Kointegračná analýza v ekonometrii

Martin Lukáčik – Juraj Pekár

Prognózovanie budúceho vývoja ekonomických ukazovateľov, ktoré sú v centre záujmu užívateľov informácií, býva často najdôležitejšou úlohou ekonomických analytikov. Tí musia pri svojom výskume rozhodnúť, aké metódy použijú pre vytvorenie najvhodnejších predpovedí. Hlavným problémom je *existujúca dichotómia* pri konštrukcii krátkodobých a dlhodobých prognóz.

Dlhodobý vývoj je determinovaný ekonomickými súvislosťami, preto sa jeho predpovedanie obvykle zveruje regresným ekonometrickým modelom s logickou štruktúrou použitých premenných na základe ekonomickej teórie. Takýmto spôsobom sme odhadovali modely aj v minulom roku a tento postup sledujeme aj v tomto roku na štvrťročných údajoch (kvôli ich dostatočnému počtu a kvôli dlhodobejšiemu trendu, ktorý je v nich zachytený).

Pre predvídanie *krátkodobého vývoja* sa využíva skôr niektorá z extrapolačných metód, lebo lepšie zachytáva existujúcu dynamiku v skúmanej štruktúre. Opäť bola súčasťou minuloročného výskumu a je aj v tohoročnom výstupe prognóza vytvorená na základe modelov využívajúcich exponenciálne vyrovnávanie pri mesačných údajoch o spotrebe (slúži pre preskúmanie bezprostredne nasledujúcich období).

V tejto časti sa pokúsime vysvetliť možné riešenie tohto problému, ktoré priniesli kointegračné modely s korekčným členom. Pre pochopenie tohto náročnejšieho postupu si musíme ozrejmiť nielen výslednú koncepciu modelov s korekčným členom, ktorú využijeme pre tvorbu predpovedí, ale aj existenciu nestacionarity a spôsoby jej overovania v časových radoch a rovnako aj princíp kointegrácie a postupy jej zisťovania.

Ak by sme mali celú vec stručne zhrnúť: ekonomická teória ponúka rozsiahle možnosti výberu pre tvorcu modelu, ale obvykle je nepostačujúca z hľadiska popisu štruktúry dynamiky, ktorá je nevyhnutná pre zachytenie mnohých drobných okamžitých reakcií. Preto klasické statické ekonometrické modely obvykle poskytujú z krátkodobého hľadiska pri porovnaní s extrapolačnými metódami slabšie výsledky pri prognózach. V rámci ekonometrie tento fakt priniesol do centra jej záujmu *modely vektorovej autoregresie* s testovaním stacionarity. Následne sa pre nestacionárne časové rady (obvyklé v ekonómii) začala využívať analýza kointegrácie, ktorá viedla k tvorbe *modelov s korekčným členom*

Modely vektorovej autoregresie sú modelmi, ktoré riešia prepojenie jednotlivých skúmaných radov ako závislosť od navzájom posunutých hodnôt, čím zabezpečia dôkladnú previazanosť analyzovaných radov. Sú špecifickým prípadom viacrovnicového modelu, v ktorom každá vysvetľovaná premenná závisí od vlastných oneskorených hodnôt a od oneskorených hodnôt všetkých do modelu zavedených premenných. Považujú sa za viacradovú analógiu Boxovej-Jenkinsovej metodológie ARIMA modelov opisovanú v metodologickej časti minulý rok.

Pre takýto typ modelov je dôležité *skúmanie stacionarity* použitých radov, lebo práve od stacionarity respektíve stupňa nestacionarity závisí, s akým typom údajov môžeme realizovať výpočty. Ak sú pôvodné napozorované rady stacionárne, tak tieto sa dajú priamo použiť pre odhad neznámych parametrov modelov, s ktorými vieme prognózovať budúci vývoj. Ak sú pôvodné rady nestacionárne, čo je častý jav pri mnohých ekonomických premenných, tak podľa použitého stupňa diferencie (čo je obvyklý spôsob pre získanie stacionárneho radu) sa v modeloch môžu použiť iba takto diferencované rady. Ale namiesto dlhodobých vzťahov medzi radmi sa skúma iba vzťah medzi ich rastom alebo tempom rastu.

