Szigma, XX. (1987-88) 2-3. sz.

MÁTYÁS LÁSZLÓ

Szimultán panelmodellek becslése

Az idősorok és keresztmetszeti adatok együttes felhasználásának a gyakorlatban legjobban bevált módja a panelmodellek alkalmazása. A hagyományos (MÁTYÁS, 1986b) és a dinamikus (MÁTYÁS, 1987/88) modellekre kidolgozott paraméterbecslési eljárások azonban — természetesen — a szimultán modellek esetén nem kielégítőek.

Az alábbi írás fő célja bemutatni a szimultán panelmodellek paramétereinek becslésére szolgáló főbb eljárásokat. A tárgyalandó témakört három szempontból is leszűkítjük:

- egyrészt nem foglalkozunk a maximum likelihood becslésekkel, csupán a legkisebb négyzetek módszerére támaszkodó eljárásokat mutatjuk be, mivel ezek gyakorlati alkalmazása a legelterjedtebb;
- másrészt kizárólag a véletlen hatású modelleket tárgyaljuk (ahol az egyed és időspecifikus hatások a residuális változóba kerülnek beépítésre), mivel az állandó hatású szimultán modellek becslése viszonylag egyszerűen elvégezhető (MÁTYÁS, 1986a);
- s végül nem foglalkozunk az identifikáció feltételeivel, mivel ezek — szempontunkból — a hagyományos feltételektől nem különböznek.

A továbbiakban először a modellt és feltételrendszerét definiáljuk, majd a korlátozott információs esztimátorokat és végül a teljes információs esztimátorokat tekintjük át részletesebben.[1]

A modell

Tekintsük az M darab szimultán egyenletből álló modellt:[2]

$$Y\Gamma + X\beta + U = 0 \tag{1}$$

ahol

Y az endogén változók megfigyeléseinek $(NT \times M)$ méretű mátrixa,

---

[1] Az írás folyamán végig felhasználjuk a MÁTYÁS (1986b) cikkben megfogalmazottakat és támaszkodunk az ott bevezetett jelölésekre.

[2] Ahol az identitásokat kiküszöböltük.

X az exogén változók megfigyeléseinek $(NT \times K)$ méretű mátrixa,

U a reziduális változók $(NT \times M)$ méretű mátrixa,

Γ az endogén változókhoz kapcsolódó strukturális paraméterek $(M \times N)$ méretű mátrixa,

β az exogén változókhoz kapcsolódó strukturális paraméterek $(K \times M)$ méretű mátrixa,

K az exogén változók száma,

N a megfigyelt egyedek száma és

T a megfigyelt idősor hossza.

Az (1) rendszer j-edik egyenlete a következő:

$$y_j = Y_j\alpha_j + X_j\beta_j + u_j = Z_j\gamma_j + u_j\,, \tag{2}$$

ahol

y_j a j-edik strukturális egyenlet eredmény változójának megfigyeléseit tartalmazó $(NT \times 1)$ méretű vektor,

Y_j a j-edik strukturális egyenlet endogén magyarázó változóinak megfigyeléseit tartalmazó $(NT \times M_j)$ méretű mátrix,

X_j a j-edik strukturális egyenlet exogén magyarázó változóinak megfigyeléseit tartalmazó $(NT \times K)$ méretű mátrix,

$Z_j = [Y_j, X_j]$,

α_j $(M_j \times 1)$ méretű paramétervektor,

β_j $(K_j \times 1)$ méretű paramétervektor,

$\gamma'_j = [\alpha'_j, \beta'_j]$

u_j a reziduális változók $(NT \times 1)$ méretű vektora,

M_j a j-edik strukturális egyenlet endogén magyarázó változóinak száma,

K_j a j-edik egyenlet exogén magyarázó változóinak száma, továbbá

$$Y_j = YH_j \quad \text{és} \quad X_j = XL_j\,,$$

ahol H_j és L_j a megfelelő szelektáló mátrixok, melyek az (1) modellből a j-edik strukturális egyenletet előállítják.

