



BUSINESS & DIPLOMACY REVIEW

**1. évfolyam, 2. szám
2023/2**

TARTALOM

- Árpád Bánhalmi: Hierarchical Clustering Combined with Neural Networks
- Berta Péter: Luxusfogyasztás, elrendezett házasság és a családi élet politikai gazdaságtana
- Dalma Boldog: Media Awareness among Students of the Budapest Business University
- Bölcskei Attila: Multidimenzionális adatvizualizáció dinamikus Chernoff-arcokkal
- Papp-Váry Árpád Ferenc: Hogyan látnak minket, kutatókat, és miként javíthatunk rajta? – A tudományos kutatókról kialakult sztereotípiák a nemzetközi és hazai felmérések alapján
- Agneš Slavić, Maja Strugar Jelača, Nemanja Berber, Dimitrije Gašić: The importance of students' soft skills based on pilot research results from Serbia
- György Szondi: The Evolution of Public Diplomacy – Definitions and Goals

Business & Diplomacy Review

Kiadó: Budapesti Gazdasági Egyetem

Szerkesztőség

1165 Budapest, Diósy Lajos u. 22-24.

<https://uni-bge.hu/hu/business-diplomacy-review>

Főszerkesztő

Sáringér János

Felelős szerkesztő

Csekő Katalin

Szerkesztőbizottság tagjai

Csekő Katalin, elnök

Bába Iván

Ferkelt Balázs

Juhász Tímea

Nagy Milada

Marinovich Endre

Marosán Bence Péter

Murádin János Kristóf

Válóczi Marianna

Tipográfia és tördelés

Nagy és Heteyi Kft.

ISSN 3004-0116

HU ISSN 3004-0116

Business & Diplomacy Review állandó lektorai

Prof. Dr. Bajomi-Lázár Péter, PhD

Dr. habil. Buday-Sántha Andrea, PhD

Dr. Csekő Katalin, PhD

Dr. habil. Domonkos Endre, PhD

Dr. Ferkelt Balázs, PhD

Dr. Gyene Pál, PhD

Dr. Harsányi Dávid, PhD

Dr. Horváth Annamária, PhD

Dr. habil. Krasztev Péter, PhD

Dr. habil. Juhász Tímea, PhD

Dr. habil. Marosán Bence Péter, PhD

Dr. habil. Murádin János Kristóf, PhD

Dr. Nagy Milada, PhD

Dr. Pólya Éva, PhD

Dr. habil. Sáringer János, PhD

Dr. Válóczy Marianna, PhD

Dr. Zelena András, PhD

TARTALOM

Hierarchical Clustering Combined with Neural Networks <i>Árpád Bánhalmi</i>	5
Luxusfogyasztás, elrendezett házasság és a családi élet politikai gazdaságtana <i>Berta Péter</i>	20
Media Awareness among Students of the Budapest Business University <i>Dalma Boldog</i>	44
Multidimenzionális adatvizualizáció dinamikus Chernoff-arcokkal <i>Bölcskei Attila</i>	61
Hogyan látunk minket, kutatókat, és miként javíthatunk rajta? – A tudományos kutatókról kialakult sztereotípiák a nemzetközi és hazai felmérések alapján <i>Papp-Váry Árpád Ferenc</i>	71
The importance of students' soft skills based on pilot research results from Serbia <i>Agneš Slavić, Maja Strugar Jelača, Nemanja Berber, Dimitrije Gašić</i>	96
The Evolution of Public Diplomacy – Definitions and Goals <i>György Szondi</i>	121

MULTIDIMENZIONÁLIS ADATVIZUALIZÁCIÓ DINAMIKUS CHERNOFF-ARCOKKAL

Bölcseki Attila

Absztrakt

A gazdasági, üzleti, műszaki világ és általában az élet növekvő adatmennyiséget generál, melynek feldolgozása, az adatok vizualizációja folyamatos kihívások elé állítja a szakembereket. Ennek eredményeképpen számos nem-standard ábrázolási módszer jelent meg a statisztikai elemzésekben. Jelen dolgozat tárgyát a k -dimenziós adatsorok szemléletessé tétele képezi, melyet az éppen 50 esztendőös ún. Chernoff-arcok alkalmazásával végezzük, az arcok egy újszerű, dinamikus geometriai implementációjának prezentálásával. A kifejlesztett technikát messzemenően javasoljuk a Kollégák szíves felhasználására mind idősorok, mind kérdőívek elemzése, mind egyéb komplex minták bemutatására és elemzésére.