Preto aby sme nestratili informáciu o dlhodobom vývoji, bola navrhnutá *koncepcia kointegrácie*. Teda možnosť využiť za istých okolností v analýze aj nestacionárne rady. Tento postup umožňuje skombinovať do jedného celku dve matematicky odlišné veci. Dlhodobú informáciu v pôvodných nestacionárnych radoch zahrnutú v kointegračnom vzťahu, ktorý je stacionárny spolu s krátkodobou informáciou obsiahnutou v dynamike diferencovaných stacionárnych radov. Tým sa vytvorí model rovnováhy s korekčným členom.

Čo je to stacionarita časového radu?

Aby sme získali uspokojivý model vyhovujúci ekonomickým aj štatistickým predpokladom, je dôležité spoznať správanie sa trendu v modelovaných premenných. Teda poznať odpovede na otázky: Sú trendy prítomné alebo absentujú v skúmaných premenných? Ak je trend prítomný, aký je, náhodný alebo nie?

Stacionaritu môžeme chápať *v striktnom zmysle* a požadovať nemennosť skúmanej náhodnej premennej v čase, bez ohľadu na to, aký si z nej vyberieme časový úsek. Pre praktický výskum postačuje skúmanie stacionarity *v slabom zmysle* ako nemennosť základných štatistických mier skúmanej náhodnej premennej v čase, teda jej strednej hodnoty, rozptylu a kovariancie.

Ekonomické časové rady často obsahujú trend a preto bežným javom je nestacionarita procesu vzhľadom na priemer. Ak je takýto trend lineárny, potom jednoduché prvé diferencie proces stacionarizujú. Ak premenná vykazuje známky rôzneho rozptylu hodnôt pri nižších a rôzneho pri vyšších úrovniach, potom hovoríme o nestacionarite procesu vzhľadom na rozptyl. Ako vhodná transformácia pre získanie stacionárneho radu sa uvádza logaritmovanie. Prezentované typy časových radov môžeme rozdeliť na dve základné skupiny, trendovo stacionárne procesy a diferenčne stacionárne procesy.

Lineárny deterministický trend v premennej uvažujeme vtedy, keď pre ňu platí:

$$y_t = \alpha_0 + \alpha_1 t + u_t, \qquad \text{kde } u_t \sim N\left(0, \sigma_u^2\right) \text{ teda náhodná zložka je stacionárna.}$$

Časový rad y_t je zjavne nestacionárny proces. Jednoduchou matematickou úpravou, odrátaním deterministickej časti modelu od procesu, sa proces stane stacionárnym:

$$\left(y_t - \alpha_0 - \alpha_1 t\right) = u_t$$

Pôvodný proces y_t, ktorý môžeme stacionarizovať takýmto spôsobom, nazývame *trendovo stacionárny proces* (TSP).

Proces náhodnej prechádzky (random walk) rozoznávame, keď platí:

$$y_t = y_{t-1} + u_t,$$

pričom y_t je stacionárny proces vzhľadom na priemer, ale určite nie vzhľadom na rozptyl, lebo po rekurentnom dosadení pre priemer a pre rozptyl platia nasledovné vzťahy:

$$E\left(y_t\right) = E\left(u_t\right) + E\left(u_{t-1}\right) + E\left(u_{t-2}\right) + \ldots = 0 + 0 + 0 + \ldots = 0$$

$$\operatorname{var}\left(y_t\right) = \operatorname{var}\left(u_t\right) + \operatorname{var}\left(u_{t-1}\right) + \operatorname{var}\left(u_{t-2}\right) + \ldots = \sigma_u^2 + \sigma_u^2 + \sigma_u^2 + \ldots = \infty$$

Vidíme zrejmý rozdiel oproti TSP, kde:

$$E\left(y_t\right) = E\left(u_t\right) = 0$$

$$\operatorname{var}\left(y_t\right) = \operatorname{var}\left(u_t\right) = \sigma_u^2$$

Teda rozptyl trendovo stacionárnych procesov je konečná hodnota, ale rozptyl náhodnej prechádzky je nekonečný.

