $\hat{p}$:

$$\hat{p} = \lim_{j\to\infty} p^{k_j} \ , \quad \hat{p} \in P \ .$$

A sorozat tagjaira fennállnak az alábbi egyenlőtlenségek:

$$\omega_{k_j-1} = C(A(p^{k_j-1}), p^{k_j}) \geq C(A(p^{k_j}), p^{k_j}) > C(A(p^{k_j}), p^{k_j+1}) = \omega_{k_j} \ .$$

Mivel $\{\omega_{k_j}\}$ alulról korlátos pozitív csökkenő sorozat, ezért létezik a limesze:

$$\lim_{j\to\infty}\omega_{k_j}=\lim_{j\to\infty}C(A(p^{k_j}),p^{k_j})=$$
$$=\lim_{j\to\infty}C(A(p^{k_j-1}),p^{k_j})=\lim_{j\to\infty}C(A(p^{k_j}),p^{k_j+1})=\omega\geq 0\ .$$

Ha $\hat{p}>0$, akkor

$$\lim_{j\to\infty}C(A(p^{k_j}),p^{k_j})=\min_{p\in P}C(A(\hat{p}),p)=$$
$$=\min_{p\in P}C(A(p),\hat{p})=C(A(\hat{p}),\hat{p})=\omega\geq 0\ ,$$

ekkor tehát $\hat{p}$ rendelkezik a (*) tulajdonsággal és az állítást bizonyítottuk. *Belátjuk még, hogy $\hat{p}>0$.* Ez következik. Vegyük észre, hogy minden $\delta>0$ számhoz létezik olyan r index, hogy $C(A(p^{k_j}),p^{k_j})<\omega+\delta$, ha $j>r$, mivel $C(A(p),p)$ folytonos függvénye p-nek és $C(A(p^{k_j}),p^{k_j})$ csökkenő sorozat, amely $\omega\geq 0$ értékhez tart. Válasszunk δ-t.

Vegyük észre, hogy $\hat{p}_j>0$ legalább egy indexre, mivel $\sum_{j=1}^n\hat{p}_j=1$. Tegyük fel az állítással ellentétben, hogy legalább egy i indexre $\hat{p}_i=0$. Vegyük észre, hogy e feltevés mellett van olyan i_0, j_0 indexpár – és ezt itt kiválasztjuk –, hogy $\hat{p}_{i_0}=0$, $\hat{p}_{j_0}>0$ és az A mátrix (i_0,j_0)-adik helye ki van töltve. Ha ugyanis minden i indexre, amelyre $\hat{p}_i=0$ fennállna, hogy $\hat{p}_j=0$, ha $j\in\{j:j\in\Gamma_i\}$, akkor az $I=\{i:\hat{p}_i=0\}$ indexhalmazra azt kapnánk, hogy $I=\bigcup_{i\in I}\{j:j\in\Gamma_i\}$, ellentmondásban azzal a feltevéssel, hogy A összefüggő. Írjuk fel $C(A(p^{k_r}),p^{k_r})$ értékét:

$$C(A(p^{k_r}),p^{k_r})=\sum_{i\in\{1,\dots,n\},j\in\Gamma_i}\Big(p_i^{k_r}\ln\frac{p_i^{k_r}}{a_{ij}p_j^{k_r}}+a_{ij}p_j^{k_r}-p_i^{k_r}\Big)=$$
$$=\sum_{\substack{i\in\{1,\dots,n\},j\in\Gamma_i\\(i,j),(j,i)\neq(i_0,j_0)}}\Big(p_i^{k_r}\ln\frac{p_i^{k_r}}{a_{ij}p_j^{k_r}}+a_{ij}p_j^{k_r}-p_i^{k_r}\Big)+$$
$$+p_{i_0}^{k_r}\ln\frac{p_{i_0}^{k_r}}{a_{i_0j_0}p_{j_0}^{k_r}}+a_{i_0j_0}p_{j_0}^{k_r}-p_{i_0}^{k_r}+p_{j_0}^{k_r}\ln\frac{p_{j_0}^{k_r}}{a_{j_0i_0}p_{i_0}^{k_r}}+a_{j_0i_0}p_{i_0}^{k_r}-p_{j_0}^{k_r}\ .$$

