REGRESSZIÓS JÁTÉKOK[1]

PINTÉR MIKLÓS
Budapesti Corvinus Egyetem

Egy kooperatív játék megoldása az egyes játékosok által együttesen elérhető eredmény bizonyos elvek szerinti elosztása a játékosok között. Egy regressziós modellben a rendelkezésre álló magyarázó változók által együttesen elérhető illeszkedés az az eredmény, amit szét szeretnénk osztani. Az egyes magyarázó változók játékelméleti módszerekkel való értékelése egyrészt hozzásegít az adott, modellezni kívánt probléma jobb megértéséhez, másrészt segít kiválasztani azoknak a változóknak a körét, amelyek az adott probléma modellezéséhez szükségesek. A cikk célja, hogy a kooperatív játékelméletben jól ismert Shapley-érték fogalmat használva értékeljük a regressziós modellek magyarázó változóit. Áttekintjük a Shapley-érték Hart és Mas-Colell–féle karakterizációját, és példákon mutatjuk be a javasolt eljárást. Konklúzió: a Shapley-érték használata regressziós modellek magyarázó változóinak értékelésére védhető, jól interpretálható módszer.

1 Bevezető

Az ökonometriai, statisztikai elemzések során gyakran felmerül az adott magyarázó változók, vagy bizonyos hatások fontosságának megállapítása. Ismerjük az elemezni kívánt jellemzők együttes hatásait, marginális hatásait, ezekből szeretnénk a változók egy tisztított, egyéni értékelését kapni, mely egyéni értékelés jól jellemzi az egyes változókat, hatásokat. A feladatot megfogalmazhatjuk úgy is, hogy szét szeretnénk osztani az egyes magyarázó változók együttes hatását az egyes változók között úgy, hogy a szétosztás rendelkezzen bizonyos tulajdonságokkal.

A játékelméletben az ilyen, vagy ehhez hasonló kérdésekre az átruházható hasznosságú kooperatív játékok megoldásai adnak választ. A feladat ott az, hogy szétosszuk az egyes játékosok (magyarázó változók) között az általuk együttesen elért eredményt (illeszkedést).

A kooperatív játékelméleti megoldások alkalmazása változóértékelésre már Chevan és Sutherland [2] cikkében felmerül, akik különböző regressziós feladatok kapcsán a magyarázó változók együttes magyarázóerejét (többszörös

[1]Köszönöm Forgó Ferenc professzornak, hogy ráirányította a figyelmemet a cikk témájára, nevezetesen arra, hogy a kooperatív játékelméleti megoldáskoncepciók használhatóak többváltozós regressziós modellek magyarázó változóinak értékelésére. Külön köszönöm a cikk bírálójának, Hajdu Ottónak, értékes észrevételeit, javaslatait. Köszönettel tartozom továbbá Forgó Ferencnek, Orosz Ágotának és Solymosi Tamásnak, akik számos megjegyzésükkel, javaslatukkal segítették munkámat. Természetesen az előforduló pontatlanságokért, hibákért egyedül én vagyok a felelős. Ez a munka az OTKA T046194 pályázat támogatásával készült. Beérkezett: 2007. május 30. E-mail: `miklos.pinter@uni-corvinus.hu`.

determinációs együttható) osztották szét az egyes változók között a Shapley-érték szerint. Chevan és Sutherland nem ismerte fel, hogy a Shapley-értéket használták értékelésükre, erre csak Stufken [17] hívta fel a figyelmet később. Ettől a pontól kezdve világos volt, hogy a különböző statisztikai, változó-értékelési stb. vizsgálatok során a kooperatív játékelmélet bizonyos megoldáskoncepciói, fogalmai használhatóak. Shorrocks [16] megmutatta, hogy a hatásértékelési (változóértékelési) módszerek egy széles csoportja tekinthető úgy, mint a Shapley-érték különböző megfogalmazásai, sőt, a szerző által az értékelési eljárásokkal szemben megfogalmazott elvárásoknak a Shapley-érték eleget tesz, tehát abban az értelemben ideális értékelési eljárás. Lipovetsky és Conklin [10] szintén a Shapley-értéket használja regressziós modellek változó-értékelésének meghatározására, különös tekintettel a multikollinearitás által okozott értékelési problémák kezelésére. Lipovetsky és Conklin nem ismerik a fenti szerzőket, így nem tudják, hogy eredményeik csak részben újak.

Számos, a Shapley-értéket használó, nem közgazdasági, nem üzleti alkalmazott írás, alkalmazás, elemzés jelent már meg. Csak iránymutatás végett, példaként megemlítünk néhányat: Cox [3], Gefeller et al. [6], Albrecht et al. [1], Wan [18], Zhang és Wan [19].

Grömping [8] a Shapley-érték statisztikai alkalmazásának egy részletes áttekintését adja. A cikk megjelenésének dátuma (2007), továbbá annak tartalma egyértelműen mutatja, hogy a Shapley-érték statisztikai alkalmazásának elmélete fontos, még formálódó, koránt sem lezárt kutatási terület.

A cikk célja, hogy elméleti oldalról alátámasszuk a Shapley-érték használatát regressziós modellek magyarázó változóinak értékelésére. Ebből a célból bevezetjük a regressziós játékok fogalmát, amely fogalom nem mintákra, hanem valószínűségi változókra épülő fogalom. A regressziós játékok ez a fajta absztrakt tulajdonsága lehetővé teszi, hogy elvi oldalról jellemezzük ezt a játékosztályt. A fő jellemzésünk a Shapley-érték Hart és Mas-Colell–féle [9] karakterizációjára épül. A 27. tételben megmutatjuk, hogy a Shapley-érték az egyetlen olyan, a regressziós játékok osztályán értelmezett értékelés, amely rendelkezik a Hart és Mas-Colell által megkövetelt, és a regressziós problémák során jól interpretálható, jogosan elvárható tulajdonságokkal.

A munka felépítése a következő: a 2. szakaszban áttekintjük a cikkben használt játékelméleti fogalmakat. A 3. szakaszban a Shapley-érték Hart és Mas-Colell–féle potenciálra támaszkodó jellemzését tárgyaljuk. A 4. szakaszban bevezetjük a regressziós játék fogalmát (22. definíció), továbbá megmutatjuk, hogy a potenciállal miként jellemezhető a Shapley-féle értékelés a regressziós játékok osztályán (27. tétel). Végül, az 5. szakaszban egy konkrét alkalmazási módszert ismertetünk. Az utolsó szakasz az összefoglalásé.

2 Kooperatív játékok

Ebben a szakaszban röviden áttekintjük a cikkben érintett kooperatív játékelméleti fogalmakat. Nem térünk ki részletesen minden fogalomra, eredményre, amit használni kívánunk, hanem támaszkodva az [5] jegyzetre, csak azokat a fogalmakat, eredményeket tárgyaljuk, amelyek közvetlenül szükségesek a cikk megértéséhez.

1. definíció. Legyen N a játékosok véges halmaza, és legyen $v : \mathcal{P}(N) \to I\!R$ olyan függvény, hogy $v(\emptyset) = 0$, ahol $\mathcal{P}(N)$ az N halmaz hatványhalmaza. Ekkor v-t karakterisztikus függvénnyel adott, átruházható hasznosságú, más néven TU (transferable utility) kooperatív játéknak (a továbbiakban röviden ,,csak" kooperatív játéknak) nevezzük.