A véletlen hatású panelmodelleknél az u_j reziduális változó felbontható az

$$u_j = (I_N \otimes L_T)\mu_j + (L_N \otimes I_T)\lambda j + v_j \tag{3}$$

formában, ahol

μ_j az egyedhatásokat képviselő $(N \times 1)$ méretű valószínűségi vektorváltozó

$\mu_j = (\mu_{1j}, \ldots, \mu_{Nj})$,

λ_j az időhatásokat képviselő $(T \times 1)$ méretű valószínűségi vektorváltozó

$\lambda_j = (\lambda_{1j}, \ldots, \lambda_{Tj})$,

v_j a tényleges residuális vektorváltozó, mérete: $(NT \times 1)$,

I_N , I_T az $(N \times N)$ illetve $(T \times T)$ méretű egységmátrixok.

Tegyük fel,hogy

$$E(\mu_j) = 0\,,\ E(\lambda_j) = 0\,,\ E(v_j) = 0\,, \quad j = 1, \ldots, M$$

$$E(\mu_j \mu_l') = \sigma^2_{\mu jl} I_N\,, \quad E(\lambda_j \lambda_l') = \sigma^2_{\lambda jl} I_T\,, \quad E(v_j v_l') = \sigma^2_{vlj} I_{NT}\,,$$

és a μ_j, λ_j, v_j vektorok páronként függetlenek minden j és l-re (j és $l = 1, \ldots, M$).

Bevezetve a (MÁTYÁS 1986b) írásban részletesen áttekintett

$$M_1 = B_n = (I_N \otimes \frac{J_T}{T}) - \frac{J_{NT}}{NT}\,, \quad M_2 = B_t = (\frac{J_N}{N} \otimes I_T) - \frac{J_{NT}}{NT}\,,$$

$$M_3 = FA = \frac{J_{NT}}{NT}\,, \quad M_4 = W^* = I_{NT} - (I_N \otimes \frac{J_N}{T}) - (\frac{J_N}{N} \otimes I_T) + FA$$

operátorokat, ahol J a csupa 1-esekből álló megfelelő méretű mátrixot jelöli, a kiinduló (1) modell residuális változóinak kovariancia mátrixa:

$$\Sigma = E[(\text{vec } u)(\text{vec } u)'] = \Sigma_\mu \otimes (I_N \otimes J_T) + \Sigma_\lambda \otimes (J_N \otimes I_T) + \Sigma_v \otimes I_{NT}\,,$$

ahol

vec a máglyázó operátort jelöli[3] és

$\Sigma_\mu = [\sigma^2_{\mu jl}]$, mérete: $(M \times M)$,
$\Sigma_\lambda = [\sigma^2_{\lambda jl}]$, mérete: $(M \times M)$,
$\Sigma_v = [\sigma^2_{vjl}]$, mérete: $(M \times M)$.

A j-edik egyenlet (3) residuális változójának konvariancia mátrixa:

$$E(u_j u_j') = \Sigma_{jj} = \sigma^2_{\mu jj}(I_N \otimes J_T) + \sigma^2_{\lambda jj}(J_N \otimes I_T) + \sigma^2_{vjj} I_{NT}\,.$$

Belátható (a MÁTYÁS, 1986b-ben látottak analógiájára, BALESTRA, & al., 1987), hogy a Σ mátrix spektrálfelbontása a következő:

$$\Sigma = \sum_{i=1}^{4} \Sigma_i \otimes M_i$$

[3] A máglyázó operátor adott vektoregyüttest „egymás alá" rendez.

$$\Sigma_1 = \Sigma_v + T\Sigma_\mu, \quad \Sigma_2 = \Sigma_v + N\Sigma_\lambda,$$

$$\Sigma_3 = \Sigma_v + T\Sigma_\mu + N\Sigma_\lambda \quad \text{és} \quad \Sigma_4 = \Sigma_v.$$

Korlátozott információs becslés

Korlátozott információs becslésnek szokás nevezni azt a becslési eljárást, mely az (1) modellt a (2) összefüggésre alapozva egyenletenként becsüli, figyelmen kívül hagyva az egyes egyenletek közötti kapcsolatból származó, Σ-ban megtestesülő információt.