Proces náhodnej prechádzky sa často uvažuje s konštantou, potom sa proces nazýva proces náhodnej prechádzky s driftom – ide o tzv. *stochastický trend*, a platí:

$$y_t = \beta_0 + y_{t-1} + u_t$$

Jednoduchou diferenciou sa proces stane stacionárnym:

$$\Delta y_t = y_t - y_{t-1} = (1-L)\, y_t = \beta_0 + u_t .$$

Takýto proces nazývame *diferenčne stacionárny proces* (DSP) alebo kvôli špeciálnemu tvaru polynómu oneskorenia (koreň sa rovná 1) sa nazýva aj *procesom s jednotkovým koreňom*.

Vo vzťahu sme použili operátor oneskorenia L, ktorý zjednodušuje zápis pri dynamických vzťahoch a pre ktorý platí:

$$LX_t = X_{t-1},\ L^2X_t = L(LX_t) = L(X_{t-1}) = X_{t-2},\ \ldots,\ L^s X_t = X_{t-s}.$$

Pre modelovanie je dôležité rozlišovať medzi stacionárnymi a nestacionárnymi procesmi. Pre porovnanie uveďme oproti už uvedenému nestacionárnemu procesu náhodnej prechádzky stacionárny autoregresný proces prvého rádu, teda $AR(1)$ proces:

$$y_t = \beta_1 y_{t-1} + u_t, \quad \text{kde } u_t \sim N\left(0, \sigma_u^2\right) \text{ a } |\beta_1| < 1 .$$

Pomocou operátora oneskorenia a s využitím vzťahu pre sumáciu môžeme ukázať, že:

$$y_t = \frac{1}{1-\beta_1 L} u_t = u_t + \beta_1 u_{t-1} + \beta_1^2 u_{t-2} + \ldots .$$

Pre priemer a pre rozptyl platia nasledovné vzťahy:

$$E(y_t) = E(u_t) + \beta_1 E(u_{t-1}) + \beta_1^2 E(u_{t-2}) + \ldots = 0 + \beta_1 0 + \beta_1^2 0 + \ldots = 0$$

$$\operatorname{var}(y_t) = \operatorname{var}(u_t) + \beta_1^2 \operatorname{var}(u_{t-1}) + \beta_1^4 \operatorname{var}(u_{t-2}) + \ldots = \sigma_u^2 \left(1 + \beta_1^2 + \beta_1^4 + \ldots\right) = \frac{\sigma_u^2}{1-\beta_1^2}$$

Dôležitým ekonomickým rozdielom je reakcia na šok v náhodnej zložke (čo predstavuje neočakávaný zásah), ktorá sa pri stacionárnych procesoch postupne stráca, ale pri nestacionárnych procesoch zostáva šok dlhodobo prítomný (má permanentný vplyv).

Preto je dôležité uviesť, že nestacionárne rady rozlišujeme podľa počtu diferencií potrebných pre získanie stacionarity. Časový rad označujeme ako *integrovaný rádu k* v tom prípade, ak nespĺňa podmienky stacionarity, ale pomocou k-násobného diferencovania je ho možné previesť na stacionárny rad. Integrovaný rad rádu k sa označuje $y_t \sim I(k)$. Preto diferencovanie integrovaného procesu y_t rádu 1, ktorý obsahuje práve 1 jednotkový koreň, znamená jeho stacionarizáciu:

$$\Delta y_t = (1-L)\, y_t = u_t \quad \Rightarrow \quad \Delta y_t \sim I(0).$$

Ak je rad y_t integrovaný rádu 2 (obsahuje 2 jednotkové korene), je potrebné, aby bol kvôli stacionarizácii diferencovaný dvakrát:

$$(1-L)^2\, y_t = \Delta y_t - \Delta y_{t-1} = \Delta^2 y_t = u_t \quad \Rightarrow \quad y_t \sim I(2), \Delta y_t \sim I(1), \Delta^2 y_t \sim I(0).$$

Ekonomická interpretácia je potom taká, že y_t má trend v tempe rastu, Δy_t má lineárny trend a $\Delta^2 y_t$ je stacionárna. V ekonomických radoch je integrovanosť vyššia ako rádu 2 veľmi zriedkavá. Najčastejšia je integrovanosť rádu 1. Integrovanosť rádu 2 sa vyskytuje občas napríklad v cenových premenných v niektorých výrazných inflačných obdobiach a podobne.