A fenti definíció mögött meghúzódó intuíció a következő: az N halmaz részhalmazai az egyes koalíciók, míg a v karakterisztikus függvény értékei az egyes koalíciók által elérhető kifizetést, hasznosságot adják meg. Látható, hogy a koalíciók tagjai együtt érnek el bizonyos kifizetési értékeket, és az elért értékeket tetszőlegesen szét tudják osztani a résztvevők között; innen az átruházható hasznosság jelző.

2. segédtétel. Legyen $\mathcal{G}^N$ az $|N|$ (N számossága) elemű játékoshalmazzal rendelkező kooperatív játékok osztálya. Ekkor tetszőleges N-re $\mathcal{G}^N$ és $I\!R^{2^{|N|}-1}$ izomorfak.

Bizonyítás. A bizonyítást az olvasóra bízzuk. □

A fenti segédtételből következik, hogy valójában nem maga a játékosok halmaza N, hanem N számossága az, ami fontos. Tehát a játékosztály fogalma nem a játékosok személyére, hanem azok számára épül. A továbbiakban feltesszük, hogy $v \in \mathcal{G}^N$ olyan kooperatív játék, amelynek játékoshalmaza $N = \{1, \dots, n\}$, és a két tér között rögzített izomorfizmusunk van, magyarán szólva, feltesszük, hogy $I\!R^{2^n-1}$ egy rögzített bázisa mellett definiáljuk v-t.

3. definíció. A $v \in \mathcal{G}^N$ kooperatív játék monoton, ha tetszőleges olyan $A, B \in \mathcal{P}(N)$ esetén, hogy $A \subseteq B$, $v(A) \leq v(B)$.

A monoton kooperatív játékban egy új játékos tetszőleges koalícióba való belépése nem csökkenti az elérhető hasznosságot.

4. definíció. A $v \in \mathcal{G}^N$ kooperatív játék szuperadditív (szubadditív), ha tetszőleges olyan $A, B \in \mathcal{P}(N)$-re, hogy $A \cap B = \emptyset$, $v(A \cup B) \geq v(A) + v(B)$ ($v(A \cup B) \leq v(A) + v(B)$). v additív, ha egyszerre szuper- és szubadditív, tehát, ha tetszőleges olyan $A, B \in \mathcal{P}(N)$-re, hogy $A \cap B = \emptyset$, $v(A \cup B) = v(A) + v(B)$.

A szuperadditív kooperatív játékokban számolhatunk a nagykoalíció (N) megalakulásával, hiszen az összes játékos összefogása olyan hasznossági szintet tud biztosítani a játékosoknak, amit más ,,részösszefogásokkal" felülmúlni nem lehet. A szubadditív játékok esetén azonban (kivéve az additív esetet) nem várható a nagykoalíció megalakulása, ezek a kooperatív játékok bizonyos értelemben patologikusak.

5. definíció. A $v \in \mathcal{G}^N$ kooperatív játék lényeges, ha $v(N) > \sum\limits_{i \in N} v(\{i\})$.

A lényegesség esetében az elvárás az, hogy a nagykoalíció által elérhető hasznossági szint haladja meg a játékosok által egyénileg elérhető hasznosságok összegét. A lényeges elnevezésre a magyarázat abban rejlik, hogy a nem lényeges játékok esetén végképp nem számolhatunk a nagykoalíció megalakulásával.

6. megjegyzés. Nagyon sok egyensúlyfogalom csak lényeges játékok esetén bír jelentéssel. Így pl. a mag, kernel, alkuhalmaz, stabil halmaz, nukleolusz csak lényeges játékok esetén tartalmas fogalom (lsd. pl. [4, 5, 11]).

Ebben a cikkben a Shapley-értéket (Shapley [14]) használjuk.

7. definíció. Legyen $v \in \mathcal{G}^N$ kooperatív játék, és legyen $v_i'(S) = v(S \cup \{i\}) - v(S)$, ahol $i \in N$, $S \in \mathcal{P}(N)$. Legyen továbbá tetszőleges $i \in N$ esetén

$$f^i_{Sh}(S) = \begin{cases} \dfrac{|S|!(|N|-|S|-1)!}{|N|!}, & \text{ha } i \notin S \\ 0 & \text{különben} \end{cases}$$

eloszlás $\mathcal{P}(N)$-en[2]. Ekkor $\phi_i(v)$, az i játékos Shapley-értéke a következő:

$$\phi_i(v) = \sum_{S \in \mathcal{P}(N)} v_i'(S)\ f^i_{Sh}(S)\ . \tag{1}$$

A Shapley-érték mögé egy jól megfogható intuíció helyezhető. A $v_i'(S)$ az az érték, ami az i játékos ceteris paribus hozzájárulása az S koalícióhoz. Tehát, ha az i játékos nem tagja az S koalíciónak, akkor az ő belépése az S koalícióba $v_i'(S)$ „többletet” hoz, ha pedig az i játékos tagja az S koalíciónak, akkor $v_i'(S) = 0$. Tegyük fel, hogy a koalíciók az egyes játékosok véletlenszerű sorrendjéből alakulnak ki (pl. érkezési sorrend), tehát minden azonos elemszámú koalíció bekövetkezésének azonos a valószínűsége. Legyen S egy tetszőleges koalíció, ekkor ha $i \notin S$, akkor az S koalíció $|S|!(|N|-|S|-1)!/|N|!$ valószínűséggel alakul meg, így az i játékos hozzájárulása ezen a koalíción keresztül várhatóan $v_i'(S)|S|!(|N|-|S|-1)!/|N|! = v_i'(S) f^i_{Sh}(S)$, tehát $\phi_i(v)$ nem más, mint v_i', f^i_{Sh}-ra vonatkozó várható értéke.

A következő példában a fent ismertetett Shapley-értéket és a hozzá kapcsolható intuíciót mutatjuk be.

8. példa. Legyen $N = \{1, 2, 3\}$, és legyen v a következő:

$$v(\{1\}) = 1\,, \quad v(\{2\}) = 0\,, \quad v(\{3\}) = 2\,,$$
$$v(\{1,2\}) = 2\,, \quad v(\{1,3\}) = 2\,, \quad v(\{2,3\}) = 3\,,$$
$$v(\{1,2,3\}) = 5\,.$$

Érkezési sorrendek:

1	1	2	2	3	3
2	3	1	3	1	2
3	2	3	1	2	1

[2]Tehát $\sum_{S \in \mathcal{P}(N)} f^i_{Sh}(S) = 1$.