Kulcsszavak: multidimenzionális adatvizualizáció, Chernoff-arc, dinamikus geometria, statisztikai elemzés

Abstract

The economic, business, technical world and life in general is producing an increasing amount of data, and the processing and visualisation of this data is a constant challenge for professionals. As a result, a number of non-standard visualisation methods have emerged in statistical analysis. The subject of this paper is the visualization of k -dimensional data sets, which is based on the 50-year-old so-called Chernoff faces, for which a novel dynamic geometric implementation of the faces will be introduced. The developed technique is highly recommended for the use of colleagues for the presentation and analysis of time series, test series and other complex patterns.

Keywords: multidimensional data visualization, Chernoff face, dynamic geometry, statistical analysis

A multidimenzionális adatvizualizáció szerepe és módszerei

Vizualitás, adatvizualizáció és matematika

Ebben a dolgozatban a vizualizáció tág fogalomkörét az információvizualizációra szűkítjük le. Ez azt jelenti, hogy azzal a céllal láttatunk, hogy elsősorban információt közöljünk vele, és ne csak szórakoztassunk általa. Az információ gazdasági, üzleti, műszaki stb. célú vizualizációja azt követeli meg, hogy az adatokat rendszerben ábrázoljuk.

Az adatvizualizáció egy elterjedt megközelítés [1] szerint nem más, mint leképezés az adat és annak vizuális megjelenítése között, pontosabban leképezés az agy két modalitása, nevezetesen a matematikai és a vizuális között. Manovich szerint az adatvizualizáció célja az, hogy nagy mennyiségű adatot mások számára áttekinthetővé tegyünk. Ennek nagy szerepe van a statisztikusok munkájában, amennyiben az adatgyűjtés után, az adatfeldolgozási fázisban az adatokat áttekinthetővé téve mintázatokat, összefüggéseket mutathat a változók között, s ennek nyomán választható ki az adekvát matematikai-statisztikai módszer.

Az adatvizualizáció tehát olyan módszer, amely számok (adatok) közötti kapcsolatokat mutat be vizuálisan, abból a célból, hogy például folyamatok trendjeit illusztráljuk, mintázatokat ismerjünk fel, megtaláljunk kiugró értékeket stb.

Minderre azért is van növekvő igény, mivel az adatok óriási és egyre fokozódó ütemben termelődnek (Big Data jelenség). Becslések szerint 2025-ben nagyjából egy hét alatt annyi adat keletkezik majd, mint 2013-ban összesen (4,3 zettabájt), holott már ez is több mint 850-szerese volt a valaha írt összes könyvben tárolt adatmennyiségnek (5 exabájt) [2].

A vizuális megjelenítés előnyeit az alábbiakban foglalhatjuk össze: segíti a főbb következtetések megjegyzését; javítja az elemző azon képességét, hogy kapcsolatokat vegyen észre; a jelenségek könnyebben bemutathatók, elmesélhetőek, ha ábrázolva vannak; illetve lehetővé teszi bonyolult számítások áttekintését.

A sok paraméteres problémák vizualizációs módszereiről

Ha adatvizualizációról van szó, szokás megkülönböztetni standard és nem-standard ábrázolási módszereket; előbbieken értve a leggyakrabban és legrégebben használt eljárásokat: a kör- és oszlopdiaagramot, a hisztogramot, az empirikus eloszlásfüggvényt és gyakorisági poligont. Nem-standard ábrázolási eljárásokkal több változót is figyelembe vevő, vagy a változásokat jobban kiemelő vizualizációkra is lehetőség nyílik.

A következőekben, az értelmezési keretet szűkítve, arra a nem-standard problémára keressük a választ, hogy miként lehet a k -dimenziós tér pontjait úgy megjeleníteni, hogy a változók közötti, feltételezett komplex összefüggések is megjelenjenek. A vizualizáció céljaul azt tűzzük ki, hogy a valószínűsíthető kapcsolatokat feltárjuk, megsejtsük, melyeket azután a megfelelő statisztikai módszerekkel igazolni is lehet. A k -dimenziós vektorok, amelyekkel dolgozunk olyan mintáknak felelnek meg, amelyek egy adott alany (jelenség, időpont stb.) k db különböző ismérv szerinti értékeit tükrözik.