Testovanie stacionarity

Testovanie stacionarity sa používa v prevedenej forme, keď v autoregresnom modeli je od obidvoch strán odpočítaná hodnota y_{t-1}, čím získame model napríklad v tvare:

$$\Delta y_t = (\beta_1 - 1)\, y_{t-1} + u_t = \delta y_{t-1} + u_t \, .$$

Dickeyov-Fullerov test (DF test) po odhade parametra δ a výpočte Studentovej štatistiky t_δ, porovnáva túto hodnotu s tabelovanou τ hodnotou. Nižšia záporná hodnota štatistiky t_δ znamená zamietnutie nulovej hypotézy a teda stacionaritu radu. Pri nezamietnutí nulovej hypotézy o nestacionarite sa môže proces testovania opakovať s modelmi vyšších diferencií, až kým nie je nulová hypotéza zamietnutá. Podľa stupňa diferencie premennej pri zamietnutej nulovej hypotéze o jej parametri sa potom určí stupeň integrácie časového radu.

Autori testu si zároveň všimli, že kritické hodnoty štatistiky závisia od typu skúmaného modelu. Preto tabelovali kritické hodnoty ešte zvlášť pre model obsahujúci absolútny člen a aj osobitne pre model s absolútnym členom a lineárnym časovým trendom. V prípade, ak je náhodná zložka autokorelovaný proces, klasický Dickeyov-Fullerov test zlyháva. Preto títo autori navrhli rozšíriť pôvodný autoregresný model prvého rádu na všeobecný autoregresný model rádu p, ktorý eliminuje problém autokorelácie práve pridanými autoregresnými členmi. Rozšírený model s absolútnym členom a lineárnym časovým trendom potom vyzerá nasledovne:

$$\Delta y_t = \mu + \alpha_1 t + \delta y_{t-1} + \sum_{j=2}^{p} \beta_{j-1} \Delta y_{t-j+1} + u_t \, .$$

Testovanie sa nazýva rozšírený Dickeyov-Fullerov test (ADF). Pôvodne tabelované kritické hodnoty platia aj pre rozšírené typy modelov. Rozhodnutie o veľkosti použitého autoregresného rádu p, ktorý najlepšie reprezentuje autokoreláciu, sa odporúča vykonať podľa Schwarzovho informačného kritéria (SC).

Vzhľadom na vyšší počet možných kombinácií pre testovanie sa používa *rozšírený Dickeyov-Fullerov test* v nasledovnej alebo analogickej sekvenčnej procedúre postupu:

1. Test začína odhadom najvšeobecnejšieho modelu s absolútnym členom a lineárnym časovým trendom.

2. Pre testovanie prítomnosti trendu sa využije podmienená hypotéza $\alpha_1 = 0$ ak $\delta = 0$ alebo združená hypotéza $\alpha_1 = \delta = 0$. Prijatie hypotézy o nulovom parametri α_1 znamená, že v rade nie je deterministický trend a pokračuje sa krokom 4.
3. Po zamietnutí hypotézy o nulovom parametri α_1 sa vykoná test o nulovom parametri δ. Jeho nevýznamnosť znamená, že rad je proces *náhodnej prechádzky s trendom*. V rámci tohto záveru sa dá dodatočne zistiť prítomnosť posunu a využije sa podmienená hypotéza $\mu = 0$ ak $\delta = 0$ alebo združená hypotéza $\mu = \alpha_1 = \delta = 0$.
4. Prejdeme k testovaniu modelu obsahujúcom absolútny člen ale bez trendu (vypadne $\alpha_1 t$).
5. Pre testovanie prítomnosti posunu sa v tomto modeli využíva podmienená hypotéza $\mu = 0$ ak $\delta = 0$ alebo združená hypotéza $\mu = \delta = 0$. Prijatie hypotézy o nulovom parametri μ znamená, že rad je proces *náhodnej prechádzky* (bez driftu a trendu).
6. Po zamietnutí hypotézy o nulovom parametri μ sa vykoná test o nulovom parametri δ. Jeho nevýznamnosť znamená, že rad je proces *náhodnej prechádzky s driftom*.
7. Prijatie hypotézy o významnosti parametra δ znamená v testovacom kroku 3 alebo 6 zastavenie procedúry s konštatovaním, že *rad je stacionárny*. Po zistení stacionarity sa klasickým *t*-testom dá overiť, či rad obsahuje posun alebo deterministický trend.