Határhozzájárulások (játékosonként):

$$\begin{matrix} 1 & 1 & 2 & 2 & 0 & 2 \\ 1 & 3 & 0 & 0 & 3 & 1 \\ 3 & 1 & 3 & 3 & 2 & 2 \end{matrix}$$

A Shapley-értékek:

$$\phi_1(v) = \frac{1}{6}\big(1+1+2+2+0+2\big) = \frac{8}{6}$$
$$\phi_2(v) = \frac{1}{6}\big(1+3+0+0+3+1\big) = \frac{8}{6}$$
$$\phi_3(v) = \frac{1}{6}\big(3+1+3+3+2+2\big) = \frac{14}{6}$$

Fontos látni, hogy a Shapley-érték azonos valószínűséggel, súllyal kezeli az egyes játékosokat (f^i_{Sh}-k). Ez az „egyenlő" kezelés a Shapley-érték normatív tulajdonsága. Természetesen használhatunk más súlyrendszert is, ekkor kapjuk a Shapley-érték egy általánosítását, az aszimmetrikus Shapley-értéket (lsd. pl. Shapley [15]).

9. megjegyzés. Világos, hogy ϕ egy lineáris leképezés, így annak konkrét formája függ mind az értelmezési tartomány, mind az értékkészlet vektorterek bázisától. Az értelmezési tartomány bázisát már rögzítettük, így továbbiakban rögzítjük az értékkészlet egy bázisát is. Ekkor ϕ jóldefiniált.

A következőkben tisztázzuk, hogy milyen tulajdonságokkal rendelkezik a Shapley-érték, illetve, milyen tulajdonságok meglétével egyenértékű a Shapley-érték. Ez utóbbi kérdés a Shapley-érték axiomatizálása, és ehhez van szükség a következő fogalmakra, eredményekre.

10. definíció. Legyen $v \in \mathcal{G}^N$ tetszőlegesen rögzített, és legyen $i, j \in N$. Ekkor az i és j játékosok ekvivalensek ($i \sim j$), ha minden olyan $S \in \mathcal{P}(N)$-re, hogy $i, j \notin S$, $v'_i(S) = v'_j(S)$.

A fenti definíció szerint két játékos ekvivalens, ha felcserélhetőek, tehát, ha egy olyan koalíciót vizsgálunk, amiben egyikőjük sincs benne, akkor mindegy, hogy melyik csatlakozik az adott koalícióhoz, mind a ketten ugyan annyi „többletet" hoznak. Könnyen látható, hogy $\sim$ ekvivalencia reláció.

11. definíció. A $v \in \mathcal{G}^N$ kooperatív játék alapjáték, ha $i, j \notin NP(v)$-ből következik, hogy $i \sim j$, ahol $NP(v) = \{i \in N \mid v'_i = 0\}$, a v játék nullajátékosainak halmaza.

Egy kooperatív játék akkor alapjáték, ha a nullajátékosokon kívül csak egyféle játékosa van. Az alapjáték fogalom használhatósága nem derül ki igazán a következőkben tárgyalásra kerülő problémáknál, de más, itt nem tárgyalt axiomatizálási koncepcióknál kulcsszerepe van.

12. következmény. *Ha $v \in \mathcal{G}^N$ alapjáték, akkor tetszőleges $\alpha \in \mathbb{R}$ esetén αv is alapjáték.*

Bizonyítás. A 11. definíció közvetlen következménye. □

Egy alapjáték tetszőleges skalárszorosa is alapjáték. Magyarán szólva, az alapjátékokat szabadon lehet „nyújtani", „zsugorítani", attól még alapjátékok maradnak.

Világos, hogy meglehetősen sok fajta alapjáték van. Ennek megmutatására nézzük a következő, [12]-ből való állítást.

13. állítás. *Legyen $v \in \mathcal{G}^N$ tetszőlegesen rögzített kooperatív játék. Ekkor $\exists v_i, \ldots, v_k \in \mathcal{G}^N$ alapjátékok, hogy $v = \sum_{i=1}^{k} v_k$.*

Bizonyítás. Elég azt megmutatni, hogy van $2^n - 1$ lineárisan független alapjáték. Legyen

$$u_T(S) = \begin{cases} 1, & \text{ha } T \subseteq S \\ 0 & \text{különben,} \end{cases}$$

ahol $S, T \in \mathcal{P}(N)$, $T \neq \emptyset$ tetszőlegesen rögzített. Az u_T játékot a T koalícióhoz tartozó egyetértési játéknak[3] nevezzük. Könnyen látható, hogy az egyetértési játékok alapjátékok.

A következő lépés annak megmutatása, hogy $\{u_T\}_{T \neq \emptyset}$ lineárisan független vektorrendszer. Indirekt tegyük fel, hogy $0 = \sum_{T \in \mathcal{P}(N) \setminus \emptyset} \alpha_T u_T$ és $\{T \mid \alpha_T \neq 0\} \neq \emptyset$. Legyen T^* a $\{T \mid \alpha_T \neq 0\}$ halmaz egy minimális eleme. Ekkor azonban

$$\left(\sum_{T \in \mathcal{P}(N) \setminus \emptyset} \alpha_T u_T \right)(T^*) = \alpha_{T^*} \neq 0 \,,$$

ami ellentmondás. □

A 13. állítás azt mondja, hogy $\mathcal{G}^N$-nek van alapjátékokból álló bázisa, amely állítást egy konkrét bázis megadásával bizonyítottunk. $\mathcal{G}^N$ alapjátékokból álló bázisainak a jellemzése, kategorizálása nyitott kérdés. A következőkben kiderül, hogy egy ilyen karakterizáció sokkal kerekebbé tenné a Shapley-érték axiomatizálásának tárgyalását.

3 Potenciál függvény

A Shapley-érték Hart és Mas-Colell [9] szerzőktől származó jellemzésével foglalkozunk a következőkben. A tárgyalás során [12]-re támaszkodunk.

14. definíció. Legyen $v \in \mathcal{G}^N$ tetszőleges kooperatív játék, és legyen $T \subseteq N$. Ekkor a v játék T-hez tartozó v^T részjátéka a következő: $v^T(S) = v(S)$, minden $S \subseteq T$-re. Látható, hogy $v^T \in \mathcal{G}^T$.

Tehát a v játék részjátékát úgy kapjuk, hogy egyszerűen kihagyunk játékosokat. Azok a koalíciók, ahol kihagyott játékosok szerepeltek, azok a részjátékban már nem lesznek, tehát ott a részjátékot nem is kell definiálni.

[3] Az elnevezés nagyon találó, hiszen a játék értéke pontosan akkor 1, ha a T koalíció tagjai „egyetértenek".

15. definíció. Legyen $\Gamma^N = \cup_{T \subseteq N} \mathcal{G}^T$, $P : A \to I\!R$ függvény, ahol $A \subseteq \Gamma^N$, és legyen tetszőleges olyan $v \in \mathcal{G}^T \subseteq A$ játékra, hogy $T \neq \emptyset$, és $|T| > 1$-re $v^{T \setminus \{i\}} \in A$,

$$P'_i(v) = \begin{cases} P(v), & \text{ha } |T| = 1 \\ P(v) - P(v^{T \setminus \{i\}}) & \text{különben} \end{cases}$$

Ekkor, ha tetszőleges olyan $v \in \mathcal{G}^T \subseteq \Gamma^N$-re, hogy $T \neq \emptyset$ és $v^{T \setminus \{i\}} \in A$ minden $i \in T$-re, fennáll, hogy

$$\sum_{i \in T} P'_i(v) = v(T) \,, \tag{2}$$

akkor P-t az A halmazon értelmezett potenciálnak nevezzük.