A szumma minden tagja nemnegatív, összegük és így tagjaik értéke is kisebb, mint $\omega+\delta$. Írjuk fel az összeg utolsó tagjára vonatkozó egyenlőtlenséget részletesebben:

$$p_{j_0}^{k_r}\ln p_{j_0}^{k_r}-p_{j_0}^{k_r}\ln a_{j_0i_0}-p_{j_0}^{k_r}\ln p_{i_0}^{k_r}+a_{j_0i_0}p_{i_0}^{k_r}-p_{j_0}^{k_r}<\omega+\delta\ .$$

De a bal oldalon lévő tagok közül $-p_{j_0}^{k_r}\ln p_{i_0}^{k_r}\to+\infty$, ha $j\to+\infty$, miközben a többi tag véges értékhez tart. Ellentmondásra jutottunk tehát azzal a feltevéssel, hogy a $\hat{p}$ prioritásvektornak lehet 0 értékű eleme. Ezzel az állítást bizonyítottuk. □

A 6. Állításból és bizonyítása gondolatmenetéből következik, hogy a (K-T-1) feladatnak van megoldása, azt azonban nem zártuk ki, hogy egynél több megoldása is lenne. Bármely megoldására azonban fennáll a következő

2. Következmény. *A (K-T-1) feladat egy (p^0, ω^0) megoldására fennáll, hogy*

$$\omega^0 = \min_{p \in P, p \in R^n_{++}} C(A(p^0), p) = C(A(p^0), p^0)\,, \quad \omega^0 = \frac{1}{n} \sum_{i=1}^{n} \sum_{j \in \Gamma_i} \left(a_{ij} - \frac{p_i^0}{p_j^0} \right) .$$

7 Numerikus vizsgálatok

Két klasszikus példát választottunk abból a célból, hogy az eredményeket más szerzők eredményeivel összevethessük. A teljesen kitöltött páros összehasonlítás mátrix-szal kapcsolatos számításokat a ,,Távolság Philadelphiától" példával, a nem teljesen kitöltött esetben pedig a ,,A család házat vásárol" példával is illusztráljuk, mindkettő Saaty híres könyvéből [1980] származik.

7.1 ,,Távolság Philadelphiától"

Ebben a példában a döntéshozó hat várost hasonlított össze azoknak Philadelphiától való relatív távolságuk tekintetében. E városok (alternatívák) sorra: Kairó, Tokió, Chicago, San Francisco, London, Montreal. A mátrixnak elég lenne csak a jobb felső háromszögét feltüntetni, a bal alsó háromszög szükségképpen a reciprok értékeket tartalmazza:

$$\begin{pmatrix} 1 & 1/3 & 8 & 3 & 3 & 7 \\ 3 & 1 & 9 & 3 & 3 & 9 \\ 1/8 & 1/9 & 1 & 1/6 & 1/5 & 2 \\ 1/3 & 1/3 & 6 & 1 & 1/3 & 6 \\ 1/3 & 1/3 & 5 & 3 & 1 & 6 \\ 1/7 & 1/9 & 1/2 & 1/6 & 1/6 & 1 \end{pmatrix}$$

Alább az első sorban a jelen dolgozatban leírt módszerrel kapott prioritásvektort, a második sorban pedig Saaty sajátérték módszerével kapott eredményt tüntettük fel. A harmadik sorban pedig az alternatívák sorrendjét, amely a két prioritásvektor esetében ugyanaz lett:

0,24368	0,42483	0,03444	0,10662	0,14868	0,03176
0,26185	0,39749	0,03343	0,11639	0,16424	0,02660
2	1	5	4	3	6

Ugyanezt a mátrixot nem teljesen kitöltötté alakítottuk Harker [1987b] példája nyomán. Kitöltöttek maradtak az (1,2), (1,3), (2,5), (3,6), (4,5) és (5,6) elemek, párjaik és a födiagonális. Az így kapott prioritásvektort alább a dolgozatban leírt módszerrel az első, illetve Harker módszerével a második sorban tüntettük fel, a harmadikban pedig ismét az alternatívák sorrendjét, amely a két vektor esetében ismét ugyanaz lett:

0,1859	0,5371	0,0261	0,0570	0,1704	0,0235
0,2339	0,4612	0,0382	0,0659	0,1734	0,0273
2	1	5	4	3	6

Megjegyezzük, hogy $C(A,p)$ optimális értéke 0,442755 a teljesen kitöltött mátrix esetében, és $C(A(p),p)$ értéke 0,015089, amikor $p=$ (0,1859; 0,5371; 0,0261; 0,0570; 0,1704; 0,0235). Ezeket az értékeket az inkonzisztencia mértékének tekintjük, és amint vártuk, lényegesen kisebb a hiányos elemeket tartalmazó mátrix esetében.