Tekintsük újra a j-edik strukturális egyenletet és transzformáljuk a 2LNM[4] esztimátor előállításánál megszokottakhoz hasonlóan, csak egy kicsit általánosabban:

$$X'Fy_j = X'FZ_j\gamma_j + X'Fu_j \tag{4}$$

ahol F egy $(NT \times NT)$ méretű mátrix. Abban az esetben ha $F = I_{NT}$, akkor a (4) transzformált modellre az ALNM[5]-et alkalmazva a hagyományos 2LNM esztimátort kapjuk. A reziduális változó (3) struktúrája következtében azonban ez az esztimátor nem megfelelő. Abban az esetben, ha $F = W^* = M_4$ akkor a (4) összefüggésre az ALNM-et alkalmazva a csoporton belüli (*within, intra*) esztimátor szimultán esetre vonatkozó megfelelőjét kapjuk:

$$\hat{\gamma}_{jw} = [Z_j'W^*X(X'W^*X)^{-1}X'W^*Z_j]^{-1}\times$$
$$\times \; Z_j'W^*X(X'W^*X)^{-1}X'W^*y_j. \tag{5}$$

Nézzük meg egy kicsit közelebbről, hogy honnan jön az (5) esztimátor. Tekintsük újra a (2) modellt és transzformáljuk úgy, hogy a (3) reziduumból az idő és egyedhatásokat kiiktassuk. A W* operátorral beszorozva a modellt ez könnyen elérhető (MÁTYÁS, 1986b):

$$W^*y_j = X'W^*Z_j\gamma_j + W^*u_j,$$

ahol $E(W^*u_ju_j'W^*) = \sigma_{vjj}W^*$. Erre alkalmazva az instrumentális változókat

$$X'W^*y_j = X'W^*Z_j\gamma_j + X'W^*u_j,$$

ahol

$$E(X'W^*u_ju_j'W^*X) = \sigma^2_{vjj}X'W^*X = \sigma^2_{vjj}\Omega.$$

A szokásos módon keresünk egy olyan P nem szinguláris mátrixot, hogy

$$P^{-1}P^{-1'} = \Omega^{-1}, \quad P^{-1}\Omega P^{-1'} = I, \quad PP' = \Omega.$$

---

[4] Kétfokozatú legkisebb négyzetek módszere.

[5] Általánosított legkisebb négyzetek módszere.

Ezzel tovább transzformálva a modellt:

$$P^{-1}X'W^*y_j = P^{-1}X'W^*Z_j\gamma_j + P^{-1}X'W^*u_j\,,$$

majd alkalmazva a ALNM-et megkapjuk az (5) esztimátort.

Az (5) esztimátor tehát egy olyan speciális 2LNM esztimátor, mely két merőleges vetítés segítségével (ezt végzi a W^* operátor) kiiktatja a μ_j és λ_j egyed- és időhatásokat, így ezután az ALNM már a szokásos módon alkalmazható. Ennek az esztimátornak nagy előnye, hogy használatához nincs szükség az ismeretlen varianciák becslésére.[6] A merőleges vetítés azonban információ veszteséggel jár, így célszerűbb az egyed- és időhatások kiiktatása helyett ezeket figyelembe venni, mivel ekkor az (5) esztimátornál jobb hatásfokú esztimátort nyerhetünk. Ez természetesen megfelelő F mátrix kiválasztásával történhet. Legyen $F = \Sigma_{jj}^{-1}$, akkor a (2) összefüggésből a megfelelő általánosított 2LNM (A2LNM) esztimátor a következő:

$$\hat{\gamma}_{j,A2LNM} = \left[Z_j'\Sigma_{jj}^{-1}X(X'\Sigma_{jj}^{-1}X)^{-1}X'\Sigma_{jj}^{-1}Z_j\right]^{-1}\times \qquad (6)$$
$$\times\left[Z_j'\Sigma_{jj}^{-1}X(X'\Sigma_{jj}^{-1}X)^{-1}X'\Sigma_{jj}^{-1}y_j\right].$$

Belátható (BALESTRA & al., 1987, WHITE, 1984), hogy az $F = \Sigma_{jj}^{-1}$ minimalizálja az A2LNM esztimátor kovariancia mátrixának a nyomát (diagonális elemeinek összegét), és ilyen értelemben kielégítőnek tekinthető (vagyis az $F = \Sigma_{jj}^{-1}$ választás ilyen értelemben optimális).