A potenciál tehát olyan függvény, amely a kooperatív játékok olyan részosztályán értelmezett, ahol nem rögzített a játékosok száma. Ebből következőleg, a potenciál alkalmas különböző játékosszámú játékok összehasonlítására is. (2) szerint, tetszőleges játékban a nagykoalíció értéke megegyezik az egyes játékosok potenciálra gyakorolt hatásainak összegével. Tehát azt mondhatjuk, hogy egy játékban a nagykoalíció értékét az egyes játékosok elhagyásával kapott játékok figyelembevételével kalkuláljuk.

Ehhez szükséges a következő fogalom.

16. definíció. Az $A \subseteq \Gamma^N$ játékosztály részjáték-zárt, ha tetszőleges $v \in \Gamma^N$ tetszőleges v^T, $T \subseteq N$, $|T| > 1$ részjátéka benne van A-ban.

A (2) egyenlőségből látható, hogy a potenciál fogalma akkor nem semmitmondó, ha tetszőleges játék tetszőleges részjátéka is benne van a vizsgált részosztályban. Ellenkező esetben a potenciál nem egyértelmű, így nem jóldefiniált.

A következő tétel, amely [12]-ből való, a Shapley-érték Hart és Mas-Colell–féle axiomatizációja.

17. tétel. *Legyen $A \subseteq \Gamma^N$ részjáték-zárt játékosztály. Ekkor az A-n értelmezett függvény pontosan akkor potenciál, ha tetszőleges $v \in \mathcal{G}^T \subseteq A$-ra és tetszőleges $i \in T$-re $P'_i(v) = \phi_i(v)$.*

Bizonyítás. Először megmutatjuk, hogy van A-n potenciál, sőt csak egyetlen egy van. Világos, hogy ha A részjáték-zárt, akkor $P'_i(v)$ létezik minden $v \in A$ esetén. $|T|$-n való indukcióval bizonyítunk. Legyen $v \in \mathcal{G}^T \subseteq A$, hogy $|T| = 1$. Ekkor legyen $P(v) = v(T)$.

Tegyük fel, hogy a $\mathcal{G}^T \subseteq A$, $|T| = k < n$ halmazbeli játékokra P jóldefiniált (tehát egyértelműen definiált), és legyen $v \in \mathcal{G}^S \subseteq A$ olyan, hogy $|S| = k + 1$, és

$$P(v) = \frac{v(S) + \sum_{i \in S} P(v^{S \setminus \{i\}})}{|S|} \,.$$

Ekkor
$$\sum_{i\in S}(P(v)-P(v^{S\setminus\{i\}})) = |S|P(v) - \sum_{i\in S} P(v^{S\setminus\{i\}}) =$$
$$= v(S) + \sum_{i\in S} P(v^{S\setminus\{i\}}) - \sum_{i\in S} P(v^{S\setminus\{i\}}) = v(S) \ ,$$
tehát P megfelel (2)-nek, sőt csak P felel meg.

A következő lépésben konkrétan megadjuk P-t. Legyen $v \in \mathcal{G}^T \subseteq A$ tetszőlegesen rögzített, $\{v_S\}_{S\neq\emptyset}$, $\mathcal{G}^T$ egy alapjátékokból álló bázisa (a 13. állítás miatt ilyen bázis létezik), ahol $T\setminus S = NP(v_S)$, és $v = \sum_{S\neq\emptyset} \alpha_S v_S$. Legyen továbbá
$$P^*(v) = \sum_{S\neq\emptyset} \frac{\alpha_S v_S(T)}{|S|} \ .$$
Legyen most $v \in \mathcal{G}^S \subseteq \Gamma^N$ olyan, hogy $|T| = 1$. Ekkor $P^*(v) = v(T)$, tehát ha $|T| = 1$, akkor $P^* = P$.

Legyen $v \in \mathcal{G}^T \subseteq A$ tetszőlegesen rögzített, ahol $|T| > 1$. Legyen továbbá $i \in T$ szintén tetszőlegesen rögzített, ekkor
$$P^*(v^{T\setminus\{i\}}) = \sum_{i\notin S,\, S\neq\emptyset} \frac{\alpha_S v_S(T)}{|S|} \ ,$$
így
$$P_i^{*'}(v) = \sum_{S\neq\emptyset} \frac{\alpha_S v_S(T)}{|S|} - \sum_{i\notin S,\, S\neq\emptyset} \frac{\alpha_S v_S(T)}{|S|} = \sum_{i\in S} \frac{\alpha_S v_S(T)}{|S|} \ . \tag{3}$$
Ekkor
$$\sum_{i\in T} P_i^{*'}(v) = \sum_{i\in T}\sum_{i\in S} \frac{\alpha_S v_S(T)}{|S|} = \sum_{S\neq\emptyset} \alpha_S v_S(T) = v(T) \ .$$
Tehát P^* potenciál, magyarán szólva $P^* = P$.

A következő lépés annak megmutatása, hogy tetszőleges $i \in T$-re $P_i'(v) = \phi_i(v)$. Legyen $v_S \in \mathcal{G}^N$ tetszőlegesen rögzített fent használt alapjáték. Ekkor a Shapley-érték, és az alapjáték fogalmak definícióiból (7. és 11. definíciók) következik, hogy
$$\phi_i(v_S) = \begin{cases} \dfrac{v_S(T)}{|S|}, & \text{ha } i \in S \\ 0 & \text{különben,} \end{cases}$$
így a (3) egyenlőségből és a Shapley-érték linearitásából következik, hogy tetszőleges $i \in T$-re
$$P_i'(v) = \sum_{i\in S} \frac{\alpha_S v_S(T)}{|S|} = \sum_{i\in T} \alpha_S \phi_i(v_S) = \phi_i(v) \ .$$
□

A 17. tételből látható, hogy a Shapley-féle értékelés potenciállal való jellemzésének egyetlen feltétele az, hogy a vizsgált játékosztály részjáték-zárt legyen. Ez sok esetben nagyon kézenfekvő és könnyen ellenőrizhető tulajdonság. Ennek illusztrálására nézzük az alábbi következményt.

18. következmény. *A P^* a(z)*

1. *Γ^N-en*
2. *szuperadditív játékok osztályán*
3. *szubadditív játékok osztályán*
4. *monoton játékok osztályán*
5. *additív játékok osztályán*

értelmezett függvény pontosan akkor potenciál, ha tetszőleges $v \in \mathcal{G}^T \subseteq \Gamma^N$-re és tetszőleges $i \in T$-re $P_i'(v) = \phi_i(v)$.

Bizonyítás. Minden említett játékosztály részjáték-zárt, így alkalmazhatjuk a 17. tételt. □

Fontos látni, hogy pl. a lényeges játékok osztálya nem részjáték-zárt, tehát azon a játékosztályon a potenciál nem karakterizálja a Shapley-féle értékelést.

4 Regressziós játékok

Ebben a szakaszban a lineáris regressziós[4] modellezési problémát vizsgáljuk. A célunk olyan módszert adni, amelynek segítségével értékelni tudjuk az egyes magyarázó változókat.