A probléma megoldása számos kreatív módszer megalkotásához vezetett, melyek közül mindössze kettőt emelünk ki.

Az egyik legkorábban javasolt módszer k változó hatásának egyetlen ábrán való ábrázolására az Andrews által [3] javasolt technika, amely Fourier-sor alakot használ a megjelenítésre, ahol a megfelelő együtthatókat a tetszőleges méretű k -dimenziós vektor x_1, x_2, x_3, \dots koordinátái adják az alábbiak szerint. (1)

$$f_{\vec{x}}(t) = x_1/\sqrt{2} + x_2 \sin t + x_3 \cos t + x_4 \sin 2t + x_5 \cos 2t + \dots; t \in (-\pi, \pi)$$

Másrészt, az egyik legszórakoztatóbb és legprovokatívabb próbálkozás arra, hogy nagy mennyiségű változó hatását egyszerre érzékeltesük, Hermann Chernofftól származik [4]. Az eljárás lényege az, hogy a k -dimenziós vektor egy képregény figura arcának különböző sajátságait (az arc formája; a szem, szemöldök, száj és orr formája, elhelyezkedése) vezérli. Ezáltal a jellegzetességek kiválóan szemléltethetővé válnak.

A kiinduló feltevés az volt, hogy az emberi elme apró különbségeket is képes detektálni az emberi arcon, így lehetőség van finom árnyalatok észlelésére és arra is, hogy egy arcon a változást észrevegyük. Nem jelent gondot az emberi elme számára az, ha a rajzolt arc nem egészen realisztikus, hiszen a rajzfilmek, animációk figuráinak

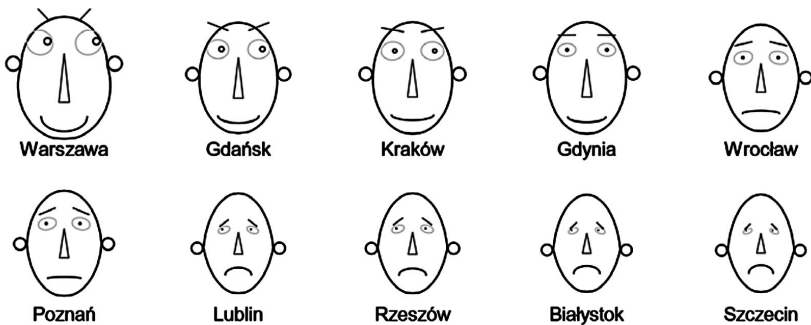
érzelmeit is könnyen át tudjuk élni. Az arcok ezen kívül könnyebben megjegyezhetők: az érzelmek, a jellegzetességek könnyebben idézhetők fel a későbbiekben, akár sok idő elteltével is.

Az arcokkal történő adatvizualizáció lehetővé teszi például, hogy klaszter analízist végezzünk, ilyenkor az összetartozó adatmintákat csoportosítjuk. Viszont időszorelemzésnél, amikor az időben egymást követő k-dimenziós adatsorokat különálló arcokként jelenítjük meg, az arc karakterének változásával azonosíthatók a folyamatban előforduló töréspontok. Könnyen lehet kiugró értékeket, mással kapcsolatba nem hozható adatsorokat is azonosítani: az arc karaktere ilyenkor minden másétól jelentősen el fog térni.

Ha a vizualizáció kínálta előnyöket nézzük, akkor biztosan állítható, hogy a Chernoff-arcok segítik a főbb következtetések megjegyzését és nagyban segítenek az elemzőnek abban, hogy kapcsolatokat azonosítson.

A módszer éppen idén lett 50 éves, de érdekessége miatt még nem szorul ráncfelvarráásra; a mai napig aktuális és hasznos eszköz az adatelemzésben, erre utal a [4] publikáció több mint 2300 hivatkozása is (2023. szeptemberi adat).

Az alábbi két ábrával a közelmúlt tudományos kínálatából merítve bemutatjuk egyrészt az implementációk, másfelől a felhasználások változatosságát.



