TÉRÖKONOMETRIA*

VARGA ATTILA

A térökonometria (spatial econometrics) az ökonometria azon részterülete, mely a keresztmetszeti és paneladatokra épülő regressziós modellekben a térbeli interakció (térbeli autokorreláció) és a térbeli struktúra (térbeli heterogenitás) által generált ökonometriai problémák kezelésével foglalkozik. A tanulmány elsősorban a térökonometria szemlélet- és módszertanbeli sajátosságainak magyar nyelven való bemutatására vállalkozik s korántsem törekszik a tudományág teljes körű elemzésére. Ennek megfelelően a térökonometriai modellek közül a széleskörűen alkalmazott lineáris modellek állnak a középpontban. A tanulmány a tér ökonometriai modellezésének, a regressziós modellek becslési eljárásainak, valamint a térbeli autokorreláció tesztelésének problémáit mutatja be.

TÁRGYSZÓ: Ökonometria; Területi elemzések; Autokorreláció.

A térökonometria az ökonometria azon részterülete, mely a keresztmetszeti és paneladatokra épülő regressziós modellekben a térbeli interakció (térbeli autokorreláció) és a térbeli struktúra (térbeli heterogenitás) által generált ökonometriai problémák kezelésével foglalkozik (*Paelinck–Klaassen*; 1979, *Anselin*; 1988a). A térökonometria bizonyos hasonlóságot mutat a térstatisztikával, melyet széleskörűen alkalmaznak bizonyos fizikai tudományokban (mint például a statisztikai mechanika, az ökológia vagy az epidemiológia), ám ugyanakkor különbözik is tőle s ez a különbség a statisztikai és ökonometriai megközelítés alapvető eltéréséből következik. Ugyanis míg az ökonometriai elemzés jellege, illetve az alkalmazott specifikációk mindig közgazdaság-elméleti modellek által meghatározottak, a statisztika kiindulópontja elsősorban az adatokban van, s így közelítésmódját korántsem jellemzi egy ekképpen értelmezett szigorúan a priori elméleti meghatározottság.

Paelink és Klaassen belga ökonometrikusok a térökonometria alapjait eredetileg a keresztmetszeti adatokat használó regionális és multiregionális modellekben fellépő térbeli autokorreláció kezelésére dolgozták ki az 1970-es évek végén (*Paelinck–Klaassen*; 1979). A térökonometria modelljeinek továbbfejlesztése, illetve azok alkalmazása kezdetben elsősorban a közgazdaságtan specializált, kifejezetten térbeli problémákkal foglalkozó terü-

* A szerző ezúton is szeretné köszönetét kifejezni *Luc Anselinnek*, a térökonometria vezető nemzetközi kutatójának, akit professzoraként, PhD disszertációs témavezetőjeként, majd kollégaként és több tanulmányban szerzőtársként tisztelhet. A tanulmány korábbi változatairól készített alapos véleményekért *Hunyadi Lászlót*, *Nemes Nagy Józsefet*, *Pintér Józsefet* és *Rappai Gábort* illeti köszönet. Az esetleges hibákért természetesen a szerző viseli a felelősséget.

letei – úgy, mint a regionális közgazdaságtudomány, a városgazdaságtan, a gazdaságföldrajz vagy az ingatlan-gazdaságtan – határain belül zajlott. Az 1990-es évek új fejleménye, hogy a térökonometria módszereinek felhasználására egyre több példával találkozhatunk a közgazdaságtan olyan ágaiban is, ahol a folyamatok térbeli vonatkozásainak kutatása hagyományosan nem tartozott az érdeklődés homlokterébe, így a világgazdaságtanban, az államháztartástanban, a munka gazdaságtanában, az agrárgazdaságtanban, vagy a technológiai fejlődés gazdaságtanában (lásd például *Anselin–Varga–Acs*; 1997, *Aten*; 1996, *Bell–Bockstael*; 1999, *Case*; 1991, *Case–Rosen–Hines*; 1993, *Holtz–Eakin*; 1994, *Murdoch–Rahmatian–Thayer*; 1993, *Murdoch–Sandler–Sargent*; 1997, *Nelson–Hellerstein*; 1997, *Topa*; 1996, *Varga*; 1998). Mindezek mellett az utóbbi években az irányadó ökonometriai irodalomban is megjelent több olyan tanulmány, mely egyes modellspecifikációk, tesztstatisztikák és becslési eljárások térben kódolt adatokra való kidolgozását célozza (lásd például *Conley*; 1996, *Driscoll–Kray*; 1998, *Pinkse–Slade*; 1998).

A térökonometriai alkalmazások iránt az 1990-es évektől megélénkült közgazdaságtani érdeklődés két tényező hatására vezethető vissza (*Anselin*; 2001). Az egyik tényezőként az a gazdaságelméletben jelentkező új keletű irányzat említhető, mely (szakítva a közgazdaságtanban mindmáig uralkodó atomisztikus, elszigetelt módon döntést hozó gazdasági szereplő képével) a hangsúlyt az egyének között kiépülő (innovációs, termelési, értékesítési) hálózatok, a gazdasági szereplők térbeli koncentrációjából fakadó extern hatások vagy a tudás-spilloverek formájában működő interakciók tanulmányozására helyezi át. Ez az új szemlélet jelentkezik például az Aoki-féle „új makroökonómiában” (*Aoki*; 1994, 1996), a Romer-Lucas-féle „új (endogén) növekedéselméletben” (*Romer*; 1990, *Lucas*; 1988), a Krugman-féle „új gazdaságföldrajzban” (*Krugman*; 1991, *Fujita–Krugman–Venables*; 1999), vagy az „innovációs rendszerek” Nelson-Lundvall-féle irányzatában (*Nelson*; 1993, *Lundvall*; 1992).[1] A térökonometria iránti megnövekedett érdeklődés tehát egyrészt az empirikus térbeli modellek kezelésére képes ökonometriai módszerek iránti igényből fakad. Az alkalmazások szaporodását másrészt a térinformatika (Geographical Information System – GIS) gyors terjedésével párhuzamosan elsősorban az Egyesült Államokban megjelent nagyméretű, térben kódolt elektronikus adatbázisok (vagyis olyan adatbázisok, melyek a megfigyelési egységek – megyék, városok, kerületek stb. – térbeli elhelyezkedésére vonatkozóan is tartalmaznak információt) ökonometriai elemzésének igénye magyarázza. A térbeli adatoknak ugyanis meghatározó sajátossága a térbeli autokorreláció, melynek kezelésére az idősoros vagy a paneladatok elemzésénél általában használt ökonometriai technikák nem alkalmasak.

Tanulmányomban elsősorban a térökonometria szemlélet- és módszertanbeli sajátosságainak magyar nyelven való bemutatására vállalkozom s korántsem törekszem a tudományág teljes körű elemzésére. Így a térökonometriai modellek közül a széleskörűen alkalmazott lineáris modellek állnak a középpontban, míg a kutatás aktív területeiként számon tartott egyéb ökonometriai problémák (például a térbeli panelmodellek, vagy a kategorizált adatok térökonometriai elemzése[2]) részletes bemutatása kívül marad a ta-

[1] Az új növekedésleméletről, az innovációs rendszerek irányzatáról és az új gazdaságföldrajzról részletesebben magyarul lásd *Acs* és *Varga* (2001) tanulmányát.

[2] A térbeli panelmodellekkel kapcsolatos módszertani kérdésekről lásd például *Anselin* (1988a), *Baltagi–Li* (1999), a térbeli probitmodellel kapcsolatban pedig *Case* (1992), *McMillen* (1992), *Bolduc–Fortin–Gordon* (1997), *Heagerty–Lele* (1998), *Pinkse–Slade* (1998), *Beron–Vijverberg* (1999).

nulmány keretein.[3] A következő rész a tér ökonometriai modellezése problémájának felvázolását és az erre adott térökonometriai megoldást mutatja be. A harmadik rész a térbeli autokorreláció statisztikai teszteléséről, míg a negyedik a lineáris modellekről és azok becslési eljárásairól ad áttekintést. A térbeli autokorrelációval összefüggő specifikációs tesztek ismertetése az ötödik rész tárgya. A tanulmány néhány záró gondolattal fejeződik be.