Legyenek η a magyarázott, és ξ_i, $i = 1, \ldots, n$ a magyarázó változók. Absztrakt formában tekintjük a modellezési feladatot, tehát nem foglalkozunk becslésekkel, feltesszük, hogy ismerjük a valószínűségi változókat, amikkel dolgozunk.

19. definíció. Legyen $N = \{\xi_1, \ldots, \xi_n\}$ a játékosok halmaza, tehát N, az n magyarázó változó halmaza.

A továbbiakban feltesszük, hogy N rögzített, amin azt értjük, hogy n magyarázó változó van a modellben. Tekintsük a következő feladatot ($S \subseteq N$ tetszőlegesen rögzített):

$$\begin{aligned} \operatorname{var}(\eta) - \operatorname{var}(\eta - \sum_{i \in S} \beta_i \xi_i) \quad &\to \quad \max \\ \beta_i \in \mathbb{R}, \quad i \in S & \end{aligned} \tag{4}$$

20. definíció. Legyen a magyarázott (η), és a magyarázó ($\xi_1, \ldots, \xi_n$) változók rögzítettek. Tetszőleges $S \in \mathcal{P}(N)$-re $v(S)$ legyen (4) megoldása.

Az egyes koalíciók értékét az általuk elért ,,illeszkedés" (4) jósága adja. Az általunk használt mérőszám az RSS-nek feleltethető meg (természetesen nem

[4]Nem lineáris regressziós problémákra teljesen analóg módon megy a felépítés, így annak tárgyalásától itt eltekintünk.

egyezik meg vele). Az alapgondolat a következő: szokásos a statisztikai irodalomban, hogy egy modell illeszkedésén a többszörös determinációs együtthatót értik. Ez a mutató azonban egy lenormázott érték (0 és 1 közé esik), ami matematikai, játékelméleti szempontból nem túl szerencsés[5]. Ugyanakkor, az itt tárgyalt megközelítésben nincsenek mintavektorok (minták), valószínűségi változókkal dolgozunk, és az RSS megfelelőjét keressük[6]. Mivel a (4) által meghatározott játékok csak egy pozitív skalár szorzóval térnek el az RSS által meghatározottól, és kúp struktúránk van, így az RSS és a (4) megközelítések ekvivalensnek tekinthetőek.

Chaven és Sutherland [2], Lipovetsky és Conklin [10] az illeszkedés mérőjeként a többszörös determinációs együtthatót használta. A (4) mérőszám előnyeként lehet felhozni (a többszörös determinációs együtthatóval szemben), hogy a többszörös determinációs együttható egy lenormázott érték, így a magyarázóváltozók abszolút értelemben nem értékelhetőek általa (több, különböző modellben szereplő magyarázó változók összevetése nehézkes).

21. következmény. *v kooperatív játék.*

Bizonyítás. Az 1, 19, 20. definíciók közvetlen következménye. □

Rögzített magyarázott és magyarázó változók esetén a v kooperatív játékban a különböző bevont magyarázó változók által meghatározott modellek adják a játékot. Tehát egy játékban több modell van, minden kooperatív játékhoz egy jól meghatározott modellcsoport tartozik.

22. definíció. A 19, 20. definíciók által meghatározott játékokat regressziós játékoknak nevezzük. A regressziós játékok osztályát $\mathcal{G}_R^N$-rel jelöljük.

A regressziós játékok tehát olyan játékok, amelyek megfeleltethetőek regressziós feladatoknak.

23. segédtétel. *A $\mathcal{G}_R^N$ játékosztály része a monoton játékok osztályának.*

Bizonyítás. A bizonyítást az olvasóra bízzuk. □

A 23. segédtétel könnyen interpretálható. Amennyiben egy rögzített modellbe egy új magyarázó változót illesztünk, akkor az új modell magyarázóereje nem lehet kisebb, mint az eredeti modellé. Vagy másképpen, egy altér és egy pont távolsága nem nőhet attól, hogy az alteret egy új vektorral bővítjük. A következőkben egy példával illusztráljuk az eddig elmondottakat.

24. példa. Legyen a kovarianciamátrix:

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

[5]Ez a tény nem derül ki az itt tárgyalt megközelítésben, de arról van szó, hogy a regressziós játékoknak nincs kúp struktúrája (az $A \subseteq \mathbb{R}^n$ halmaz kúp, ha tetszőleges $\alpha \in \mathbb{R}_+$ esetén $\alpha A \subseteq A$), ha a többszörös determinációs együtthatót használjuk az értékelésnél, ami más axiomatizációs eljárásoknál (Shapley, Young) kifejezetten hátrányos.

[6]Igazából, az itt használt absztrakt modellben nincs semmi kétség afelől, hogy mi legyen a mérőszám, hiszen nem kell becsülnünk, ismerjük a valószínűségi változókat, és egyszerűen csak legközelebbi pontokat keresünk (pontosabban legkisebb távolságokat).

A kovarianciamátrix főátlójában rendre a magyarázott (η), és a magyarázó változók (ξ_1, ξ_2, ξ_3) varianciái találhatóak. A főátlón kívüli elemek a szokásos kovarianciák. A fenti kovarianciamátrixból látható, hogy ξ_1-nek nincs közvetlen hatása η-ra.

A megfelelő v regressziós játék:

$$\begin{aligned}
&v(\{1\}) = 0\,, \quad v(\{2\}) = \frac{1}{4}\,, \quad v(\{3\}) = \frac{1}{3}\,,\\
&v(\{1,2\}) = \frac{1}{3}\,, \quad v(\{1,3\}) = \frac{1}{3}\,, \quad v(\{2,3\}) = \frac{3}{8}\,,\\
&v(\{1,2,3\}) = \frac{2}{5}\,.
\end{aligned}$$

Látható, hogy v olyan monoton játék, amely nem szuperadditív, nem szubadditív, és nem is lényeges. A v regressziós játék Shapley-értéke :

$$\left(\frac{16}{720}, \frac{121}{720}, \frac{151}{720}\right).$$

Komponensenként:

$$\begin{aligned}
\phi_1(v) &= \frac{1}{3}\,0 + \frac{1}{6}\left(\frac{1}{12} + 0\right) + \frac{1}{3}\,\frac{1}{40} = \frac{16}{720}\,,\\
\phi_2(v) &= \frac{1}{3}\,\frac{1}{4} + \frac{1}{6}\left(\frac{1}{3} + \frac{1}{24}\right) + \frac{1}{3}\,\frac{1}{15} = \frac{121}{720}\,,\\
\phi_3(v) &= \frac{1}{3}\,\frac{1}{3} + \frac{1}{6}\left(\frac{1}{3} + \frac{1}{8}\right) + \frac{1}{3}\,\frac{1}{15} = \frac{151}{720}\,.
\end{aligned}$$

Vegyük észre, hogy a 24. példában ξ_1 és η korrelálatlanok, ám ξ_1 Shapley-értéke nem nulla. Világos, hogy ha ξ_1 és η függetlenek lennének, akkor a Shapley-érték nulla lenne. Tehát, a Shapley-érték bizonyos esetekben meg tudja különböztetni a korrelálatlanságot és a függetlenséget. Ez a jelenség leírható a parciális korreláció fogalmával is, de utalva Grömpingre [8], ez inkább elvárás, mint új tulajdonság. Tehát nem arról van szó, hogy a Shapley-érték ebben a tekintetben újat hoz, hanem arról, hogy teljesíti az ebben a jelenségben megtestesülő elvárásokat.