Princíp kointegrácie

Pri analýze ekonomických časových radov sa vyskytujú situácie, keď základné skúmané štatistiky odhadnutého modelu naznačujú veľmi dobré výsledky (štatisticky významné parametre a vysoké R^2) s výnimkou Durbinovej-Watsonovej štatistiky. Takéto odhady sú zavádzajúce, ak časové rady v rovnici obsahujú trend alebo nestacionárny náhodný proces, lebo odhady metódy najmenších štvorcov nie sú konzistentné. Ako jednoduché pravidlo diagnostiky signalizujúce falošnú regresiu je uvedený vzťah $R^2 > DW$. Preto vzniká otázka, či sa nemôžu napriek nestacionarite využiť pri analýze pôvodné úrovňové premenné. Odpoveďou sa stala koncepcia kointegrácie.

Dva nestacionárne rady sa považujú za *kointegrované* rádu (d, b), pričom $d \geq b \geq 0$, ak oba rady sú integrované rádu d a existuje lineárna kombinácia týchto dvoch radov, ktorá je integrovaná rádu $(d - b)$. Matematicky sa dá celá koncepcia zapísať nasledovne:

nech $y_t \sim I(d) \wedge x_t \sim I(d)$, potom ak $\exists(\lambda_1 y_t + \lambda_2 x_t) \sim I(d-b) \Rightarrow y_t, x_t \sim CI(d,b)$.

Označenie *CI* symbolizuje kointegráciu.

Predmetom záujmu je taký prípad kointegrácie, kde $d = b$, čo znamená stacionaritu lineárnej kombinácie $\lambda_1 y_t + \lambda_2 x_t$ a prináša možnosť odhadovať premenné bez diferencovania.

Uvedenú koncepciu je možné zovšeobecniť pre prípad viacerých premenných. Potom k radov x_{1t}, x_{2t}, ... x_{kt} považujeme za *kointegrované* rádu (d, b), pričom $d \geq b \geq 0$, ak všetky rady sú integrované rádu d a existuje lineárna kombinácia týchto radov, ktorá je integrovaná rádu $(d - b)$. Rozdiel oproti prípadu s dvomi premennými predstavuje kointegrujúci vektor, lebo v prípade viacerých premenných nie je jednoznačný. Totiž pre prípad k premenných môže existovať až k–1 lineárne nezávislých vektorov. Podľa počtu kointegrujúcich vektorov sa potom rozoznáva *hodnosť kointegrácie*.

Spôsob zisťovania kointegrácie Engleovym-Grangerovym testom

Najznámejším a najjednoduchším testom prítomnosti kointegrácie medzi dvomi premennými je *Engleova-Grangerova testovacia procedúra*. Test prebieha nasledovne:

V prvom kroku sa pomocou metodológie testovania jednotkových koreňov zistí rád integrácie obidvoch premenných. Môžu nastať tri rozličné prípady:

1. ak rád integrácie oboch premenných je rovnaký (mimo stacionarity), test kointegrácie pokračuje ďalším krokom,
2. ak rád integrácie dvoch premenných je rozdielny, premenné nie sú kointegrované,
3. ak sú obidve premenné stacionárne, testovacia procedúra končí.

V druhom kroku sa pomocou bežnej metódy najmenších štvorcov odhadne rovnica dlhodobej rovnováhy nazývaná tiež *rovnica kointegrácie*:

$$y_t = \beta_0 + \gamma_0 x_t + \varepsilon_t .$$

Aby boli premenné y_t a x_t kointegrované, musia byť náhodné poruchy v rovnici dlhodobej rovnováhy stacionárne.

Preto sa *v treťom kroku* takisto využije testovanie jednotkových koreňov, napríklad prostredníctvom rozšíreného Dickeyovho-Fullerovho testu, pre časový rad reziduálov e_t získaných z odhadu rovnice dlhodobej rovnováhy. Testujeme model:

$$\Delta e_t = \delta e_{t-1} + \sum_{j=2}^{p} \delta_j \Delta e_{t-j+1} + \nu_t .$$

Prijatie nulovej hypotézy o nestacionarite radu reziduálov vyplývajúcej z akceptácie $\delta = 0$, teda ak t_δ je väčšie ako tabelovaná hodnota, znamená, že premenné y_t a x_t nie sú kointegrované. Zamietnutie značí kointegráciu skúmaných premenných.