A TÉR ÖKONOMETRIAI REPREZENTÁCIÓJA

A térbeli adatok adekvát ökonometriai elemzésének alapfeltétele a megfigyelési egységek közötti térbeli relációk megfelelő módon való leképezése. A tér (közgazdasági szempontból) kétdimenziós jellegéből következően ugyanis azok az ökonometriai megoldások, melyek az „egydimenziós" idősoros adatokra épülő elemzéseknél megfelelőképpen használhatók, a térbeli problémák megragadására nem feltétlenül alkalmasak. A térbeli keresztmetszeti adatbázisok összeállításánál gyakran használt megoldás például az, amikor a megfigyelési egységek betűrendben vagy statisztikai kódszámok alapján kerülnek sorbarendezésre. Mint a későbbiekben látni fogjuk, bizonyos esetekben az idősoros elemzések ezen „egydimenziós" technikájának térbeli adaptációja elfogadható, az esetek többségében viszont nem eredményez megfelelő megoldást.

Idősoros elemzéskor a megfigyelési egységeket időbeli előfordulásuk alapján rendezzük, így az egységek sorrendje megfelel azok időbeli sorrendjének. Az időbeli folyamatok egyirányúsága miatt a valóságnak ez a leképezése teljesen megfelelő. Ez a leképezés lehetővé teszi, hogy mind a megfigyelési egységek egymás közötti függőségéből (időbeli korreláció), mind azok valószínűség-eloszlás függvényeinek különbözőségeiből (időbeli heterogenitás) eredő ökonometriai problémákat kezelni tudjuk. Az időbeli megfigyelések tehát statisztikai–ökonometriai szempontból a legjobban az „idővonal" egymást követő pontjaiként képezhetőek le. Térbeli adatok esetén a térkép a valóság hasonlóan absztrakt reprezentációja. Azok a módszerek, melyek alkalmasak az egydimenzióban mozgó időbeli folyamatok elemzésére, nem feltétlenül alkalmazhatók egy kétdimenziós jelenség vizsgálatára. Ha a térbeli megfigyelési pontok egymástól függetlenek, akkor e pontok egymáshoz viszonyított helyzetének ismerete nem indokolja speciális ökonometriai módszerek alkalmazását. Akkor viszont, ha a megfigyelési egységeket egymás közötti korreláció jellemzi, vagy azok egymástól igen eltérő jellegzetességekkel bírnak (térbeli heterogenitás), a relatív térbeli pozíciók figyelmen kívül hagyása téves eredményekre vezethet.

Mind a térstatisztikusok mind a geográfusok evidenciaként kezelik azt, hogy a térbeli adatok egyik meghatározó jellemzője azok heterogenitásra és korrelációra való hajlama. A térbeli heterogenitás a térbeli egységek sajátosságainak, másoktól való különbözőségeinek következménye. Városok, megyék stb. ugyanis igen jelentősen különbözhetnek egymástól bizonyos gazdaságilag meghatározó jellemzők szempontjából. Így például jelentős különbségek tapasztalhatóak a megfigyelési egységek között az ott élő munkaerő minőségében, számában, az ott található természeti erőforrások mennyiségében, vagy akár a kutatás–fejlesztési kiadásokban.

A megfigyelési egységek térbeli függősége (vagy térbeli autokorrelációja – „spatial autocorrelation") inkább szabálynak, mint kivételnek tekinthető. A térstatisztika vezető

[3] A jelenleginél jóval részletesebb elemzésekhez a következő áttekintő munkákat ajánlom az olvasó figyelmébe: *Anselin* (1988a, 2001), *Anselin–Florax* (1995), *Anselin–Rey* (1997), *LeSage* (1999).

nemzetközi tekintélye, *Cressie* véleménye szerint térbeli adatok esetén „a dependencia minden irányban jelen van és minél szétszórtabban helyezkednek el a megfigyelési egységek, annál gyengébb a függőség" (*Cressie*; 1993, 3. old.). A térbeli dependencia lényegében függvényszerű relációt jelent ugyanazon változó különböző helyeken mért értékei között. Míg az adatok időbeli függősége esetén a két megfigyelési egység közötti reláció egyirányú, addig a térbeli függőség kétirányú reláció kettőnél több térbeli egység között, vagyis egy változó bármely lokációban mért értéke nemcsak, hogy hat a többi lokációkban mért értékekre, de ugyanakkor maga is determinált az egyéb lokációkban mért értékek által.

Hasonló értékek csoportosulása a térben (az ún. klaszterek kialakulása) a megfigyelési egységekben mért értékek között a pozitív térbeli autokorreláció létrejöttét eredményezheti. Példaként hozható a nagyvárosok tehetősebbek vagy éppen szegények lakta negyedeinek csoportulása a térben vagy a világ nagy innovációs klasztereinek jelensége, ahol az egyes klaszterek számos, a technológiai innovációban sikeres nagyváros térbeli „sűrűsödésének" eredményei. Negatív térbeli autokorreláció esetén az egyes megfigyelési egységek szomszédságában azoktól különböző értékeket mérünk, ami az értékek egyfajta sakktáblaszerű térbeli eloszlását eredményezi. A kétfajta autokorreláció közül a pozitív autokorreláció jelenségét jóval könnyebb megérteni, míg a negatív autokorrelációra nem minden esetben tudunk intuitív magyarázatot találni.

Az adatokban észlelt pozitív térbeli autokorreláció két eltérő okkal magyarázható: az egyik technikai jellegű, míg a másik szubsztantív, vagyis a társadalmi interakciók természetéből következik. A technikai ok a számbavétel esetleges hibáiból adódik. Térbeli adatok aggregációja tipikusan politikai határok (országok, megyék, városok) mentén történik, ugyanakkor a vizsgálat tárgyát képező folyamatok esetleg teljesen más mértéket követnek a térben. Mindezek következményeként az egymással szomszédos térbeli egységekben a vizsgált változó értékei igen közel esnek egymáshoz, ami a statisztikai–ökonometriai elemzésben pozitív térbeli autokorrelációként jelenik meg (példaként lehet megemlíteni azokat az eseteket, amikor a természeti adottságok kihasználására alapuló gazdasági tevékenységek – bányászat, mezőgazdaság stb. – több földrajzi egység területét ölelik fel, vagy amikor az országhatárokon átnyúló együttműködések következményeként a szomszédos, de az országhatárok által elválasztott térbeli egységek között bizonyos hasonlóságok alakulnak ki, ahogyan az például az osztrák-magyar határ mentén megfigyelhető).

A szubsztantív magyarázat szerint a társadalmi interakciók sajátos természetéből következik a térbeli autokorreláció, hiszen a megfigyelési egységek közötti kapcsolatok következtében a vizsgált változó igen hasonló értékei jöhetnek létre egymáshoz közel fekvő lokációkban még akkor is, ha az adatok aggregációja megfelelő. Mindez abból a közismert tényből következik, miszerint a társadalmi interakciók kialakulását a térbeli közelség jelentősen elősegíti. Az ekképpen létrejött autokorreláció erőssége tehát a térben nem konstans, hanem az egységek közötti távolsággal fordított arányban változik. Érvényes tehát az, amit *Tobler* (1979) a geográfia első törvényének nevez, vagyis, hogy „minden mindennel összefügg, de közeli jelenségek sokkal inkább összefüggenek egymással, mint távoli dolgok" (*Tobler*; 1979).

Amint az az előzőkben megállapítást nyert, abban az esetben, ha a térbeli adatokat homogenitás vagy térbeli függetlenség jellemzi, az időbeli adatoknál használt ökonometriai megoldások használata elfogadható. Amennyiben azonban a térbeli adatok heterogének és (vagy) közöttük autokorreláció tapasztalható, akkor az alkalmazandó regressziós eljárás szempontjából a megfigyelések térbeli pozíciója kiemelt jelentőségre tehet szert, s a térbeli helyzet megfelelő ökonometriai–statisztikai reprezentációja a további elemzés alapfeltételévé válhat. Az adatok térbeli heterogenitása eredményezheti a regressziós egyenletben jelentkező heteroszkedaszticitást. A térbeli heterogenitás kezelése így viszonylag kisebb kihívással jár, hiszen az esetek többségében a heteroszkedaszticitással kapcsolatos problémáknál széles körben alkalmazott ökonometriai módszerek, mint például a véletlen koefficiens modellek (random coefficient models), a hiba komponens modellek (error component models), vagy az adaptív szűrés (adaptive filtering) módszerei jól használhatók.[4] Mindezek mellett a térbeli adatelemzés irodalmában kifejezetten térspecifikus módszereket is találunk. E modellekben a tér reprezentációjára két alapvető megoldás található: a megfigyelési egységeket vagy azok *koordinátái* révén [ahogyan ezzel a módszerrel például a trend felszín analízisben (trend surface analysis) vagy a térbeli paraméter kiterjesztés modelljében (spatial expansion model) találkozhatunk] vagy pedig *dummy változók* felhasználásával [mint például a térbeli variancia-analízisben (spatial analysis of variance) vagy a térben váltakozó regresszió módszerében (spatially switching regression)] építik be az elemzésbe.[5] Míg a térbeni heterogenitás kezelése lényegében nem igényel speciális ökonometriai közelítésmódokat, addig ez a térbeli autokorreláció tárgyalásáról nem mondható el. Mindezek miatt a továbbiakban a tanulmány a térbeli függőség kezelésének módszertanára összpontosít.