25. következmény. *A regressziós játékok osztálya nem része sem a szuper-, sem a szubadditív, sem a lényeges játékok osztályának.*

Bizonyítás. Lsd. a 24. példát. □

Egyetlen dolgot tudunk mondani.

26. segédtétel. *Ha a magyarázó változók korrelálatlanok* ($\mathrm{corr}_{i,j} = 0$, *$\forall i, j \in N$-re), akkor a generált regressziós játék additív.*

Bizonyítás. A bizonyítást az olvasóra bízzuk. □

A 26. segédtétel nem megfordítható. Könnyen konstruálható olyan példa, ahol az adott magyarázó változók korreláltak, a generált játék mégis additív.

Más oldalról, minél erősebben korrelált két magyarázó változó, annál jobban hasonlít a Shapley-értékük egymásra. Tehát, ha két magyarázó változó teljesen korrelált, akkor a Shapley-értékük megegyezik. Ugyanakkor, könnyen megadható olyan példa, ahol két magyarázó változó korrelálatlan, a Shapley-értékük mégis megegyezik. A fentiekből is kitűnik, hogy a kovarianciamátrixra (regressziós feladatra) vonatkozó fogalmak nem karakterizálják a generált játékot. Tehát, (általában) közvetlen, a kovarianciamátrixra épülő Shapley-féle értékelésre nem látunk esélyt.

A 25. következmény és a 6. megjegyzés azt mutatja, hogy a regressziós játékok osztályán miért a Shapley-értékkel próbáljuk a magyarázó változókat értékelni. Természetesen vannak a Shapley-értéken kívül még olyan megoldáskoncepciók, amelyek nem lényeges játékok esetén is tartalommal bírnak, de „első körös”, a „legnépszerűbb” megoldás koncepciók közül a Shapley-érték az egyetlen, amely ezzel a tulajdonsággal bír.

27. tétel. *P, a Γ_R^N játékosztályon értelmezett függvény pontosan akkor potenciál, ha $P_i' = \phi_i^S$, minden $i \in N$-re.*

Bizonyítás. A 17. tétel alkalmazhatóságához, csak azt kell látnunk, hogy tetszőleges $v \in \mathcal{G}_R^T \subseteq \Gamma_R^N$, $|T| > 1$ regressziós játék, tetszőleges $i \in T$ játékos elhagyásával kapott részjátéka regressziós játék. Ezt úgy értelmezhetjük, hogy egy tetszőleges regressziós modellből, annak tetszőleges magyarázó változóját elhagyva regressziós modellt kapunk, ami egészen nyilvánvaló. $\square$

A potenciál fogalma (7. definíció) jól interpretálható a regressziós játékok esetében[7]. Tegyük fel, hogy adott egy regressziós modell. Ekkor az illeszkedése, értéke ennek a modellnek az összes magyarázó változó bevonásával elérhető illeszkedés. Hogyan osszuk szét, ezt ez értéket az egyes magyarázó változók között, más szavakkal, hogyan értékeljük ezeket a valószínűségi változókat?

A potenciál azt mondja, hogy adott magyarázó változók értéke az ő elhagyásával kapott modell illeszkedése (értéke) és az adott modellünk értékének különbsége legyen. Tehát egy magyarázó változó érjen annyit, amennyi a ceteris paribus hozzájárulása az adott modell értékéhez. Az az elvárás pedig, hogy egy modell értéke a belőle egy magyarázó változó elhagyásával kapott modellek és a teljes modell értékkülönbségeinek összege legyen, igen természetes.

A Shapley-értéknek a regressziós játékok esetén, a 7. definíció tárgyaláskor adott interpretációtól eltérő magyarázatot is lehet adni. Szokásos a statisztikai, ökonometriai feladatoknál az optimális modell kiválasztásához, tehát a magyarázó változók értékeléséhez az ú.n. *stepwise* módszereket használni. Ezek alapvetően vagy egyre bővebb, vagy egyre szűkebb magyarázó változó halmazzal rendelkező modellek összevetésével mérik meg az utolsónak bevett, vagy elhagyott magyarázó változó fontosságát (értékét).

[7]Itt jegyezzük meg, hogy a potenciál fogalma hasonlít ugyan a parciális determinációs együttható statisztikai fogalomhoz, de mind elméleti, mind gyakorlati szemszögből vizsgálva különbözik attól.

A Shapley-érték ezzel szemben az összes lehetséges modellt, tehát az összes lehetséges egyre bővülő, vagy egyre szűkülő modellsorozatot elemzi, és az egyes magyarázó változók hatásainak azonos súllyal vett összegét rendeli az adott magyarázó változóhoz. Tehát a Shapley-érték úgy interpretálható, mintha megvizsgáltuk volna a modell összes lehetséges felépítését, és csak ez után értékeljük az egyes magyarázó változókat. Ebben az értelemben a Shapley-érték nagyon hasonlít a stepwise módszerekhez, a különbség az, hogy míg a stepwise módszerek lokálisak (egy láncszemhez kötöttek), addig a Shapley-érték globális (az összes lánc, összes láncszeméhez kötődik).

Más oldalról, ahogy a 7. definíció után már említettük, a Shapley-érték tulajdonképpen egy várható érték. Ebben az értelemben tehát, a Shapley-féle értékelés azt mondja, hogy az adott magyarázó változó várhatóan mennyivel járul hozzá az adott modell magyarázóerejéhez (illeszkedés).

Ezek után, a potenciál azt mondja, hogy a Shapley-érték fent ismertetett tulajdonságú értékelése megkapható úgy, mint az egyes magyarázó változók ceteris paribus hozzájárulása az adott modell értékéhez.

Egy tulajdonságra van még szükségünk. A Shapley-érték potenciálfüggvénnyel való jellemzéséhez a regressziós játékok osztályának részjáték-zártsága kell. A játékos (magyarázó változó) elhagyhatósága meglehetősen kézenfekvő tulajdonság. Ha egy regressziós modellből kiveszünk egy magyarázó változót, természetszerűleg egy regressziós modellt kapunk, sőt magának a regressziós játéknak fogalma is erre a tulajdonságra épül.

Az előzőekből következőleg, a Shapley-érték Hart és Mas-Colell–féle axiomatizálása nagyon természetes, így a Shapley-érték használata magyarázó változók értékelésére igen kézenfekvő és védhető.

5 Alkalmazás

A következőkben (gyakorlati) példákon mutatjuk be az előzőekben ismertetett változóértékelési módszert. Alkalmazások esetén nem valószínűségi változókat kapunk, hanem csak mintákat, így a Shapley-féle értékelést is „csak" becsüljük a konkrét modellcsoport esetén. Amennyiben a magyarázó változók paramétereinek becslése torzítatlan, annyiban maga a regressziós játék is torzítatlanul becsült, így a Shapley-féle értékelés is torzítatlanul becsült. A becslések egyéb tulajdonságait nem tárgyaljuk.