Nézzük meg ennek az esztimátornak a logikáját az (5) esztimátornál látottakhoz hasonlóan. A j-edik strukturális egyenlet kovariancia mátrixát az R transzformáció segítségével diagonalizálva ($R^{-1}R^{-1'} = \Sigma_{jj}^{-1}$, $R^{-1}\Sigma_{jj}^{-1}R^{-1'} = I$, $RR' = \Sigma_{jj}$):

$$R^{-1}y_j = R^{-1}Z_j\gamma_j + R^{-1}u_j\,. \qquad (7)$$

Vezessük most be a továbblépés érdekében a *lényegében aszimptotikusan hatásos* esztimátor fogalmát. Egy $\tilde{\beta}$ aszimptotikusan torzítatlan esztimátort akkor nevezünk *lényegében aszimptotikusan hatásos*abbnak egy másik $\hat{\beta}$ aszimptotikusan torzítatlan esztimátornál, ha tetszőleges $\delta > 0$ esetén van olyan mintaelemszám $((NT)(\delta))$, hogy minden $n > (NT)(\delta)$ mintánál az

$$(1+\delta)avar\,\hat{\beta}_n - avar\,\tilde{\beta}_n$$

különbség pozitív szemi-definit, ahol $avar$ az aszimptotikus kovariancia mátrixot jelöli (WHITE, 1984). Ebből a definícióból kiindulva a (7) modell optimális instrumentális változóinak nevezzük azokat a változókat, mellyekkel a modellt transzformálva, majd erre a ALNM-et (vagy KLNM-et) alkalmazva egy *lényegében*

[6] Megjegyzendő, hogy az (1) szimultán modell (2) j-edik egyenlete kovariancia mátrixának diagonalizálása útján nyert paraméterbecslések megegyeznek az (5) esztimátor nyújtotta becslésekkel MÁTYÁS, 1986a, BALTAGI, 1981.

aszimptotikusan hatásos esztimátorhoz jutunk. Belátható (WHITE, 1984), hogy a (7) modell esetén az ilyen optimális instrumentális változókat a $R^{-1}X$ mátrix tartalmazza. Ennek megfelelően a (7) modellt ezzel transzformálva:

$$(R^{-1}X)'R^{-1}y_j = X'R^{-1}R^{-1}Z_j\gamma_j + X'R^{-1}R^{-1}u_j,$$

majd az ALNM-et alkalmazva visszakapjuk a (6) esztimátort, ami értelemszerűen *lényegében aszimptotikusan hatásos.*

Az A2LNM esztimátor használhatóságához szükség van a Σ_{jj} kovariancia mátrix becslésére. Ez az $\hat{u}_j = y_j - Z_j\hat{\gamma}_{j,w}$ becsült residuum vektor segítségével könnyen előállítható:

$$\hat{\sigma}^2_{i,jj} = \frac{1}{m_i}\hat{u}'_j M_i \hat{u}_j, \qquad i = 1,2,4$$

ahol m_i a megfelelő M_i operátorok rangja ($m_1 = N-1$, $m_2 = T-1$, $m_4 = (N-1)(T-1)$), továbbá

$$\hat{\sigma}^2_{3,jj} = \hat{\sigma}^2_{1,jj} + \hat{\sigma}^2_{2,jj} - \hat{\sigma}^2_{4,jj},$$

és ebből

$$\hat{\Sigma}_{jj} = \sum_{i=1}^{4} \hat{\sigma}^2_{i,jj} M_i .$$

Ezek után az A2LNM esztimátor már minden további nélkül használható.