A térbeli autokorreláció ökonometriai tárgyalásához olyan térreprezentációra van szükség, amely képes az egységek egymáshoz viszonyított helyzetét megragadni. A megfigyelési egységek szomszédsági fokok szerinti kategorizálása széleskörűen használt kiindulópont a relatív térbeli helyzetek leképezésében. Az elsőfokú szomszédság fogalma az egységek közötti legnagyobb közelség (legkisebb távolság) megragadására szolgál. Egymással közös határral bíró, vagy egymástól egy bizonyos kritikus távolságon belül elhelyezkedő megfigyelési egységeket nevezünk elsőfokú szomszédoknak. A másod- harmad- stb. fokú szomszédság pedig a közös határon, vagy kritikus távolságon alapuló felfogásnak a távolabb fekvő egységekre való kiterjesztéseként adódik (a másodfokú szomszéd az elsőfokú szomszéd szomszédja és így tovább).

A megfigyelési egységek relatív térbeli pozícióinak megragadására szolgálnak az úgynevezett térbeli súlymátrixok (spatial weights matrices). A térbeli súlymátrix (**W**) dimenziója a megfigyelési egységek számával (N) egyenlő. A térstruktúra legegyszerűbb leképezése az úgynevezett bináris szomszédságmátrixokkal (binary contiguity matrices) lehetséges, amikor a mátrix bármely elemének (w_{ij}) két lehetséges értéke 1 és 0. Amikor i és j egymással (a választott szomszédsági fok szerint) szomszédos, akkor az érték 1 lesz, egyébként pedig 0. Megállapodás szerint a mátrix diagonális elemei nullával egyenlők. A bináris szomszédságmátrixok szimmetrikusak.

[4] A heteroszkedasztikus reziduummal összefüggő ökonometriai problémák kezelésére általánosan alkalmazott módszerek színvonalas áttekintését magyar nyelven lásd *Kőrösi*, *Mátyás* és *Székely* (1990) könyvében, valamint *Pintér* (1991) tanulmányában.

[5] Bővebben e módszerekről lásd például *Anselin* (1998a), *Casetti* (1997), *Griffith* (1988), *Haining* (1990) vagy *Kristensen* (1996).

A következő mátrix egy 4x4-es bináris szomszédságmátrix példája:

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

A mátrixok másik osztályába tartoznak az úgynevezett inverz távolságalapú súlymátrixok (inverse distance weights matrices). Ekkor a mátrix egy-egy eleme az i és j lokációk közötti távolság reciprokának valamely fokú hatványával egyenlő.[6] A következő, 4x4-es mátrix erre szolgál példaként, ahol d_{ij} az i-edik és a j-edik megfigyelési egység közötti távolságot reprezentálja.

$$\mathbf{W} = \begin{pmatrix} 0 & \frac{1}{(d_{1,2})^2} & \frac{1}{(d_{1,3})^2} & \frac{1}{(d_{1,4})^2} \\ \frac{1}{(d_{2,1})^2} & 0 & \frac{1}{(d_{2,3})^2} & \frac{1}{(d_{2,4})^2} \\ \frac{1}{(d_{3,1})^2} & \frac{1}{(d_{3,2})^2} & 0 & \frac{1}{(d_{3,4})^2} \\ \frac{1}{(d_{4,1})^2} & \frac{1}{(d_{4,2})^2} & \frac{1}{(d_{4,3})^2} & 0 \end{pmatrix}.$$

A súlymátrixok minden egyes elemének a saját sorában elhelyezkedő elemek összegével való osztása révén kapjuk az úgynevezett sorstandardizált súlymátrixokat (row-standarized weights matrices). Értelemszerűen a sorstandardizált súlymátrixok minden egyes eleme nulla és egy közötti értéket vehet fel s az elemek összege minden egyes sorban 1-gyel egyenlő. A sorstandardizált mátrixok használata két okkal magyarázható: egyrészt az általuk kapott eredmények könnyebben értelmezhetők, másrészt a maximum likelihood becsléssel összefüggő számítások csak ezek alkalmazásával végezhetők el, ahogyan az a későbbiekben bemutatásra kerül. A térbeli autokorreláció statisztikai–ökonometriai kezelésének mélyebb megvilágítása céljából a következő fejezet a térbeli függőség tesztelésének módszertanába nyújt betekintést.

A TÉRBELI AUTOKORRELÁCIÓ STATISZTIKAI TESZTELÉSE

A térbeli autokorreláció[7] jelenlétére az egymáshoz igen hasonló értékek térbeli csoportosulásából (pozitív térbeli autokorreláció), vagy az egymással szomszédos megfigyelési egységeknél igen eltérő értékek jelentkezéséből (negatív térbeli autokorreláció) következtethetünk. A pusztán vizuális benyomás azonban nem helyettesítheti annak statisz-

[6] A térbeli súlymátrixokról jóval részletesebben lásd a következő munkákat: *Anselin* (1988a), *Cliff* és *Ord* (1981) és *Upton* és *Fingleton* (1985).

[7] A térbeli autokorreláció statisztikai vonatkozásairól magyar nyelven lásd még *Nemes Nagy* (1998).

tikai tesztelését, hogy vajon az adatok adott térbeli konstellációja pusztán a véletlen eredménye, vagy pedig térbeli autokorreláció jelenlétének következménye. A térbeli autokorreláció tesztelésére két típusú próbacsoport terjedt el a gyakorlatban: az egyik csoportba az úgynevezett globális próbafüggvények, míg a másikba a lokális statisztikák tartoznak. A globális tesztek az adott változó térbeli autokorrelációra való hajlamát tesztelik, a teljes adatsor figyelembevételével, míg a lokális próbák az egyes megfigyelési egységek szintjén vizsgálják a helyi (pozitív, vagy negatív) autokorreláció jelenlétét.

A továbbiakban a globális próbafüggvények közül a Moran-féle I-próba (Moran's I statistic) (*Moran*; 1948, *Cliff–Ord*; 1981), míg a lokális tesztek közül az Anselin-féle Lokális Moran-próba (Local Moran) (*Anselin*; 1995) kerül bemutatásra.[8]

A Moran-féle I-próba próbafüggvénye:

$$I = [N/S_0]\ [\Sigma_{i,j}\, w_{ij}\, (x_i - \mu)(x_j - \mu)\ /\ \Sigma_i(x_i - \mu)^2], \qquad /1/$$

ahol N a megfigyelési egységek száma, x_i és x_j az x változó két pontban mért értékei, μ ezek várható értéke, w_{ij} a térbeli súlymátrix egy eleme, S_0 pedig normalizáló faktor:

$$S_0 = \Sigma_{i,j}\, w_{ij}, \qquad /2/$$

vagyis a súlyok összege. Ha $\mathbf{W}$ sorstandardizált súlymátrix (ahogyan az általában javasolt), akkor $S_0 = N$, és a próbafüggvény a következő egyszerűbb formát ölti:

$$I^* = \Sigma_{i,j}\, w_{ij}\, (x_i - \mu)(x_j - \mu)\ /\ \Sigma_i(x_i - \mu)^2, \qquad /3/$$

ami bár hasonló, de mégsem teljesen azonos a korrelációs együtthatóval. I^* várható értéke ugyanis $[-1/(N-1)]$-el egyenlő, ami a mintaelemszám növekedésével a nullához közelít (*Anselin*; 1992). I^*-nak a várható értéknél nagyobb értékei pozitív, míg annál kisebb értékei negatív térbeli autokorrelációra utalnak.