A 2. szakaszban nem a regressziós négyzetösszeget, hanem annak egy pozitív számszorosát használtuk. A továbbiakban azonban az RSS-t fogjuk használni. Tekintettel azonban az előzőekben mondottakra, minden elméleti eredményünk érvényben marad.

A következőkben ismertetésre kerülő módszerek *csak egy lehetséges alkalmazásai a Shapley-féle értékelésnek, a cél csak az illusztrálás, nem több.* Az előző szakaszban leírt eredmények másképpen is használhatóak magyarázó változók értékelésére, fontos továbbá, hogy maguk a javasolt módszerek nem követelnek meg különösebb játékelméleti ismereteket.

Az első példa egy a gretl [7] programmal együtt letölthető, Ramanathan

[13] könyvhöz tartozó idősor. Az idősor 1959 és 1989 között az USA-beli Oregon állam puhafa kitermelését tárgyalja. A magyarázni kívánt változó a teljes puhafa kitermelés az adott évben milliárd board feet-ben[8] y. Öt magyarázó változónk van: a puhafa export (USA-n kívülre szállított fa) az adott évben millió board feet-ben x_1, a megkezdett lakásépítések száma az USA-ban az adott évben millió darab x_2, a papír- és faipari termékek termelési indexe az adott évben x_3, a rönkfa árak az észak-nyugat–csendes-óceáni partvidéken az adott évben \$/1000 board feet x_4, a termelői árindex (összes termékre) az adott évben x_5 (az adatok leírása megtalálható az adatfájlban).

A feladat elemzése során tengelymetszet alkalmazása látszik szükségesnek, tehát a konstans minden modellben szerepel. A modellek paraméterbecslését a hagyományos legkisebb négyzetek (a továbbiakban OLS) becsléssel végeztük. Autokorreláció, heteroszkedaszticitás esetén az OLS becsléssel kapott paraméterek nem feltétlenül a legkedvezőbbek (a becslés hatásossága veszhet el). Ebben a konkrét példában nem foglalkozunk ezzel a problémával, hiszen a cél a módszer bemutatása, működésének illusztrálása. Általában azonban az egyes becslési eljárások szabadon használhatóak az egyes modellekben, tehát elképzelhető pl., hogy a modellcsoport egyik elemében OLS-t, a másikban GLS-t (általánosított legkisebb négyzetek módszere) stb., vagy esetleg ismert eloszlások esetén maximum likelihood becslést használunk. Mivel itt a mintákból az igazi modellparamétereket csak becsülni tudjuk, így a 4. szakaszban leírtaknak megfelelően, a megfelelő modellek magyarázóerejét becsüljük meg.

A Shapley-féle értékelés:

$$(5.3326; 5.2641; 7.2784; 1.0128; 5.7077)\,, \tag{5}$$

tehát a változók fontossági sorrendben: x_3, x_5, x_1, x_2, x_4. Azt is megállapíthatjuk, hogy pl. x_3 közel nyolcszor olyan fontos, mint x_4.

Ha két magyarázó változó erősen korrelált, akkor a Shapley-féle értékelésük közel esik egymáshoz. Tekintsük meg ezért a magyarázó változók korrelációs-mátrixát:

$$\begin{pmatrix} 1.0000 & 0.2015 & 0.9413 & 0.6833 & 0.8492 \\ 0.2015 & 1.0000 & 0.2011 & -0.0428 & -0.0110 \\ 0.9413 & 0.2011 & 1.0000 & 0.6779 & 0.9154 \\ 0.6833 & -0.0428 & 0.6779 & 1.0000 & 0.6842 \\ 0.8492 & -0.0110 & 0.9154 & 0.6842 & 1.0000 \end{pmatrix}. \tag{6}$$

Látható, hogy az x_1, x_3, x_5 változók erősen korrelálnak egymással, tehát mindhárom együttes szerepeltetése nem feltétlenül indokolt. Az x_2 változó azonban alig korrelál a többi magyarázó változóval, így annak fontossága felértékelődik.

Tekintsük azt a leszűkített regressziós játékot, ahol csak az x_1, x_2, és x_4 magyarázó változók a játékosok, és az egyes modellek értékeit úgy számoljuk, hogy az x_3, x_5 változók mindig szerepelnek a regresszióban. Tulajdonképpen

[8] 1 board feet = kb. 3,744 cm^3.

három magyarázó változós modellcsoportot elemezünk, melyet feltételes (x_3, x_5 változók minden regresszióban szerepelnek) értékelésnek nevezhetünk.

A kapott Shapley-féle értékelés (az értékek összege: az összes magyarázó változók együttes magyarázóerejéből levonva a feltételül szabott két magyarázó változó x_3, x_5 által együttesen elért magyarázóerőt, $RSS_N - RSS_{\{x_3,x_5\}}$):

$$(0.3429; 0.6652; 0.6724) \ . \tag{7}$$

Megváltozott a meghagyott változók erősorrendje: itt x_4 a legerősebb, míg az eredeti Shapley-féle értékelésben a leggyengébb volt. Másrészt látható, hogy ez a három változó összesen is csak 1.6805-tel tudja növelni a magyarázóerőt[9] (6.83%), ami nagyon csekély. Látható továbbá, hogy az eredeti modellben x_2 „bevétele" 5.2641-gyel növeli a magyarázóerőt (21.4%), ami jelentősnek tűnik. Kiszűrve azonban x_3, x_5 magyarázó változók hatását jelentős visszaesést tapasztalunk, ami arra utal, hogy x_2 már korántsem viselkedik függetlenül az x_3, x_5 együttestől. Érdekes azonban, hogy x_4 egészen sokat megőrzött értékéből, így x_4 fontosabb változónak tűnik, mint azt az eredeti (nem feltételes) értékelés sugallta (lsd. a 24. példa utáni szövegrészt).

A második példánk szintén egy a gretl programmal együtt letölthető, Ramanathan könyvhöz tartozó idősor. Az idősor az USA-ban eladott új autók állományának, negyedéves bontásban, elemzésére szolgál. A magyarázni kívánt változó az eladott új autók száma 1000 db-ban y. Öt magyarázó változónk van: népesség millió fő x_1, az egy főre eső elkölthető jövedelem ezer dollárban, 1982-es bázisévvel x_2, az új autók árindexe 1982-es bázisévvel x_3, az elsődleges, bankok által alkalmazott kamatláb (%) x_4, munkanélküliségi ráta (%) x_5 (az adatok leírása megtalálható az adatfájlban).

Az előző példához hasonlóan csak OLS becslést használunk és tengelymetszetet teszünk a modellekbe. A modellszelekciós kritériumok két modellt javasoltak:

$$y = 9541.4 - 57.9x_1 - 595.9x_2 - 34.3x_4 \ , \tag{8}$$

$$y = 13386 - 82x_1 + 659x_2 + 11x_3 - 39x_4 \ . \tag{9}$$

A Shapley-féle értékelés ($TSS = 5719200$):

$$(692900; 559000; 593000; 1125600; 546800) \ , \tag{10}$$

tehát a változók fontossági sorrendben: x_4, x_1, x_3, x_2, x_5. Egy új módszer illusztrálása céljából végezzünk további elemzéseket ezen a példán.