Ak nejaké premenné sú kointegrované, potom medzi nimi existuje dlhodobý vzťah. Z krátkodobého hľadiska môžu byť v nerovnováhe. Dynamika krátkodobých nerovnovážnych vzťahov medzi premennými sa dá vyjadriť pomocou *modelu s korekčným členom*. Takýto model, ktorý spája krátkodobé a dlhodobé správanie premenných sa môže všeobecne zapísať:

$$\Delta y_t = f\left(\Delta y_{t-1}, \Delta y_{t-2}, \ldots, \Delta x_t, \Delta x_{t-1}, \ldots\right) + \lambda e_{t-1} + u_t ,$$

kde $y_t \sim I(1), x_t \sim I(1)$ sú nestacionárne rady, ale kointegrované. Odchýlky od dlhodobej rovnováhy reprezentuje $e_{t-1} = y_{t-1} - \beta_0 - \gamma_0 x_{t-1} \sim I(0)$, teda oneskorený reziduál, ktorý je stacionárny. Zároveň λ je koeficient krátkodobého prispôsobovania.

Zjavným nedostatkom Engleovej a Grangerovej procedúry je, že sa nedá použiť v prípade použitia viacerých premenných a predpokladu viacerých kointegračných vzťahov. Tieto problémy rieši Johansenova procedúra, ktorá pracuje s viacerými radmi súčasne.

Johansenova procedúra

Uvažujme nasledovný vektorovo autoregresný (VAR) model bez absolútneho člena:

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + \ldots + \mathbf{A}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t = \mathbf{y}_t = \sum_{j=1}^{p} \mathbf{A}_j \mathbf{y}_{t-j} + \boldsymbol{\varepsilon}_t .$$

Predpokladajme, že všetky premenné sú integrované rádu 1, potom môžeme model zapísať:

$$\Delta\mathbf{y}_t = \mathbf{B}\mathbf{y}_{t-1} + \sum_{j=1}^{p} \mathbf{B}_j \Delta\mathbf{y}_{t-j+1} + \mathbf{v}_t ,$$

kde $\mathbf{B} = -\left(\mathbf{I} - \mathbf{A}_1 - \mathbf{A}_2 - \ldots - \mathbf{A}_p\right) = -\left(\mathbf{I} - \sum_{i=1}^{p} \mathbf{A}_i\right)$ a $\mathbf{B}_j = -\sum_{i=j+1}^{p} \mathbf{A}_i$ pre j = 1, 2, ..., p–1.

Pretože $\mathbf{y}_{t-j}$ sú všetko rady rovnako integrované rádu jedna, potom musia byť premenné $\Delta\mathbf{y}_{t-j}$ stacionárne. Ak existujú medzi nimi kointegračné vzťahy usporiadané v matici $\mathbf{B}$, tak aj $\mathbf{B}\mathbf{y}_{t-1}$ je stacionárne a model môže byť konzistentne odhadnutý.

Johansen zhrnul nasledovné závery:

– Ak hodnosť matice $\mathbf{B}$ je nula, potom všetky prvky matice sú nulové. Mechanizmus s korekčným členom $\mathbf{B}\mathbf{y}_{t-1}$neexistuje a rovnako ani dlhodobý vzťah. Premenné nie sú kointegrované a VAR model by mal byť formulovaný pomocou diferencií.

– Ak hodnosť matice $\mathbf{B}$ je rovná m, všetky jej riadky sú lineárne nezávislé a vektorový proces $\mathbf{y}_t$ je stacionárny, lebo všetky premenné sú integrované rádu 0. VAR model by mal byť formulovaný v pôvodných úrovniach premenných.

– Ak hodnosť matice $\mathbf{B}$ je rovná $r < m$, potom jej riadky nie sú lineárne nezávislé a maticu $\mathbf{B}$ môžeme zapísať ako $\mathbf{B} = \mathbf{D}\mathbf{C}^{\mathrm{T}}$, teda ako súčin *kointegrujúcej matice* $\mathbf{C}$, ktorej stĺpce zodpovedajú vektorom kointegrácie a *matice prispôsobenia* $\mathbf{D}$. V prípade, keď $\mathbf{y}_t \sim I(1)$, tak $\mathbf{C}^{\mathrm{T}}\mathbf{y}_{\mathbf{t}} \sim I(0)$ a hodnosť matice $\mathbf{B}$ je určená počtom kointegrujúcich vektorov. VAR model by mal byť formulovaný ako vektorový model s korekčným členom.