A globális térbeli autokorreláció tesztelésére használatos bármely próba (melyek közül most a Moran-féle I-próba került bemutatásra) ugyanarra az alapelvre épül, mely az úgynevezett Gamma-indexben (*Hubert–Constanzo–Gale*; 1985) nyer általános formát:

$$\Gamma = \Sigma_{i,j}\, w_{ij}\, c_{ij}\ . \qquad /4/$$

A Gamma-index jelöléseit alkalmazva, a Moran-féle I-próba függvényében a c_{ij} a következőképpen határozható meg:

$$c_{ij} = (x_i - \mu)(x_j - \mu)/\Sigma_i\, (x_i - \mu)^2. \qquad /5/$$

A Gamma-indexre alapuló globális próbák a következő elven alapulnak. Mivel a térbeli autokorreláció vagy hasonló értékek (pozitív térbeli autokorreláció), vagy pedig nagyon különböző értékek (negatív térbeli autokorreláció) térbeni csoportosulását jelenti, ezért a próbák a megfigyelési egységekben mért értékek között fennálló tendenciát (vagyis, hogy azok

[8] A térbeli autokorreláció egyéb globális és lokális próbafüggvényeiről lásd részletesebben például *Anselin* (1995), *Getis* és *Ord* (1992) vagy *Cliff–Ord* (1981).

tendenciaszerűen hasonlók, vagy pedig különbözők-e egymástól) a megfigyelési egységek térbeli helyzeteiben észlelt tendenciákkal (vagyis, hogy a hasonló, vagy különböző értékek általában térben „közel", vagy pedig „távol" helyezkednek-e el egymástól) vetik össze. Ha az /5/ egyenletben $c_{ij} > 0$ (vagyis mindkét lokációban a mért érték nagyobb, vagy kisebb mint a várható érték), és w_{ij} értéke is nullától különbözik (vagyis i és j szomszédai egymásnak), akkor hasonló értékek találhatók egymással szomszédos lokációkban, ami pozitív térbeli autokorrelációt jelent. Abban az esetben, ha különböző értékek egybeesése tapasztalható a térben, akkor ugyanazon i és j lokációkra a c_{ij} negatív értéket, míg a w_{ij} nullától különböző értéket mutat, ami a negatív térbeli autokorreláció jelenlétére utal.

A Moran-féle I-próba szignifikanciájának, vagyis annak eldöntésére, hogy vajon egy változó értékeinek adott térbeli eloszlása térbeli autokorreláció jelenlétének eredménye-e vagy sem, két alapvető módszer használatos. Mindkét módszer az értékek adott térbeli eloszlását az általában véletlenszerűen kialakuló térbeli eloszlásokhoz viszonyítja, s ha az adott térbeli eloszlás extrémnek minősül a véletlenszerűen előfordulóhoz képest, akkor ezt szignifikáns térbeli autokorrelációnak (azaz az értékeknek a véletlenszerűen előfordulótól eltérő mértékű térbeli csoportosulásának) minősíti. Természetesen az, hogy mi minősül extrémnek, részben a változónak a null-hipotézis (azaz a térbeli autokorreláció nélküli eset feltételezése) melletti valószínűségeloszlásától, részben a választott szignifikanciaszinttől függ.

A leggyakrabban használatos eljárás az értékek normális eloszlásának feltételezésére épül. Ekkor a Moran-féle I-próba szignifikanciájának megállapítása a standardizált normális eloszlású valószínűségi változó értékén alapul, melyet a szokásos módon képezünk:

$$Z_I = [I - E(I)] \,/\, SD(I), \qquad /6/$$

ahol $E(I)$ az I várható értéke, $SD(I)$ pedig az I szórása. Az I szignifikanciájának megállapítása pedig az általában szokásos módon történik.

Az értékek normális eloszlásának feltevése azonban igen sok esetben túlságosan erős feltételezés. Alternatívaként az úgynevezett permutációs módszer használata ajánlott. A módszer null-hipotézise szerint annak valószínűsége, hogy a mért értékek milyen permutációban jelennek meg az egyes lokációk között, bármely permutációra vonatkozólag ugyanaz. E feltételezés alapján a változó mért értékeinek a lokációk közötti véletlenszerű permutációihoz kiszámítjuk az I megfelelő értékeit, melynek eredményeként az I értékeinek úgynevezett mesterséges referenciaeloszlása hozható létre. Az extremitás fokának megállapításakor a tényleges I értéket viszonyítjuk a véletlen permutációk eredményeként számított I értékek eloszlásához. Egy egyszerű hüvelykujjszabály szerint a döntés alapja az úgynevezett pszeudo-szignifikancia szint. Ezt a következőképpen számítjuk: $(T+1)/(M+1)$, ahol T azoknak az eseteknek a száma, amikor a referenciaeloszlás által adott értékek nagyobbak vagy egyenlők, mint a mért I érték, míg M a permutációk számát jelöli. A pszeudo-szignifikancia szint alacsony értéke annak jele, hogy az értékek adott térbeli eloszlása extrém esetnek minősül ahhoz képest, ami a null-hipotézis alapján elvárható lenne, ami (az I előjelétől függően) pozitív, vagy negatív térbeli autokorreláció jelenlétét igazolja.

A térbeli autokorreláció globális próbái kiválóan alkalmazhatók az adatok térbeli csoportosulásra való hajlamának megállapítására. Sok esetben viszont nemcsak az általános, de a helyi tendenciák ismerete is szükséges. Előfordulhat, hogy az adatokat általában nem jellemzi térbeli autokorreláció, mégis található néhány határozott térbeli csoportosulás (klaszter).

Az egyes helyi csoportosulások szignifikanciájának megállapítására a térbeli autokorreláció lokális próbafüggvényeit alkalmazhatjuk. Ezek közül a következőkben az Anselin-féle Lokális Moran-próba kerül bemutatásra. A próba függvénye:

$$I_i = z_i \Sigma_j w_{ij} z_j, \qquad /7/$$

ahol I_i az i-edik megfigyelési egységre számított Lokális Moran próbafüggvény értéke, $z_i = x_i - \mu$, vagyis z_i a változó értékének és a várható értéknek a különbsége az i-edik lokációban, z_j pedig ugyanez a j-edik lokációra. Mivel **W** sorstandardizált, ezért a $\Sigma_j w_{ij} z_j$ tag az i-edik megfigyelési egységhez tartozó szomszédsági értékek súlyozott átlagát jelenti. Ha z_i és $\Sigma_j w_{ij} z_j$ értékei egyként pozitívak (negatívak), akkor I_i értéke pozitív autokorrelációt mutat, ami magas (alacsony) értékek térbeli koncentrálódását jelenti az i-edik megfigyelési egység környezetében. Az I_i szignifikanciájának meghatározásához a globális Moran I-próbafüggvényében is alkalmazott permutációs módszer használatos.

TÉRBELI AUTOKORRELÁCIÓ A LINEÁRIS TÉRÖKONOMETRIAI MODELLEKBEN

Amint a következőkből is látható, abban az esetben, ha az adatokat térbeli autokorreláció jellemzi, a hagyományosan használt becslési módszerek nem alkalmazhatók. A térbeli autokorreláció ökonometriai modellezésének két legelterjedtebb módszere a térbeli késleltetés modellje (spatial lag model), illetve a térbeli hiba autokorreláció modellje (spatial error model).

A térbeli késleltetés modellje

A súlymátrixoknak a térbeli regressziós modellekben való alkalmazásával egy változó bizonyos pontban mért értékét ugyanezen változónak a tér más pontjain mért értékeivel hozzuk összefüggésbe. Az idősoros elemzés mintájára térbeli adatokra is értelmezhető a késleltetés koncepciója, azzal a különbséggel, hogy itt a késleltetés nem az időben, hanem a térben való „elcsúszásként” értelmezett.