Tekintsük a magyarázó változók korrelációs-mátrixát:

$$\begin{pmatrix} 1.0000 & 0.9568 & 0.9797 & -0.0324 & -0.2195 \\ 0.9568 & 1.0000 & 0.9059 & -0.1484 & -0.4563 \\ 0.9797 & 0.9059 & 1.0000 & 0.0765 & -0.0887 \\ -0.0324 & -0.1484 & 0.0765 & 1.0000 & 0.2949 \\ -0.2195 & -0.4563 & -0.0887 & 0.2949 & 1.0000 \end{pmatrix} . \tag{11}$$

[9] TSS=30.8184

Az x_1, x_2, x_3 változók erősen korrelálnak egymással, míg az x_4, x_5 változók gyengén korreláltak. Az x_1, x_2, x_3 változókat vonjuk össze egy szuperváltozóba, amit úgy kapunk, hogy csak azokat a koalíciókat engedjük meg, ahol a fent említett három változó együtt szerepel vagy együtt nem szerepel. Az így kapott három ,,változós" modellcsoport Shapley-féle értékelése (az első érték az új szuperváltozó értékelése):

$$(2036600; 1103800; 376800) . \tag{12}$$

Látható, hogy ha az eredeti értékelésben (lsd. (10)) az első három változó (amiket összevontunk) értékeléseit összeadjuk, akkor kisebb értéket kapunk (32.26%), mint az új modellben, ahol ez a három magyarázó változó együttes erejét mérjük (35.61%). Ez a jelenség azt mutatja, hogy az eredeti modellcsoport Shapley-féle értékelése alulbecsüli x_1, x_2, x_3 változók együttes fontosságát.

6 Összegzés

A cikk célja, hogy megmutassuk, játékelméleti, pontosabban kooperatív játékelméleti fogalmakkal kezelhetőek változóértékelési problémák. *Semmiképpen sem tekintjük a cikket teljesnek abban az értelemben, hogy teljes körűen és minden pontban helyesen használta az ökonometriai, statisztikai módszereket, bár nem is ez volt a cél.* Azt céloztuk meg, hogy elméletileg megalapozzuk, és példákon keresztül megmutassuk, hogy az ökonometria, statisztika területen is használhatóak játékelméleti fogalmak, eredmények.

A cikk fő eredménye, hogy a regressziós játék fogalmának bevezetésével egy olyan jól definiált játékosztályt kaptunk (22. definíció), amely elméleti szempontból jól jellemezhető. A jellemzések az alkalmazások során (5. szakasz) jól mutatják, hogy mely utakon érdemes elindulni, mely játékelméleti fogalmak, eredmények átültetésére van esély. Elméleti értelemben a 27. tétel a cikk fő eredménye, amely azt mutatja meg, hogy a Shapley-érték alkalmazása magyarázó változók értékelésére regressziós modellekben védhető módszer.

Ami a további kutatásokat illeti, gazdag lehetőségeket látunk konkrét, *valós modellezési*[10], és elméleti ökonometriai (módszertani) elemzéseknek a tárgyalt területen. A játékelmélet oldaláról nézve, további megoldás koncepciók használata, más axiomatizálási megközelítések érvényességének vizsgálata, illetve újabb ökonometriai, statisztikai problémák játékelméleti megközelítésében látunk kutatási lehetőségeket.

[10] Terjedelmi okokból nem került bele ebbe a munkába a Shapley-érték alkalmazása tényleges modellszelekciós problémákra. Ennek a területnek tárgyalása elméleti szempontból az ú.n. semi-value–k ismertetését, gyakorlati szemszögből pedig a különböző modellszelekciós kritériumok bevezetését, tárgyalását igényelné, ami több, mint megduplázta volna a cikk terjedelmét.

Irodalom

1. Albrecht, J., D. Francois, K. Schoors: A Shapley decomposition of carbon emissions without residuals, *Energy Policy*, **30**, 727–736. (2002)
2. Chevan, A., M. Sutherland: Hierarchical Partitioning, *The American Statistician*, **45**, 90–96. (1991)
3. Cox, L. A. Jr.: A new measure of attributable risk for public health applications, *Management Science* **31**, 800–813. (1985)
4. Forgó F., J. Szép, F. Szidarovszky: *Introduction to the theory of games: concepts, methods, applications*, Kluwer Academic Publishers, Dordrecht, (1999)
5. Forgó F., Pintér M., Simonovits A., Solymosi T.: *Játékelmélet* (elektronikus jegyzet), http://www.bke.hu/~opkut/letoltheto_anyagok.html (2006)
6. Gefeller, O., M. Land, G. E. Eide: Averaging Attributable Fractions in the Multifactorial Situation: Assumptions and Interpretation, *Journal of Clinical Epidemiology* **51**, 437–441. (1998)
7. A gretl programcsomag, http://gretl.sourceforge.net/
8. Grömping, U.: Estimators of Relative Importance in Linear Regression Based on Variance Decomposition, *The American Statistician* **61**, 139–146. (2007)
9. Hart, S., A. Mas-Colell: Potential, value, and consistency, *Econometrica* **57**, 589–614. (1989)
10. Lipovetsky, S., M. Conklin: Analysis of Regression in Game Theory Approach, *Applied Stochastic Models in Business and Industry* **17**, 319–330. (2001)
11. Owen, G.: *Game Theory*, Academic Press, Inc. (1982)
12. Peleg, B., P. Sudhölter: *Introduction to the Theory of Cooperative Games*, Kluwer Academic Publishers, Boston/Dordrecht/London (2003)
13. Ramanathan, R.: *Bevezetés az ökonometriába alkalmazásokkal*, Panem Könyvkiadó, Budapest (2003)
14. Shapley, L. S., *A Value for n-Person Games*, Contributions to the Theory of Games Volume II (Annals of Mathematical Studies **28**, szerk.: Kuhn, H. W.–Tucker, A. W.) 307–317. (1953)
15. Shapley, L. S.: *A Comparison of Power Indices and a Nonsymmetric Generalization* P-5872, The Rand Corporation, Santa Monica, CA. (1977)
16. Shorrocks, A. F.: *Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value*, working paper (1999)
17. Stufken, J.: *On Hierarchical Partitioning*, The American Statistician **46**, 70–71. (1992)
18. Wan, G.: *Poverty Accounting by Factor Components*, United Nations University Research Paper No. 2006/63 (2006)
19. Zhang, Y., G. Wan: *Why do Poverty Rates Differ From Region to Region*, United Nations University Research Paper No. 2005/56 (2005)

REGRESSION GAMES

A solution of a TU coalitional game is an allocation of the payoff (utility) achieved by the players together. In a regression model, the evaluation of the explanatory variables can be an allocation of the overall fit got by those together. Therefore a regression model can be taken as a TU coalitional game, in which the explanatory variables are the players. The various solution concepts of TU coalitional games can help the modeler in recognizing the important explanatory (regressor) variables and make her possible to understand the examined model more. In this paper we build a solid mathematical background for this problem. We use the Shapley value for evaluating the explanatory variables in regression models.