Postup Johansenovej procedúry určenia počtu kointegrujúcich vektorov:

1. Pomocou ADF testu sa určí rád integrácie všetkých premenných.
2. Pomocou SC, prípadne na základe ďalších testov, sa určí rád p vektorovo autoregresného modelu formulovaného v pôvodných úrovniach.
3. Odhadne sa model, kde $\Delta\mathbf{y}_t$ je vysvetlená pomocou $\Delta\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-2}, \ldots, \Delta\mathbf{y}_{t-p+1}$. Na základe Johansenových záverov o hodnosti matice $\mathbf{B}$ bolo uvedené, že hodnosť kointegrujúcej matice je rovná počtu kointegrujúcich vektorov. Úloha zistenia hodnosti kointegrujúcej matice r sa podľa algebraickej teórie zredukuje na test významnosti charakteristických koreňov. Pre testovanie sa využívajú dve základné štatistiky koeficientu vierohodnosti. Prvou štatistikou je test stopy a druhou štatistikou je test maxima charakteristických hodnôt. V oboch testoch sa testovanie končí prvou hypotézou s nevýznamným výsledkom. Zodpovedajúca hodnota r z nulovej hypotézy je hodnosťou kointegrujúcej matice.

S tretím krokom Johansenovej procedúry je spojená jedna komplikácia. Zámerne sme neuviedli tvar odhadovaného modelu. Dôvodom je, že v tomto kroku sa rozhodujeme o prítomnosti konkrétnych deterministických zložiek v modeli. Ich výber je dôležitý preto, lebo ovplyvňuje rozdelenie relevantných štatistík používaných pri kointegračnej analýze. Ak uvažujeme všeobecný vektorový model s korekčným členom (VECM), kde $\mathbf{B} = \mathbf{D}\mathbf{C}^{\mathrm{T}}$ v tvare:

$$\Delta\mathbf{y}_t = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 t + \mathbf{B}\mathbf{y}_{t-1} + \sum_{j=1}^{p-1} \mathbf{B}_j \Delta\mathbf{y}_{t-j} + \mathbf{Q}\mathbf{x}_t + \boldsymbol{\varepsilon}_t,$$

tak existuje päť rôznych možných špecifikácií deterministických zložiek:

1. $\boldsymbol{\alpha}_0 = \mathbf{0}, \boldsymbol{\alpha}_1 = \mathbf{0}$ – čo znamená nulový priemer pre stacionárnu časť VECM a nenulový priemer pre zložky, ktoré sú integrované,
2. $\boldsymbol{\alpha}_0$ je ohraničené, $\boldsymbol{\alpha}_1 = \mathbf{0}$ – z čoho vyplýva, že konštanty patria do kointegrujúceho priestoru a teda nenulový priemer má stacionárna časť aj integrované komponenty,
3. $\boldsymbol{\alpha}_0$ je neohraničené, $\boldsymbol{\alpha}_1 = \mathbf{0}$ – čo generuje nulový priemer pre stacionárnu časť VECM a lineárny trend pre zložky, ktoré sú integrované,

4. $\boldsymbol{\alpha}_0$ je neohraničené, $\boldsymbol{\alpha}_1$ je ohraničené – čo znamená, že obidve zložky uvažovaného modelu obsahujú lineárny trend,
5. $\boldsymbol{\alpha}_0$ je neohraničené, $\boldsymbol{\alpha}_1$ je neohraničené – čo predstavuje lineárny trend stacionárnej časti modelu a kvadratický trend časti, ktorá je integrovaná.

Pre každý uvedený prípad Johansen tabeloval kritické hodnoty, pričom upozorňuje, že posun v rozdelení môžu spôsobiť aj korigujúce umelé premenné, hoci sa často ignoruje ich vplyv.

Aj napriek zložitosti, vďaka podpore aktuálnych ekonometrických programov, nie je proces modelovania Johansenovej procedúry až tak náročný ako vysvetľovaná sprievodná teória.