Belátható (BALESTRA, 1987, felhasználva HSIAO, 1974 eredményeit), hogy abban az esetben, ha

1. a μ_j, λ_j és v_j valószínűségi vektorváltozók elemei függetlenek,

2. X nem sztochasztikus és

$$\lim_{T,N\to\infty} \frac{1}{NT} X'W^* X = A,$$

ahol A pozitív definit,[7]

3.

$$\lim_{T,N\to\infty} \frac{N}{T} = 1,$$

akkor az (5) és (6) esztimátorok aszimptotikusan (megegyezően) normális eloszlásúak, 0 várható értékkel és

$$\sigma^2_{4,jj}[(\Pi H_j L_j)' A (\Pi H_j L_j)]^{-1}$$

kovariancia mátrixszal, ahol Π a redukált forma paramétereinek mátrixa.

---

[7] Csak akkor lehet pozitív definit, ha a modellben nincs szabad konstans, ellenkező esetben ugyanis az $X'W^*X$ mátrix szinguláris.

A fentiekből következik, hogy a (6) esztimátor (asszimptotikusan) nem hatásosabb mint az (5) esztimátor, ami annak tudható be, hogy N és $T \to \infty$ esetén a residuális változó kovariancia struktúrájából származó információ elenyésző az X és Y megfigyelés mátrixban foglalt információhoz képest. Vegyük figyelembe, hogy a Σ_{jj} mátrix spektrálfelbontása (MÁTYÁS, 1986b):

$$\Sigma_{jj} = (\sigma^2_{vjj} + T\sigma^2_{\mu jj} + N\sigma^2_{\lambda jj})FA + \\ + (\sigma^2_{vjj} + T\sigma^2_{\mu jj})B_n + (\sigma^2_{vjj} + N\sigma^2_{\lambda jj})B_t + \sigma^2_{vjj}W^*,$$

és így

$$\Sigma^{-1}_{jj} = \frac{1}{\sigma^2_{vjj} + T\sigma^2_{\mu jj} + N\sigma^2_{\lambda jj}}FA + \\ + \frac{1}{\sigma^2_{vjj} + T\sigma^2_{\mu jj}}B_n + \frac{1}{\sigma^2_{vjj} + N\sigma^2_{\lambda jj}}B_t + \frac{1}{\sigma^2_{vjj}}W^*.$$

Ennek megfelelően világos, hogy ha N és $T \to \infty$ akkor az (5) esztimátor és a (6) esztimátor asszimptotikusan megegyezik. Ennek tudatában kellően nagy minta esetén (ha N és T is nagy) célszerűbb az (5) esztimátor használata, mivel ekkor nem kell a varianciák becslésével foglalkozni.[8] Véges minta esetén (ha a minta nem kellően nagy) elméletileg célszerűbb az A2LNM esztimátor használata, mivel ekkor a residuális változó kovariancia struktúrájából származó információnak lényeges szerepe lehet.

Teljes információs becslés

Az (1) modell teljes információs becslésének szokás nevezni azt az eljárást, mely a modell paramétereit úgy becsüli, hogy az egyes egyenletek közötti kapcsolatra vonatkozó (a Σ mátrixban megtestesülő) információt felhasználja. Az általánosított háromfokozatú LNM esztimátorhoz (A3LNM) a hagyományos 3LNM[9] esztimátornál megszokott módon juthatunk.

Legyen

$$Z_* = \begin{pmatrix} Z_1 & & \\ & \ddots & \\ & & Z_M \end{pmatrix}, \quad \gamma = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_j \end{pmatrix}, \quad y = \operatorname{vec} Y, \quad u = \operatorname{vec} U.$$

Ekkor az (1) modell felírható az

$$y = Z_*\gamma + u$$

formában. Legyen

---

[8] Azt, hogy mi tekinthető nagy N-nek és T-nek további *Monte-Carlo* vizsgálatokkal kell eldönteni.