7.2 „A család házat vásárol"

Az eredeti mátrix Saaty könyvében teljesen kitöltött, itt arra a nem teljesen kitöltött változatára vonatkozó számítási eredményeket mutatjuk be, amelyet Bozóki, Fülöp és Rónyai használt [2010]. Ezt a nem teljesen kitöltött mátrixot az alábbi táblázat tartalmazza, ha a kivastagított elemeket elhagyjuk. A mátrix sorai és oszlopai által képviselt szempontok sorra a következők voltak: méret, közlekedés, szomszédság, kor, a kert, korszerűség, állapot, ár.

A táblázatot kitöltöttük a számításunk eredményeként kapott $p=$(0,1439; 0,0695; 0,1063; 0,0318; 0,0242; 0,0525; 0,0657; 0,5061) prioritásvektor segítségével, a hiányzó elemeket pótoltuk tehát a kivastagított értékekkel.

$$\begin{pmatrix}
1 & 5 & 3 & 7 & 6 & 6 & 0,3333 & 0,25 \\
0{,}2 & 1 & \mathbf{0{,}6539} & 5 & \mathbf{2{,}8754} & 3 & \mathbf{1{,}0588} & 0{,}1429 \\
0{,}3333 & \mathbf{1{,}5293} & 1 & \mathbf{3{,}3451} & 3 & \mathbf{2{,}0264} & 6 & \mathbf{0{,}2101} \\
0{,}1429 & 0{,}2 & \mathbf{0{,}2989} & 1 & \mathbf{1{,}3146} & 0{,}25 & \mathbf{0{,}4841} & 0{,}125 \\
0{,}1667 & \mathbf{0{,}3478} & 0{,}3333 & \mathbf{0{,}7607} & 1 & \mathbf{0{,}4608} & 0{,}2 & \mathbf{0{,}0478} \\
0{,}1667 & 0{,}3333 & \mathbf{0{,}4935} & 4 & \mathbf{2{,}1701} & 1 & \mathbf{0{,}7991} & 0{,}1667 \\
3 & \mathbf{0{,}9445} & 0{,}1667 & \mathbf{2{,}0658} & 5 & \mathbf{1{,}2514} & 1 & \mathbf{0{,}1297} \\
4 & 7 & \mathbf{4{,}7602} & 8 & \mathbf{20{,}9332} & 6 & \mathbf{7{,}7081} & 1
\end{pmatrix}$$

Az említett szerzők két módszer segítségével határozták meg a hiányzó elemek értékét. Az első a sajátérték módszer, a második az LLSM – a logaritmikus legkisebb négyzetek módszere – kiterjesztése nem teljesen kitöltött páros összehasonlítás mátrixok esetére. A két eredmény – mind a prioritásvektor, mind a hiányzó elemek értékei tekintetében lényegesen eltér. Az általunk kapott eredmény az LLSM segítségével kapott eredményhez közelebb áll, a kapott prioritásvektorok nem azonosak természetesen, de az egyes alternatívák általuk meghatározott sorrendjei is inkább csak a kis prioritású alternatívák esetében cserélődnek fel. A tanulság, ha egyáltalán van, talán az lehet, hogy érdemes több módszert alkalmazni és a konkrét alkalmazás illetve a konkrét páros összehasonlítás mátrix ismeretében további elemzéseket végezni az egyes alternatívák fontosságát illetően.

8 Összefoglalás

A dolgozatban a döntéselméletben fontos szerepet játszó páros összehasonlítás mátrix prioritásvektorának meghatározására új megközelítést alkalmaztunk: az A páros összehasonlítás mátrix és a prioritásvektor által definiált B konzisztens mátrix közötti eltérést a Kullback-Leibler relatív entrópia-függvény segítségével mértük. Ezen eltérés minimalizálása teljesen kitöltött mátrix esetében konvex programozási feladathoz vezetett, nem teljesen kitöltött mátrix esetében pedig egy fixpont problémához. A Kullback-Leibler eltérésfüggvényt minimalizáló prioritásvektor egyben azzal a tulajdonsággal is rendelkezik, hogy az A mátrix elemeinek összege és a B mátrix elemeinek összege közötti különbség éppen az eltérésfüggvény minimumának az n-szerese, ahol n a feladat mérete. Így az eltérésfüggvény minimumának értéke két szempontból is lehet alkalmas az A mátrix inkonzisztenciájának a mérésére. A feladat vizsgálata során alkalmazott bizonyítások egy része konstruktív volt, egyben eljárást adott a prioritásvektor kiszámítására. Nyitva maradt jó néhány kérdés, amelyek további elemzést igényelnek, pl. a nem teljesen kitöltött páros összehasonlítás mátrix esetében a 6. Állítás bizonyításában szereplő sorozat – és nem csak a részsorozat – konvergenciájának a vizsgálata. Érdekesnek látszik egy kísérlet a Burg entrópia mint eltérésfüggvény alkalmazására is.