Az idősoros elemzés során a késleltetett érték a késleltetési operátor révén kerül kifejezésre, vagyis például

$$L^k y_t = y_{t-k}, \qquad /8/$$

ahol $L^k y_t$ az y_t k időszakkal korábbi értékét jelöli. Térbeli adatok esetén sajnos nem ilyen egyértelmű a helyzet, ami annak következménye, hogy a késleltetés bármely szomszédsági fokon belül többirányú is lehet. Hacsak a mintaelemszám nem kifejezetten magas, e megoldás a térben késleltetett változók nagy számát eredményezheti, ami miatt a szadadságfok olyan kicsire zsugorodhat, ami akár a regressziós egyenlet becslését is lehetetlenné teheti. A probléma megoldására a térbeli késleltetés során az adott szomszédsági fokba tartozó megfigyelési egységekben mért értékek súlyozott átlagát vesszük figyelembe, vagyis

$$L^s x_i = \Sigma_j w_{ij} x_j, \qquad /9/$$

ahol L^s az s-edik fokú szomszédságnak megfelelő térbeli késleltetési operátor, x_j az x-nek a j-edik megfigyelési egységben mért értéke, ahol j az s-edik szomszédsági foknak megfelelő eleme a térnek, w_{ij} pedig a sorstandardizált térbeli súlymátrix megfelelő eleme.[9]

A térbeli késleltetés modelljének általános formája:

$$\mathbf{y}_{(N\times 1)} = \rho\mathbf{W}_{(N\times N)}\mathbf{y}_{(N\times 1)} + \mathbf{X}_{(N\times K)}\boldsymbol{\beta}_{(K\times 1)} + \boldsymbol{\varepsilon}_{(N\times 1)}, \qquad /10/$$

ahol **y** az eredményváltozó értékeinek vektora, **W** sorstandardizált súlymátrix, **Wy** az eredményváltozó térben késleltetett értékeinek vektora, **X** az exogén változók mátrixa, ρ a térbeli autoregressziós paraméter, $\boldsymbol{\beta}$ az exogén változók paramétervektora, $\boldsymbol{\varepsilon}$ pedig az egymástól független és azonos valószínűségeloszlású hibatagok vektora. A hibatagok várható értéke **0**, szórásnégyzete pedig σ^2.

Míg az időbeli adatokra épülő késleltetéses modell becslésére (abban az esetben, ha hibaautokorreláció nem tapasztalható) alkalmazható a Legkisebb Négyzetek Módszere (Ordinary Least Squares – OLS) (lásd például *Greene*; 1993), mindez nem igaz a térbeli késleltetés modelljének becslésére. Ennek oka a térben késleltetett függő változó endogenitása (hiszen a térbeli autokorreláció kétirányúsága következtében **Wy** nemcsak meghatározza az **y**-t, de maga is meghatározott az **y** által). Következésképpen a térben késleltetett függő változó korrelál a hibataggal, s így az OLS-becslőfüggvény torzított és inkonzisztens lesz (*Anselin*; 1988a). Az OLS helyett leggyakrabban használt két becslési eljárás a maximum likelihood módszer, illetve az instrumentális változók módszere.

A maximum likelihood becslés[10] alkalmazásának alapfeltétele a reziduális változó normális eloszlása. A becslési eljárás során azokat a paramétereket keressük, melyek mellett a függőváltozó megfigyelt értékei együttes előfordulásának valószerűsége maximális. A függőváltozó likelihood fügvénye a normális eloszlást követő reziduális változó együttes sűrűségfüggvényének transzformációjaként adódik. A likelihood függvényt a reziduumok együttes sűrűségfüggvényének és a Jakobi-mátrix determinánsának a szorzataként kapjuk. A Jakobi mátrix a reziduális változónak a függőváltozó szerinti elsőfokú deriváltjait tartalmazó mátrix. A térbeli késleltetés modelljében a Jakobi mátrix determinánsának formája: $|\mathbf{I} - \rho\mathbf{W}|$.

Így a logaritmikus likelihood fügvény a következő formát ölti:

$$L = \ln|\mathbf{I} - \rho\mathbf{W}| - N/2 \ln(2\pi) - N/2 \ln(\sigma^2) -$$
$$-(\mathbf{y} - \rho\mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \rho\mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})/2\sigma^2. \qquad /11/$$

A logaritmikus likelihood függvény $\boldsymbol{\beta}$, ρ, és σ^2 szerinti maximalizálása eredményezi a paraméterek azon értékeit, melyek mellett a függőváltozó adott mintájának előfordulási valószerűsége a legnagyobb. Mivel e probléma a térbeli késleltetés modelljére analitikusan nem megoldható, ezért a gyakorlatban numerikusan történik a paraméterek meghatározása.

Az egyébként igen bonyolult számításokat két tény jelentősen könnyíti. Egyrészt, mivel a $\boldsymbol{\beta}$ és σ^2 kifejezhető ρ függvényeként ezek helyettesíthetők az előbbi logaritmikus likelihood

[9] A térbeli késleltetés koncepcójának igen részletes magyarázatára lásd *Anselin* (1988a).

[10] A maximum likelihoodnak a térbeli késleltetés modelljében történő alkalmazásáról igen részletes elemzés olvasható a következő munkákban: *Anselin* (1988a), *Cliff* és *Ord* (1981) vagy *Ord* (1975).

függvénybe, ami az úgynevezett koncentrált likelihood függvényt eredményezi. Ugyanakkor a Jakobi-mátrix determinánsa az előzőnél lényegesen egyszerűbb formában is kifejezhető (*Ord*; 1975). Eszerint $\ln|\mathbf{I} - \rho\mathbf{W}| = \Sigma_i \ln(1 - \rho\omega_i)$, ahol ω_i a **W** saját értékeit jelöli. Ennek az eredménynek igen nagy a gyakorlati jelentősége, hiszen a Jakobi-mátrix determinánsának ezen formája lényegesen megkönnyíti a paraméterek kiszámítását.

A koncentrált likelihood függvény az előzők alapján a következő formát ölti:

$$L_c = \Sigma_i \ln(1 - \rho\omega_i) - (N/2)\ln[(\mathbf{e_0} - \rho\mathbf{e_L})' (\mathbf{e_0} - \rho\mathbf{e_L})/N], \qquad /12/$$

ahol $\mathbf{e_0}$ reziduum vektor abban a regressziós egyenletben, ahol **y** a bal oldali és **x** a jobb oldali változó, $\mathbf{e_L}$ pedig szintén reziduum vektor akkor, amikor **Wy** szerepel bal oldali és **x** jobb oldali változóként. A ρ elfogadható értékei az $1/\omega_{max}$-tól az $1/\omega_{min}$-ig terjedő sávban találhatók, ahol ω_{min} a **W** legkisebb, ω_{max} pedig a **W** legnagyobb saját értéke. A ρ azon értékének meghatározása, melynél a korábbi koncentrált likelihood függvény a maximális értéket veszi fel, numerikusan történik. Ezek után behelyettesítéssel adódik a $\boldsymbol{\beta}$ és σ^2 megfelelő értéke. A becslés további részleteiről lásd *Anselin* (1988a).

Ha a reziduum nem normális eloszlást követ, akkor a leggyakrabban használt alternatív becslési eljárás a kétfokozatú legkisebb négyzetek (2SLS) módszere. A 2SLS alkalmazását a térben késleltetett függőváltozó endogenitása indokolja. Az endogenitásból következően e változó korrelál a reziduummal. A 2SLS során instrumentális változók felhasználásával az endogén változót annak becsült értékével helyettesítjük, mely már mentes a reziduummal való autokorerrelációtól. Instrumentális változókként az egyenlet első fokon késleltetett magyarázó változóinak használata javasolt az irodalomban (**WX**).[11]

A térbeli hiba autokorreláció modellje

A térbeli autokorreláció ökonometriai modellezésére különböző lehetőségek adódnak. A térbeli késleltetés előzőkben vázolt modelljében a térbeli függőség közvetlen módon kerül modellezésre. Ez a közelítésmód más néven szubsztantív térbeli függőség-modellezésként is szerepel az irodalomban,[12] utalva arra, hogy a térben késleltetett független változó becsült paraméterének értéke és annak szignifikanciája explicit módon tudósít arról, hogy a független változó adott értékei kialakulásában a változónak a tér más pontjain mért értékei milyen szerepet játszanak.

Alternatív lehetőségként kínálkozik a térbeli függőség olyan megközelítése, amikor a regressziós egyenlet a magyarázó változóknak és a független változónak a térbeli autokorreláció hatásaitól „megtisztított" viszonyát tárja elénk. Ekkor a térbeli autokorreláció zavaró, vagyis kiiktatandó tényezőként szerepel a modellben. A térbeli hiba autokorreláció modellje tehát a hibatagok között jelentkező térbeli autokorreláció korrekcióját szolgálja.