[9] Háromfokozatú legkisebb négyzetek módszere.

$$D = diag < \Sigma_{11} \ldots \Sigma_{MM} >, \quad \tilde{X} = I_M \otimes X.$$

Ekkor a (6) esztimátor levezetésénél látott transzformációt elvégezve majd az ALNM-et alkalmazva kapjuk a (1) modell általánosított háromfokozatú LNM esztimátorát (A3LNM)

$$\hat{\gamma}_{\mathrm{A3LNM}} = [Z_*' D^{-1} \tilde{X}(\tilde{X}' D^{-1} \Sigma D^{-1} \tilde{X})^{-1} \tilde{X}' D^{-1} Z_*]^{-1} \times$$
$$\times [Z_*' D^{-1} \tilde{X}(\tilde{X}' D^{-1} \Sigma D^{-1} \tilde{X})^{-1} \tilde{X}' D^{-1} y].$$

Az esztimátor használhatóságához szükséges az ismeretlen varianciák becslése. Ennek egyik konzisztens útja a következő:

$$\hat{\sigma}^2_{ijl} = \frac{1}{m_i}(y_j - Z_{j\gamma_j,\mathrm{A2LNM}})' M_i (y_l - Z_{l\gamma_l,\mathrm{A2LNM}}) \qquad i = 1, 2, 4$$

$$\hat{\sigma}^2_{3jl} = \hat{\sigma}^2_{1jl} + \hat{\sigma}^2_{2jl} - \hat{\sigma}^2_{4jl}$$

$$\hat{\Sigma}_{jl} = \sum_i \hat{\sigma}^2_{ijl} M_i, \quad \hat{\Sigma} = [\hat{\Sigma}_{jl}], \quad \hat{D} = diag < \hat{\Sigma}_{11} \ldots \hat{\Sigma}_{MM} >.$$

A hagyományos technikai eszközök segítségével belátható, hogy a (6) esztimátor tárgyalásánál bevezetett 1., 2. és 3. feltételek esetén a A3LNM esztimátor aszimptotikusan normális eloszlású, 0 várható értékkel és

$$(\tilde{P}'(\Sigma_v^{-1} \otimes A)\tilde{P})^{-1}$$

kovariancia mátrixszal, ahol

$$\tilde{P} = \begin{pmatrix} \Pi H_1 L_1 & & \\ & \ddots & \\ & & \Pi H_M L_M \end{pmatrix}.$$

Abban az esetben, ha az egész rendszer éppen identifikált, bármely j-edik strukturális egyenlet paramétereinek A3LNM esztimátora megegyezik a A2LNM esztimátorral. Az (1) modell paramétereinek teljes információs maximum likelihood esztimátorának (bonyolult) levezetésével pedig kimutatható (BALESTRA, P. & al. 1987), hogy ez az esztimátor aszimptotikusan megegyezik az A3LNM esztimátorral.

Az A2LNM és A3LNM esztimátorok látszólagos bonyolultságuk ellenére számítástechnikai eszközökkel könnyen előállíthatók, így gyakorlati alkalmazásuk nem ütközik nehézségbe. Ez azért is kedvező, mert a szimultán modellek nagyon információigényesek, és ez az igény hatékonyan elégíthető ki az idősorok és keresztmetszeti adatok együttes felhasználásával, ami a szimultán panelmodellek felhasználását ösztönözheti.

(Beérkezett: 1988. május 3-án.)

Irodalom

BALESTRA, P. – VARADHARAJAN–KRISHNAKUMAR, J. (1987) Full Information Estimation of a System of Simultaneous Equation with Error Component Structure. *Econometric Theory* (Vol 3) N1.

BALTAGI, B.H. (1981) Simultaneous Equation with Error Components. *Journal of Econometrics* (Vol 17) N2.

HSIAO, C. (1974) Statistical Inference for a Model with Both Random Cross-Sectional and Time Effects. *International Economic Review* (Vol 15) N1.

MÁTYÁS L. (1986) Idősorok és keresztmetszeti adatok együttes felhasználása az ökonometriai vizsgálatokban. Kandidátusi értekezés. Budapest.

MÁTYÁS L. (1986) A panelmodellek becslése. *Szigma*, (Vol 19) N4.

MÁTYÁS L. (1987/88) Dinamikus panelmodellek becslése. *Szigma* (Vol 20) N2.

WHITE, H. (1984) *Asymptotic Theory for Econometricians*. Academic Press Inc. 1984.

Estimation of Simultaneous Panel Models

The purpose of the article is to show how the parameter estimation procedures for traditional simultaneous models can be extended to panel models. By analysing the properties of the parameter estimators thus produced the author wishes to help to the practice of model builders.