Irodalom

1. Bozóki S., Fülöp J. Poesz A. (2011). On pairwise comparison matrices that can be made consistent by the modification of a few elements, *Central European Journal of Operations Research 19*, 157–175.
2. Bozóki, S., Fülöp, J., Rónyai, L. (2010). On optimal completion of incomplete pairwise comparison matrices, *Mathematical and Computer Modelling 52*, 318–333.
3. Bregman, L. (1967). The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming, *USSR Comput. Math. and Math Phys. 7*, 200–217.
4. Choo, E. and Wedley, W. (2004). A Common framework for deriving preference values from pairwise comparison matrices, *Computers and Operations Research 31*, 893–908.
5. Csató László (2012). Ranking by pairwise comparisons for Swiss-system tournaments, *Central European Journal of Operations Research*, published first online: DOI 10.1007/s10100-012-0261-8.
6. Csiszár, I. (1967). Information type measures of differences of probability distributions and indirect observations, *Studia Sci. Math. Hungar. 2*, 299–318.
7. Farkas, A., Lancaster, P. and P. Rózsa (2003). Consistency adjustments for pairwise comparison matrices, *Numerical Linear Algebra with Applications, 10*, 689–700.
8. Fedrizzi, M., S. Giove(2007). Incomplete pairwise comparison and consistency optimization *European Journal of Operational Research 183*, 303–313
9. Gass, S. I. and T. Rapcsák (2004). Singular value decomposition in AHP, *European Journal of Operational Research 154*, 573–584.
10. Harker, P. T. (1987a). Derivatives of the Perron root of a positive reciprocal matrix: with application to the Analytic Hierarchy Process, *Applied Mathematical Computation 22*, 217–232.
11. Harker, P. T. (1987b). Incomplete pairwise comparisons in the analytic hierarchy process, *Mathematical Modelling*, 9, 837–848.

12. Ishizaka, A., A. Labib (2011). Review of the main developments in the Analytic Hierarchy Process, *Expert Systems with Applications 38(11)*, 14336–14345.
13. Iusem, A. N., Svaiter, B. F. & Teboulle M. (1994). Entropy-like proximal methods in convex programming. *Math. of Operations Research 19*, 790–814.
14. Kéri Gerzson (2005). Kritériumok páros összehasonlitás mátrixokra, *Szigma 36*, 139–148.
15. Kiwiel, K. C. (1997). Proximal minimization methods with generalized Bregman functions. *SIAM J. Control and Optimization 35*, 1142–1168.
16. Krovák, J. (1987). Ranking alternatives – comparison of different methods based on binary comparison matrices, *European Journal of Operational Research 32*, 86–95.
17. Mangasarian, O. L. (1969). *Nonlinear programming.* McGraw-Hill Book Company, New York.
18. Saaty, T. (1980). *The Analytic Hierarchy Process.* New York: McGraw-Hill.
19. Saaty, T. (2003). Decision-making with the AHP: Why is the principal eigenvector necessary? *European Journal of Operational Research, 145*, 85–91.
20. Temesi József (2011). Pairwise comparison matrices and the error-free property of the decision maker, *Central European Journal of Operations Research*, Vol 19, No 2, 239–249.

APPLYING THE KULLBACK-LEIBLER RELATIVE ENTROPY FUNCTION FOR DETERMINING PRIORITIES FOR THE PAIRWISE COMPARISON MATRIX

In this paper we apply a new approach for determining a priority vector for the pairwise comparison matrix which plays an important role in Decision Theory. The divergence between the pairwise comparison matrix A and the consistent matrix B defined by the priority vector is measured with the help of the Kullback-Leibler relative entropy function. The minimization of this divergence leads to a convex program in case of a complete matrix, leads to a fixed-point problem in case of an incomplete matrix. The priority vector minimizing the divergence also has the property that the difference of the sums of elements of the matrix A and the matrix B is n times the minimum of the divergence function where n is the dimension of the problem. Thus we developed two reasons for considering the value of the minimum of the divergence as a measure of inconsistency of the matrix A.