1. ábra: Példa Chernoff-arc ábrázolásra 2022-es ingatlanpiaci elemzés céljából, ahol a paraméterek a különböző időpontokban mért értékeknek felelnek meg
Forrás: [5], Fig. 5, részlet



Adminisztrátor



Diák

2. ábra: Chernoff-arcok egy 2022-es manilai felsőoktatási kérdőív feldolgozására a [6] publikációból

Forrás: [6], Fig.1

A Chernoff-arcok geometriájáról

Hermann Chernoff eredeti munkájában [4] 18 olyan paramétert vezetett be, amelyek segítségével az arc szerkezete változtatható. Ez azt is jelenti, hogy minden általa előállított arc megfeleltethető egy 18 dimenziós folytonos eseménytér egy mintájának. Chernoff paraméterei a fenti cikk függelékéből olvashatók is. Ezek alapján megállapítható, hogy mely paraméter arc mely területét vezérli. Leszögezzük, hogy Chernoff (és a későbbiekben mi is ezt fogjuk tenni) csakis tengelyesen szimmetrikus arcokat készített, azaz az arc bal és jobb oldala egymás tükörképe. Emiatt a paraméterek zöme csak az arc egyik felének geometriáját határozza meg, a másik oldal tükrözéssel adódik. Példaként bemutatjuk a [4] cikk egyik adatvizualizációját, az első Chernoff-arcokat:

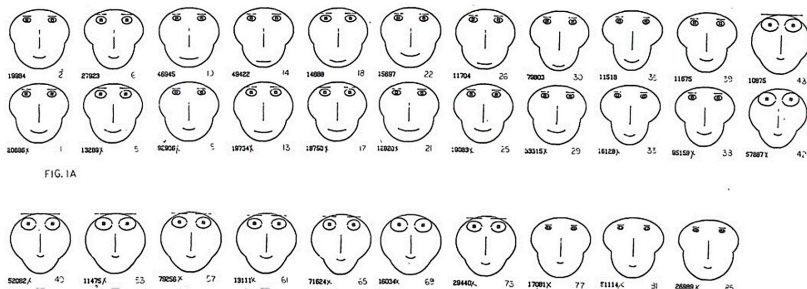


FIG. 1A

3. ábra: Példa az eredeti Chernoff-arcokra

Forrás: [4], 362. oldal (részletek az ottani 1/A és 1/B ábrákból)

Feladatunknak éreztük, hogy a dinamikus geometriai szoftverek adta lehetőségeket kihasználva újra gondoljuk a Chernoff-arcok felépítését és lehetővé tegyük, hogy a felhasználó egyrészt a csúszkákat változtatva könnyen beállíthasson adatvizualizációkat az adatsoraihoz, másrészt magukat a változtatásokat valós időben nyomon tudja követni, a látványban örömet lelje. Ilyen irányú törekvés ismereteink szerint ezidáig nem történt a témában.

Célként egy, az eredetinél realiztikusabb, részletesebb, nagyobb felhasználói élményt nyújtó, ugyanakkor a képregények báját megőrző konstrukció kidolgozását foglalmaztuk meg. Eszközként a szabadon elérhető GeoGebra dinamikus geometriai szoftvert használtuk [7], mely számos geometriai probléma szemléltetésében már bizonyított [8, 9].