Model s korekčným členom

Nech model má tvar:

$$y_t = \beta_0 + \gamma_0 x_t + \gamma_1 x_{t-1} + \beta_1 y_{t-1} + u_t,$$

kde náhodná zložka *u* má normálne rozdelenie $u_t \sim N\left(0, \sigma_u^2\right)$ a platí preň podmienka $|\beta_1| < 1$.

Po nasledovných matematických operáciách získame:

$$y_t = \beta_0 + \gamma_0 x_t + \gamma_1 x_{t-1} + \beta_1 y_{t-1} + u_t$$
$$y_t - y_{t-1} + y_{t-1} = \beta_0 + \gamma_0 x_t - \gamma_0 x_{t-1} + \gamma_0 x_{t-1} + \gamma_1 x_{t-1} + \beta_1 y_{t-1} + u_t$$
$$\Delta y_t + y_{t-1} = \beta_0 + \gamma_0 \Delta x_t + \gamma_0 x_{t-1} + \gamma_1 x_{t-1} + \beta_1 y_{t-1} + u_t$$
$$\Delta y_t = \beta_0 + \gamma_0 \Delta x_t + \left(\gamma_0 + \gamma_1\right) x_{t-1} + \left(\beta_1 - 1\right) y_{t-1} + u_t$$
$$\Delta y_t = \beta_0 + \gamma_0 \Delta x_t + \left(\beta_1 - 1\right)\left[y_{t-1} - \frac{\left(\gamma_0 + \gamma_1\right)}{1 - \beta_1} x_{t-1} \right] + u_t$$

Parameter pri x_{t-1} je dlhodobý parameter a môžeme ho ľahko pre odhad ohraničiť. Ak uvažujeme predpoklad, že z dlhodobého hľadiska sa $y_t = x_t$, tak z toho vyplýva ohraničujúca podmienka pre dlhodobý parameter v tvare $\frac{\gamma_0 + \gamma_1}{1 - \beta_1} = 1$. Po jej zavedení dostaneme:

$$\Delta y_t = \beta_0 + \gamma_0 \Delta x_t + \left(\beta_1 - 1\right)\left[y_{t-1} - x_{t-1} \right] + u_t .$$

Takýto model sa nazýva *model s korekčným členom* (Error Correction Model – ECM) a člen $\left(y_{t-1} - x_{t-1}\right)$ je *korekčný člen*, ktorý nie je ničím iným ako dlhodobým vzťahom zapísaným pomocou indexov *t*-1. Model naznačuje, že zmena premennej y_t z predchádzajúceho obdobia závisí na zmene asociovanej s pohybom vysvetľujúcej premennej x_t okolo rovnováhy a korekcie odchýlky od tejto dlhodobej rovnováhy. Parameter pri korekčnom člene je negatívny, lebo $|\beta_1| < 1$ a nazýva sa *koeficientom krátkodobého prispôsobovania*. Hodnota blízka 1 znamená rýchlu konvergenciu k rovnováhe a hodnoty blízke 0 pomalú konvergenciu.

Model s korekčným členom obsahuje v svojej špecifikácii pôvodné premenné aj ich diferencie. Preto je v prípade správnej špecifikácie považovaný za model prinášajúci presnejšie prognózy ako klasické štatistické modely vytvorené na základe metodológie Boxa a Jenkinsa alebo bežné ekonometrické modely.

Zoznam literatúry

1. DAVIDSON, R. – MACKINNON, J.: *Estimation and Inference in Econometrics*. New York: Oxford University Press 1993.
2. DICKEY, D. A. – FULLER, W. A.: Distributions of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of American Statistical Association* 74, 1979, s. 427-431.
3. ENGLE, R. F. – GRANGER, C. W. J.: Co-integration and error correction representation, estimation and testing. In.: *Econometrica* 55, 1988, s. 251-276.
4. GRANGER, C. W. J. – NEWBOLD, P.: *Forecasting Economic Time Series*. London: Academic Press 1987.
5. JOHANSEN, S.: Statistical Analysis of Co-integration Vectors, In.: *Journal of Economic Dynamics and Control* 12, 1988, s. 231-254.
6. STEWART, J.: *Econometrics*, 2nd ed., Oxford: Blackwell Publishers 1991.
7. VOGELVANG, B.: *Econometrics*: *theory and applications with EViews*. Harlow : Prentice Hall 2005.