A térbeli hiba autokorreláció modelljének általános formája:

$$\mathbf{y}_{(N\times 1)} = \mathbf{X}\boldsymbol{\beta}_{(K\times 1)} + \boldsymbol{\varepsilon}, \qquad /13/$$

[11] A térbeli késleltetés modelljének instrumentális változók módszerével való becsléséről lásd még *Land* és *Deane* (1992), *Kelejian* és *Robinson* (1993), *Kelejian* és *Prucha* (1998). Az instrumentális változók módszerének általánosabb vonatkozásairól pedig magyar nyelven bővebben *Kőrösi*, *Mátyás* és *Székely* (1990) kínál áttekintést.

[12] Lásd például *Anselin* (2001).

és

$$\boldsymbol{\varepsilon}_{(N \times 1)} = \lambda \mathbf{W} \boldsymbol{\varepsilon} + \zeta, \qquad /14/$$

ahol **y** az eredményváltozó értékeinek vektora, **X** az exogén változók mátrixa, $\boldsymbol{\beta}$ az exogén változók paramétervektora, $\boldsymbol{\varepsilon}$ az autoregresszív hibatagok vektora, **W** sorstandardizált súlymátrix, λ az autoregresszív hibatagok térben késleltetett értékeinek paramétere, ξ pedig az egymástól független és azonos valószínűségeloszlású hibatagok vektora. A független hibatagok várható értéke **0**, szórásnégyzete pedig σ^2.

Az idősoros esethez hasonlóan, az OLS becslőfüggvény torzítatlan, de nem hatásos a térbeli hiba autokorreláció esetében sem.[13] Szemben viszont az idősoros esettel, az ott igen gyakran használt általánosított legkisebb négyzetek módszere (GLS) a térbeli adatokra nem alkalmazható (*Anselin*; 1988a). Ennek oka a térbeli autokorreláció szimultán jellege, ami miatt a térbeli hiba autokorreláció esetén a maximum likelihood a leggyakrabban használatos becslési eljárás.

A térbeli hiba autokorreláció esetén a logaritmikus likelihood függvény a következő alakot ölti:

$$L = \ln\left|\mathbf{I} - \lambda\mathbf{W}\right| - N/2\ \ln(2\pi) - N/2\ \ln(\sigma^2) - (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \lambda\mathbf{W})'(\mathbf{I} - \lambda\mathbf{W})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})/2\sigma^2 \qquad /15/$$

a korábban már ismertetett jelöléseket alkalmazva. A likelihood függvényt λ, $\boldsymbol{\beta}$ és σ^2 szerint maximalizálva kapjuk a paraméterek maximum likelihood becslését.[14] A térbeli hiba autokorreláció modellje alternatív becslési eljárásainak kifejlesztése – mint például az általánosított momentum módszer (GMM) alkalmazási lehetőségeinek kidolgozása – napjaink térökonometriai kutatásainak egyik igen aktív területe.[15]

A TÉRBELI AUTOKORRELÁCIÓ TESZTELÉSE ÖKONOMETRIAI MODELLEKBEN

A legmegfelelőbb térökonometriai modellspecifikáció meghatározását számos teszt segíti. A diagnosztikák első csoportjába azok a próbák tartoznak, amelyek a térbeli hiba autokorreláció vagy a térbeli késleltetés modelljei közötti döntést teszik lehetővé. Ezek a próbák a térbeli autokorrelációnak az OLS-módszerrel becsült regressziós egyenletben való jelenlétét vizsgálják. A második csoportban pedig olyan diagnosztikák találhatóak, melyek az alkalmazott térökonometriai modell megfelelőségét tesztelik abból a szempontból, hogy vajon a szóban forgó modell a térbeli autokorreláció problémáját megnyugtatóan kezeli-e vagy valamilyen formában a függőség még mindig jelen van a regressziós egyenletben. A

[13] A különböző becslési eljárásokról, illetve az esztimátorok tulajdonságairól részletesebben lásd *Anselin* (1988a), *Cliff* és *Ord* (1981) és *Ord* (1975).

[14] A térbeli késleltetés modelljéhez hasonlóan $\boldsymbol{\beta}$ és σ^2 a λ függvényeként adódik, ami a következő koncentrált likelihood függvényt eredményezi: $L_c = \Sigma_i \ln(1 - \lambda\omega_i) - (N/2)\ln[(1/N)\mathbf{e}'(\mathbf{I} - \lambda\mathbf{W})'(\mathbf{I} - \lambda\mathbf{W})\mathbf{e}]$, ahol **e** a GLS becslés reziduum vektora, míg a többi jelölés a fentiekkel megegyezik. A λ elfogadható értékei az $1/\omega_{max}$-tól az $1/\omega_{min}$-ig terjedő sávban találhatók, ahol ω_{min} a **W** legkisebb, ω_{max} pedig a **W** legnagyobb saját értéke A λ azon értékének meghatározása, melynél a fenti koncentrált likelihood függvény a maximális értéket veszi fel, numerikusan történik. Ezek után behelyettesítéssel adódik a $\boldsymbol{\beta}$ és σ^2 megfelelő értéke. A becslés további részleteiről lásd *Anselin* (1988a).

[15] Bővebben lásd *Conley* (1996), *Kelejian* és *Prucha* (1999).

következőkben a legfontosabb próbák bemutatása következik.[16] A próbafüggvények összefoglaló bemutatását lásd a táblában.

Az ökonometriai modellekben leggyakrabban használt térbeli autokorreláció-tesztek

A próba neve	Próbafüggvény	Valószínűség-eloszlás	Forrás
MORAN	$\mathbf{e'We/e'e}$	$N(0,1)$	*Cliff–Ord* (1981)
LM-ERR	$(\mathbf{e'We}/s^2)^2/T$	χ^2 (1)	*Burridge* (1980)
LM-ERRLAG	$(\mathbf{e'We}/s^2)^2$ / $[\mathrm{tr}(\mathbf{W'W} + \mathbf{W}^2) - \mathrm{tr}(\mathbf{W'W} + \mathbf{W}^2)\mathbf{A}^{-1}\mathrm{var}(\rho)]$	χ^2 (1)	*Anselin* (1988a)
LM-LAG	$(\mathbf{e'Wy}/s^2)^2/(\mathrm{RJ}_{\rho-\beta})$	χ^2 (1)	*Anselin* (1988b)
LM-LAGERR	$(\mathbf{e'B'BWy})^2/(H_\rho - \mathbf{H}_{\theta\rho}\mathbf{Var(\theta)H'}_{\theta\rho})$	χ^2 (1)	*Anselin–Bera–Florax–Yoon* (1996)

A térbeli hiba autokorrelációnak az OLS modellben való tesztelésére széleskörűen alkalmazott módszer a Moran-féle *I*-próba következő formája:

$$I = \mathbf{e'We/e'e}, \qquad /16/$$

ahol **e** az OLS-reziduumok vektora, **W** pedig sorstandardizált ($N \times N$) méretű súlymátrix.

A statisztikai hipotézisvizsgálat a normális valószínűségeloszlás bázisán történik.[17] Ahogyan azt *Anselin* és *Rey* (1991) kimutatta, a Moran-féle *I*-próba ismertetett formája nem megbízható olyan esetekben, ha egyéb specifikációs problémák (például heteroszkedaszticitás) is jellemzik az empirikus modellt. Így a Moran-féle *I*-próba akkor is jelezhet térbeli autokorrelációt, amikor a gond egyéb okokból származik. Mindezek alapján a Moran *I*-próbát egyfajta általános specifikációs probléma-indikátorként ajánlatos lefuttatni, de a hiba-autokorreláció specifikus tesztelésére egyéb próbák elvégzése ajánlott. A Moran-féle *I*-próba kiváltására számos asszimptotikus próba található az irodalomban. A térbeli hiba autokorreláció tesztelésére *Burridge* (1980) fejlesztette ki a következő Lagrange Multiplikátor (LM) próbát:

$$\text{LM-ERR} = (\mathbf{e'We}/s^2)^2/T, \qquad /17/$$

ahol **e** az OLS-reziduumok vektora, **W** térbeli súlymátrix, $s^2 = \mathbf{e'e}/N$ és $T = \mathrm{tr}(\mathbf{W'W} + \mathbf{W}^2)$, ahol tr a mátrix nyoma (trace) mátrix műveletet jelzi. A próbafüggvény χ^2 eloszlást követ, egyes szabadságfokkal.