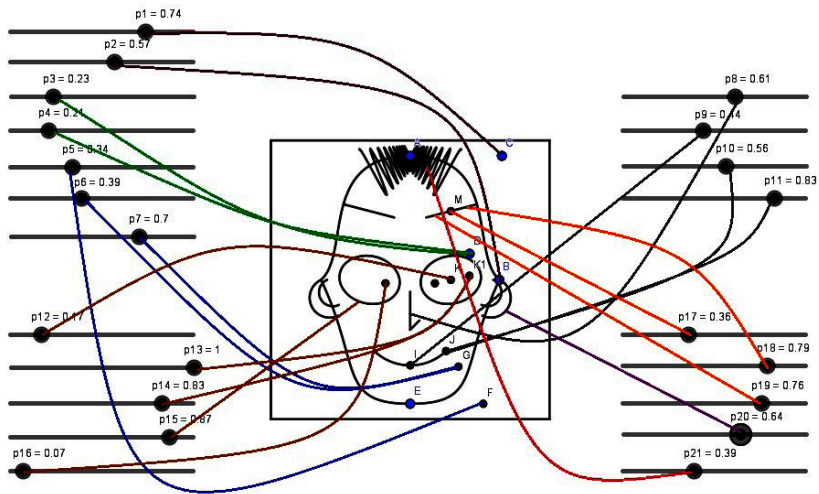
Megoldásunkban az arc formáját szintén tükörszimmetrikusnak tekintettük, melyet azonban ellipszisek helyett, a nagyobb fokú alakíthatóságot lehetővé tévő köbös spline görbékkel valósítottunk meg, melyhez 7 paraméterre volt szükség. Az orr középpontja megoldásunkban a szimmetriatengelyen van, a középpont és legalsó pont közötti felső negyedelőpontban. Formája nem egyszerű szakasz, hanem egy hosszabb függőleges, és egy ezzel 45 fokot bezáró, negyed olyan hosszú szakasz együttese. A száj leírásához szükséges három paraméter a körív forma középpontjának és egyik végpontjának helyzetét jellemzi. A szem ellipszis alakjának leírásához négy paraméterre, a pupillához még egy továbbira volt szükség. A szemöldök megadása újabb három paraméterrel lehetséges, melyek a magasságot, hosszát

és ferdeséget jellemzik. Megoldásunkban többletként még füleket (egy paraméterrel) és haját is alkottunk, mely utóbbi leírását az

$$f(x) = x \sin\left(\frac{1}{x}\right)$$

függvény egy affin transzformálta segítségével oldottuk meg, megfelelő paraméterezéssel. (2)

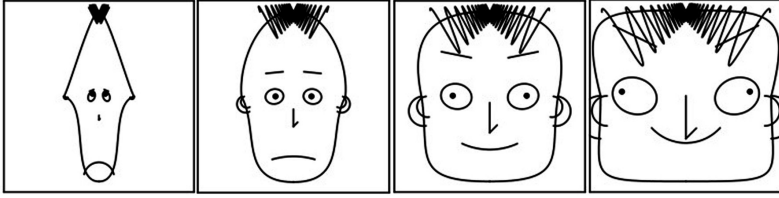
A paraméterek és az arc meghatározó adatai közötti kapcsolatokat szemléletesen az alábbi ábrán lehet követni.



4. ábra: A 21 paraméter kapcsolata az arc meghatározó adataival
Forrás: saját szerkesztés

Hangsúlyozzuk, hogy a 21 paraméter mindegyike a másiktól függetlenül hangolható, így az arcok úgyszólván végtelen változatosságban állíthatók elő.

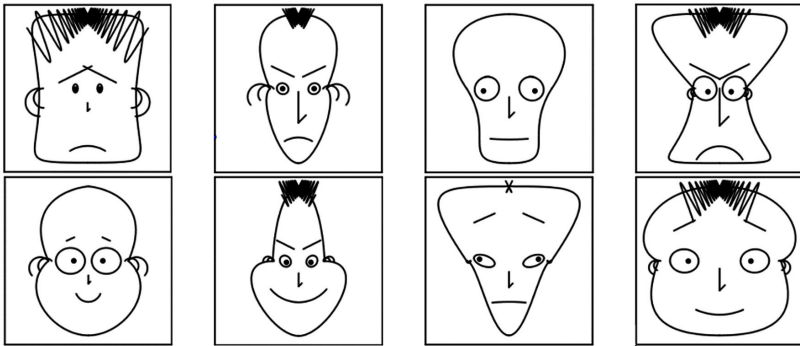
Illusztrációképpen bemutatunk egy ábrát, ahol a paramétereket mind egyforma értékekre állítottuk be, mégpedig sorban 10, 40, 70 majd végül 100%-ra.



5. ábra: Arcok, ahol az összes paraméter értéke azonos, mégpedig rendre 10, 40, 70 és végül 100%

Forrás: saját szerkesztés

További illusztrációként néhány randomizáltan választott paraméter segítségével előálló arcot mutatunk be.



6. ábra: Random paraméterekkel megjelenített Chernoff-arcok

Forrás: saját szerkesztés

Összefoglalás

A bemutatott, éppen 50 éves jubileumát ünneplő Chernoff-arcok módszerére jelen publikáció egy új, dinamikus geometriai eljárást kínál. Maga az alapelv éppen az ellenkezője a képek digitalizálásának, amely ezekből számokat generál abból a célból, hogy mondjuk a betáplált arcok megkülönböztethetők legyenek. Chernoff-arcok segítségével éppen számokból készítünk képeket, arcokat, hogy a sorsorokban rejlő különbségek láthatóvá, könnyebben értelmezhetővé váljanak.

Úgy érezzük, hogy a bemutatott eljárással segíthet nyújthatunk mindazoknak a kollégáknak, akik nagy mennyiségű, összetartozó adatot szeretnének feldolgozni. Segítség lehet például, amikor egy kérdőívre adott válaszok jelentik az ábrázolandó (legfeljebb 21 db) változót. A válaszokból készült adatsorok egy-egy arcként megjelenítve ilyenkor összefüggések és mintázatok megsejtését teszik lehetővé. Ugyancsak felhasználhatóak a Chernoff-arcok amikor egy adatot az idő függvényében szeretnénk ábrázolni, legfeljebb 21 időpontban felvett értékeken keresztül. Ilyenkor az arcok az idősorban hasonlóan viselkedő adatkészleteket tudják jelezni.

Egy másik alkalmazásban komplex adatsorokat feleltethetünk meg egy-egy arcnak, melyeket akár időbeli lefutásukkal is vizsgálhatunk. Ilyenkor az arc változásával töréspontokat, ugrásszerű változásokat lehet észrevenni, ábrázolni.

Végezetül, alkalmas a technika minden olyan esetben, amikor nagy mennyiségű összetartozó adatot egyetlen, sokat kifejező ábrába akarunk sűríteni, mégpedig mindezt szórakoztató módon.

Hivatkozások

1. Manovich, L. (2020). *Cultural analytics*. Mit Press.
2. Arbia, E. (2023). *The regulation of digital platforms: the case of pagoPA*. (Doctoral thesis)
3. Andrews, D. F. (1972). Plots of high-dimensional data. *Biometrics*, 125-136.
4. Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. *Journal of the American statistical Association*, 68(342), 361-368.
5. Kozicki, B., Stajniak, M., Magniszewski, M., Lorek, M., & Mitkow, S. Z. (2022). Multidimensional Analysis of Real Estate Prices in Seventeen Cities in Poland in Terms of Economic Security. *Journal of Security and Sustainability Issues*, 12(1).
6. Lopez-Conde, R. C., Nalipay, J. N., Ancho, I. V., Abulon, E. L. R., Rungduin, T. T., Montealegre, M. A. C., & Madronero, J. A. (2022). Utilizing Chernoff Faces in Modeling Responses in the Evaluation of Trimester Scheme Impementation. *Dalat University Journal of Science*, 125-138.

7. GeoGebra, <https://www.geogebra.org/>
8. Talata, I., Bölcskei, A., Budai, L., & Keresztes, É. R. (2023). Lineáris programozási feladatok vizualizációja GeoGebrával. In: Ország, Adrienn; Baják, Szabolcs (szerk.) I. Csernyák László konferencia közleményei Budapest, Magyarország : Budapesti Gazdasági Egyetem (BGE), 24-39.
9. Talata, I., & Bölcskei, A. (2022). Elemi majdnem-Johnson-poliéderek és számítógépes modellezésük. DIMENZIÓK: MATEMATIKAI KÖZLEMÉNYEK, 3-14.

Publikálás a Business & Diplomacy Review folyóiratban

A Folyóirat az alábbi témákhoz kapcsolódó kéziratokat fogad be lektorálásra:

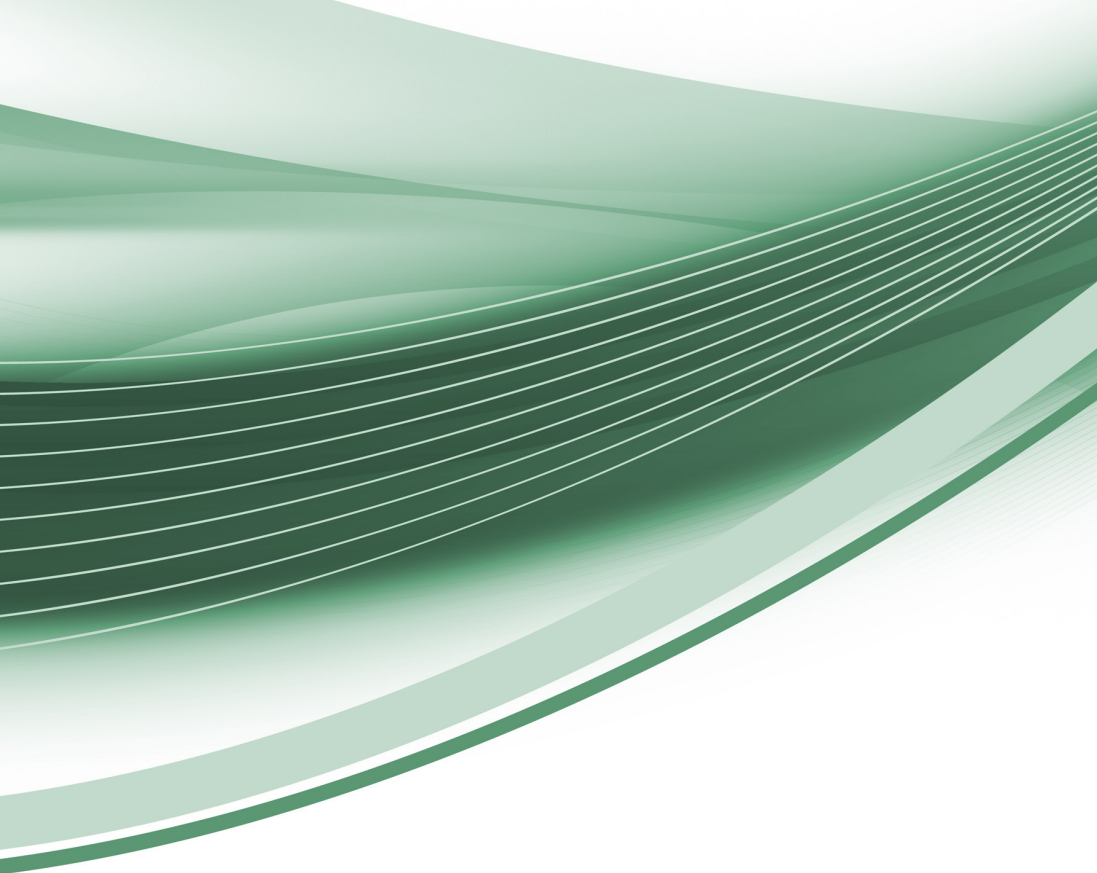
- gazdaság,
- kereskedelem,
- pénzügyek,
- nemzetközi kapcsolatok,
- diplomácia,
- nemzetközi kommunikáció
- társadalomtudományok;
- közgazdasági tudományok.

A tanulmány absztraktja minden esetben angol nyelvű. A tanulmány magyar vagy angol nyelvű, amelynek terjedelme 30 000–40 000 leütés (az absztrakttal együtt).

A kéziratok előzetes befogadásának feltételei:

- a kézirat és annak szerzői megfelelnek a Folyóirat etikai szabályainak;
- a kézirat, illetve ahhoz tartalmában nagyon hasonló tanulmányt még nem publikálták;
- a benyújtott kézirat megfelel a formai követelményeknek.

A Folyóirat a szerzőknek tiszteletdíjat nem fizet. A Folyóirat minden egyes befogadott kézirat esetében kettős vaklektorálást alkalmaz, ami azt jelenti, hogy az anonimalizált anyagot a szerzők által nem ismert lektorok értékelik. A Folyóirat csak abban az esetben fogad be kéziratot publikálásra, ha azt mind a két vaklektor publikálásra ajánlja, és vaklektor(ok) által kért javításokat/kiegészítéseket a szerző(k) végrehajtotta/ák. Amennyiben az egyik lektor javításokkal publikálásra ajánlja a kéziratot, míg a másik nem, akkor a javítások után a témában jártas újabb vaklektornak kell értékelnie az anyagot. Akkor minősül egy tanulmány tartalmában nagyon hasonlóknak egy korábbi tanulmányhoz képest, ha azok egyezősége 60% felett van. A szerzők minden egyes esetben kötelesek a vaklektorok által írt kifogásokra/javaslatokra tételesen írásban reagálni.



Kiadja: Budapesti Gazdasági Egyetem, Budapest

ISSN 3004-0116
HU ISSN 3004-0116