A térben késleltetett függő változó formájában megjelenő térbeli autokorreláció jelenlétére a következő Lagrange Multiplikátor teszt alkalmazása a legelterjedtebb (*Anselin*; 1988b):

$$\text{LM-LAG} = (\mathbf{e'Wy}/s^2)^2/(\mathrm{RJ}_{\rho-\beta}), \qquad /18/$$

ahol **e** az OLS-reziduumok vektora, **y** a regresszió függő változójának vektora,

$$\mathrm{RJ}_{\rho-\beta} = [T + (\mathbf{WX}\boldsymbol{\beta})'\mathbf{M}(\mathbf{WX}\boldsymbol{\beta})/s^2], \qquad /19/$$

[16] A diagnosztikák jóval részletesebb ismertetéséhez lásd *Anselin* (1988a, 1988b) valamint *Anselin* és *Florax* (1995).

[17] A hipotézisvizsgálattal kapcsolatban további részletekkel *Cliff* és *Ord* (1981) munkája szolgál.

ahol $\mathbf{WX}\boldsymbol{\beta}$ az OLS-modell-előrejelzések térben késleltetett értéke, valamint

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}', \qquad /20/$$

a projekciós mátrix. A próba χ^2 eloszlást követ, egyes szabadságfokkal.

A térbeli hiba autokorrelációnak a térbeli késleltetés modelljében való jelenlétére a következő Lagrange Multiplikátor teszt (*Anselin*; 1988b) használatos:

$$\text{LM-ERRLAG} = (\mathbf{e}'\mathbf{We}/s^2)^2 / [\text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2) - \text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)\mathbf{A}^{-1}\text{var}(\rho)], \qquad /21/$$

ahol **e** a térbeli késleltetés modellje maximum likelihood reziduumainak vektora,

$$\mathbf{A} = (\mathbf{I} - \rho\mathbf{W}), \qquad /22/$$

ahol ρ a térbeli autoregressziós paraméter, melynek becsült aszimptotikus szórásnégyzete var(ρ). A próbafüggyvény χ^2 eloszlást követ, egyes szabadságfokkal. Amennyiben az LM-ERRLAG teszt szignifikáns értéket mutat, akkor a szóban forgó késleltetéses modell nem használható. Megoldásként vagy a regressziós egyenlet szerkezetében való változtatás vagy pedig alternatív súlymátrixok használata javasolt.

Az LM-LAGERR próba a térben késleltetett függő változó formájában megjelenő térbeni dependencia jelenlétét teszteli a hiba autokorreláció modelljében (*Anselin–Bera–Florax–Yoon*; 1996):

$$\text{LM-LAGERR} = (\mathbf{e}_{(N\times 1)}'\mathbf{B}'\mathbf{BWy})^2/(H_\rho - \mathbf{H}_{\theta\rho(1\times[K+2])}\mathbf{Var}(\boldsymbol{\theta})\mathbf{H}'_{\theta\rho(1\times[K+2])}), \qquad /23/$$

ahol **e** a térbeli hiba autokorreláció modellje maximum likelihood reziduumainak vektora, **y** a függő változó vektora,

$$\mathbf{B} = \mathbf{I} - \lambda\mathbf{W}, \qquad /24/$$

$$\boldsymbol{\theta}_{([K+2]\text{x}1)}' = [\boldsymbol{\beta}_{(K\text{x}1)}', \lambda, s^2], \qquad /25/$$

$\mathbf{Var}(\boldsymbol{\theta})$ pedig a $\boldsymbol{\theta}$ paraméter vektor becsült szórásnégyzete:

$$H_\rho = \text{tr}\mathbf{W}^2 + \text{tr}(\mathbf{BWB}^{-1})'(\mathbf{BWB}^{-1}) + $$
$$+ (1/s^2)(\mathbf{BWX}\boldsymbol{\beta})'(\mathbf{BWX}\boldsymbol{\beta})), \qquad /26/$$

$$\mathbf{H}'_{\theta\rho(1\times[K+2])} = [\mathbf{A}_{1(K\times 1)}, \text{A}_2, 0]', \qquad /27/$$

ahol:

$$\mathbf{A}_{1(K\times 1)} = (1/s^2)(\mathbf{BX})'\mathbf{BW}\,\mathbf{X}\boldsymbol{\beta}, \qquad /28/$$

és

$$\text{A}_2 = \text{tr}(\mathbf{WB}^{-1})'\mathbf{BWB}^{-1} + \text{tr}\mathbf{W}^2\mathbf{B}^{-1}. \qquad /29/$$

A próbafüggvény χ^2 eloszlást követ, egyes szabadságfokkal. Amennyiben az LM-LAGERR teszt szignifikáns értéket mutat, akkor a szóban forgó térbeli hiba autokorreláció modell nem használható. Megoldásként vagy a regressziós egyenlet szerkezetében való változtatás vagy pedig alternatív súlymátrixok használata javasolt.

A vizsgált jelenség természetének leginkább megfelelő térbeli súlymátrix kiválasztása minden térökonometriai elemzés kritikus pontja. Habár léteznek tesztek, melyek a mátrixválasztáshoz segítséggel szolgálhatnak (lásd *Anselin*; 1984), általánosan alkalmazott egyértelmű megoldás a problémára még nem született. A térbeli autokorreláció tesztelése során ezért több, a teret különbözőképpen strukturáló súlymátrixot alkalmazunk. Amennyiben specifikációs probléma merül fel, mindig az a mátrix választandó, amelyiknél a térbeli autokorreláció problémája a leghatározottabban jelentkezik.

*

A közgazdaságtan fokozódó érdeklődése a gazdasági jelenségek térbeli vonatkozásai iránt előtérbe helyezi a térbeli modellek ökonometriai tesztelése során megbízhatóan alkalmazható empirikus módszereket. Tanulmányomban a térökonometria szemléletének bemutatására és néhány, a gyakorlatban gyakran használt térökonometriai modell és specifikációs teszt ismertetésére vállalkoztam. Az itt közöltek természetesen semmiképpen sem tekinthetők teljesnek. A térökonometria számos jelenlegi kutatási problémája (mint például a térbeli autokorreláció problémájának kezelése az egyéb, a lineáristól eltérő modellekben) kívül maradt az elemzés keretein. Befejezésként essék néhány szó a szoftverkínálatról.[18] A nemzetközi piacon létező legteljesebb programcsomag a térökonometriai és térstatisztikai elemzésekre kifejlesztett SpaceStat szoftver (*Anselin*; 1992, 1998). Gyakran használt még *Pace* és *Barry* (1998) Matlab rutinja a térökonometriai modellek maximum likelihood becslésére, valamint a Mathsoft S-Plus-hoz elkészített S+Spatialstats (*MathSoft*; 1996) a térbeli hiba autokorreláció modelljének becslésére.

IRODALOM

ACS Z. – VARGA A. (2001): Térbeliség, endogén növekedés és innováció. *Tér és Társadalom*, 14. évf. 23–39. old.

AOKI, M. (1994): New macroeconomic modelling approaches. Hierarchical dynamics and mean-field approximation. *Journal of Economic Dynamics and Control*, 18. évf. 865–877. old.

AOKI M. (1996): *New approaches to macroeconomic modelling*. Cambridge University Press, Cambridge.

ANSELIN, L. (1984): Specification tests on the structure of interaction in spatial econometric models. *Papers of the Regional Science Association*, 54. évf. 165-182. old.

ANSELIN, L. (1988a): *Spatial econometrics: methods and models*. Kluwer Academic Publishers, Boston.

ANSELIN, L. (1988b): Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical Analysis*, 20. évf. 1-17. old.

ANSELIN, L. (1992): *SpaceStat, a software program for the analysis of spatial data.* National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA.

ANSELIN, L. (1995): Local indicators of spatial association – LISA. *Geographical Analysis*, 27. évf. 93-115. old.

ANSELIN, L. (1998): *SpaceStat Version 1.90.* http://www.spacestat.com

ANSELIN, L. (2001): Spatial econometrics. In: *Baltagi, B.* (szerk.): *A companion to theoretical econometrics*. Basil Blackwell, Oxford, 310–330. old.

ANSELIN, L. – FLORAX, R. (szerk.) (1995): *New directions in spatial econometrics*. Springer-Verlag, Berlin, 21–74. old.

ANSELIN, L. – BERA, A. – FLORAX, R. – YOON, M. (1996): Simple diagnostics tests for spatial dependence. *Regional Science and Urban Economics*, 26. évf. 77–104. old.

ANSELIN, L. – HUDAK, S. (1992): Spatial econometrics in practice. A review of software options. *Regional Science and Urban Economics*, 22. évf. 509-536. old.

[18] A szoftverkínálat részletes bemutatására lásd *Anselin* és *Hudak* (1992).

ANSELIN, L. – REY, S. (1991): Properties of tests for spatial dependence in linear regression models. *Geographical Analysis*, 23. évf. 112-131. old.

ANSELIN, L. – REY, S. (szerk.) (1997): A special issue on spatial econometrics. *International Regional Science Review*, 20. évf.

ANSELIN, L. – VARGA A. – ACS Z. (1997): Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics*, 42. évf. 422–448. old.

ATEN, B. (1996): Evidence of spatial autocorrelation in international prices. *Review of Income and Wealth*, 42. évf. 149–163. old.

BALTAGI, B. – LI, D. (2001): Prediction in the panel data model with spatial correlation. In: *Anselin, L. – Florax, R.* (szerk.): *Advances in spatial econometric*s. (Kézirat.)

BELL, K. – BOCKSTAEL, N. (1999): Applying the generalized moments estimation approach to spatial problems involving micro-level data. *The Review of Economics and Statistics*, 81. évf.

BERON, K. – VIJVERBERG, W. (2001): Probit in a spatial context: a Monte Carlo approach. In: ANSELIN, L. – FLORAX, R. (szerk.): *Advances in spatial econometric*s. (Kézirat.)

BOLDUC, D. – FORTIN, B. – GORDON, S. (1997): Multinomial probit estimation of spatially interdependent choices: an empirical comparison of two new techniques. *International Regional Science Review*, 20. évf. 77–101. old.

BURRIDGE, P. (1980): On the Cliff-Ord test for spatial correlation. *Journal of the Royal Statistical Society B*, 42. évf. 107-108. old.

CASE, A. (1991): Spatial patterns in household demand. *Econometrica*, 59. évf. 953–965. old.

CASE, A. (1992): Neighborhood influence and technological change. *Regional Science and Urban Economics*, 22. évf. 491–508. old.

CASE, A. – ROSEN, H. – HINES, J. (1993): Budget spillovers and fiscal policy interdependence: evidence from the States. *Journal of Public Economics*, 52. évf. 285–307. old.

CASETTI, E. (1997): The expansion method, mathematical modelling and spatial econometrics. *International Regional Science Review*, 20. évf. 9–32. old.

CLIFF, A. – ORD, J. (1981): *Spatial processes, models and applications*. Pion, London.

CONLEY, T. (1996): *Econometric modelling of cross-sectional dependence*. PhD Dissertation. Department of Economics, University of Chicago, Chicago, IL.

CRESSIE, N. (1993): *Statistics for spatial data*. Wiley, New York.

DRISCOLL, J. – KRAAY, A. (1998): Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80. évf. 549–560. old.

FUJITA, M. – KRUGMAN, P. – VENABLES, A. (1999): *The spatial economy*. MA, MIT Press, Cambridge.

GRIFFITH, D. (1988): *Advanced spatial statistics.* Kluwer Academic Publishers, Boston.

GETIS, A. – ORD, J. (1992): The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24. évf. 189-205. old.

GREENE, W. (1993): *Econometric analysis.* MacMillan Publishing Company, New York.

HAINING, R. (1990): *Spatial data analysis in the social and environmental sciences*. Cambridge University Press, Cambridge.

HOLTZ-EAKIN, D. (1994): Public-sector capital and the productivity puzzle. *Review of Economics and Statistics*, 76. évf. 12–21. old.

HUBERT, G. – CONSTANZO, C. – GALE, N. (1985): Measuring association between spatially defined variables: an alternative procedure. *Geographical Analysis*, 17. évf. 36-46. old.

HEAGERTY, P. – LELE, S. (1998): A composite likelihood approach to binary spatial data. *Journal of the American Statistical Association*, 93. évf. 1099–1111. old.

KELEJIAN, H. – ROBINSON, D. (1993): A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model. *Papers in Regional Science*, 72. évf. 297–312. old.

KELEJIAN, H. – PRUCHA, I. (1998): A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17. évf. 99–121. old.

KELEJIAN, H. – PRUCHA, I. (1999): A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*.

KŐRÖSI G. – MÁTYÁS L. – SZÉKELY I. (1990): *Gyakorlati ökonometria*. Közgazdasági és Jogi Könyvkiadó, Budapest.

KRISTENSEN, G. (szerk.) (1996): *Symposium on the expansion method in the context of the family of models and methodologies with a focus on parametric variation, I-III*. Department of Economics, Odense University. Odense.

KRUGMAN, P. (1991): Increasing returns and economic geography. *Journal of Political Economy*. 99. évf. 483–499. old.

LAND, K. – DEANE, G. (1992): On the large-sample estimation of regression models with spatial-or network-effects terms: a two stage least squares approach. In: *Marsden, P.* (szerk.) *Sociological Methodology*, Jossey-Bass, San Francisco. 221–248. old.

LESAGE, J. (1999): Spatial econometrics. In: *Web book of regional science*. http://www.rri.wvu.edu/WebBook/LeSage/spatial/spatial.html

LUCAS, R. (1988): On the mechanics of economic development. *Journal of Monetary Economics* 22, évf. 3–42. old.

LUNDVAL, B. (szerk.) (1992): *National systems of innovation*. Pinter, London.

MCMILLEN, D. (1992): Probit with spatial autocorrelation. *Journal of Regional Science* 32, évf. 335–348. old.

MURDOCH, J. – RAHMATIAN, M. – THAYER, M. (1993): A spatially autoregressive median voter model of recreation expenditures. *Public Finance Quarterly*, 21. évf. 334–350. old.

MURDOCH, J. – SANDLER, T. – SARGENT, K. (1997): A tale of two collectives: sulphur versus nitrogen oxides emission reduction in Europe. *Economica*, 64. évf. 281–301. old.

MORAN, P. (1948): The interpretation of statistical maps. *Journal of the Royal Statistical Society B*, 10. évf. 243-251. old.

MATHSOFT (1996): *S+SpatialStats user's manual for Windows and Unix*. WA: MathSoft, Inc., Seattle.

NELSON, G. – HELLERSTEIN, D. (1997): Do roads cause deforestation? Using satellite images in econometric analysis of land use. *American Journal of Agricultural Economics*, 79. évf. 80–88. old.

NELSON, R. (szerk.) (1993): *National innovation systems*. Oxford University Press, Oxford, New York.

NEMES NAGY J. (1998): *Tér a társadalomkutatásban*. Ember–Település–Régió, Hilscher Rezső Szociálpolitikai Egyesület, Budapest.

ORD, J. (1975): Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70. évf. 120-126. old.

PAELINCK, J. – KLAASSEN, L. (1979): *Spatial econometrics*. Saxon House, Farnborough.

PACE, R. – BARRY, R. (1998): *Spatial statistics toolbox 1.0*. Real Estate Research Institute, Lousina State University, Baton Rouge, LA.

PINKSE, J. – SLADE, M. (1998): Contracting in space: an application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85. évf. 125–154. old.

PINTÉR J. (1991): A heteroszkedaszticitás diagnosztizálása. *Statisztikai Szemle*, 69. évf. 1. sz. 16–36. old.

ROMER, P. (1990): Endogenous technological change. *Journal of Political Economy*, 98. évf. S71–S102. old.

TOBLER, W. (1979): Cellular geography. In: *Gale, S. – Olsson, G.* (szerk.): *Philosophy and geography*. Reidel, Dordrecht, 379-386. old.

TOPA, G. (1996): Social interactions, local spillovers and unemployment. PhD Dissertation, Department of Economics, University of Chicago.

UPTON, G. – FINGLETON, B. (1985): *Spatial data analysis by example.* Vol. 1: *Point pattern and quantitative data*. Wiley, New York.

VARGA A. (1998): *University research and regional innovation: A spatial econometric analysis of academic technology transfers.* Kluwer Academic Publishers, Boston.

SUMMARY

Spatial econometrics is a subfield of econometrics dealing with the incorporation of spatial interaction and spatial heterogeneity into regression analysis. This paper provides a general introduction to spatial econometrics for the Hungarian professional audience by focusing primarily on the distinguishing characteristics of the methodology, on estimation of the most widely used linear spatial econometric models as well as on problems related to testing spatial autocorrelation.