ETLA ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Kalevankatu 3 B, 00100 Helsinki 10, Finland, tel. 601322

Keskusteluaiheita
Discussion papers

Jukka Lassila

LINEAARISTEN SIMULTAANIYHTÄLÖMALLIEN PARAMETRIEN ESTIMOINTI RAJOITETUN INFORMAATION MENETELMILLÄ

No 8 22.11.1976

SISÄLLYSLUETTELO

LINEAARISTEN SIMULTAANIYHTÄLÖMALLIEN PARAMETRIEN ESTIMOINTI RAJOITETUN INFORMAATION MENETELMILLÄ

1. Johdanto

Ekonometristen mallien parametrien estimointiteorian tutkimuksesta on pääosan muodostanut ns. simultaaniyhtälöitä koskeva tutkimus. Simultaaniyhtälömalleissa useiden muuttujien arvot määräytyvät samanaikaisesti, ja jonkin yhtälön selitettävä muuttuja voi olla toisissa yhtälöissä selittäjänä. Mallin määrittelystä riippuen saattavat estimointiongelmat vaihdella huomattavasti; hankaluuksia aiheuttavat esim. epälineaariset mallit ja monimutkaiset viivejakaumat sekä suurten mallien parametrien estimointi "alamittaisista" otoksista, jolloin havaintoja on vähemmän kuin eksogeenisiä selittäviä muuttujia.

Tässä tutkimuksessa tarkastellaan varsin yksinkertaista, useissa ekonometrian oppikirjoissa esitettyä mallia. Tässä mallissa, joka on täsmennetty luvussa 2 , ei ole epälineaarisia osia eikä viivejakaumia. Viimeisessä luvussa mallia hieman vaikeutetaan: havaintoja oletetaan olevan vähemmän kuin eksogeenisiä muuttujia.

Tutkimuksen päätarkoituksena on selvittää, miten teoreettisen mallin sisältämää, erilaisista oletuksista koostuvaa informaatiota voidaan käyttää hyväksi estimaattoreita muodostettaessa, ja miten jotkut yleisesti tunnetut ja käytetyt estimaattorit tätä informaatiota käyttävät. Tutkimuksessa käsitellään ainoastaan rajoitetun informaation menetelmiä; erona täyden informaation menetelmiin rajoitetun informaation menetelmät vaativat vain tarkasteltavan yhtälön spesifioimista, ei mallin kaikkien yhtälöiden. Rajoitetun informaationkin menetelmät vaativat kuitenkin tietoa siitä, mitä (eksogeenisiä) muuttujia sisältyy koko malliin.

Kaavoissa ja todistuksissa on käytetty matriisimerkintöjä. Lukemisen yksinkertaistamiseksi useimmat todistukset on siirretty liitteisiin. Todistukset on jaoteltu luvuittain.

Tutkimuksessa on kaksivaiheinen pienimmän neliösumman estimaattori ja rajoitetun informaation maximum likelihood-estimaattori johdettu aiemmasta hieman poikkeavalla tavalla. Luvussa 2 olevaa ehdollisen parhaan harhattoman estimaattorin teoreemaa en ole aiemmin nähnyt esitetyn; teoreema on yksinkertainen Aitkenin yleistetyn Gauss-Markov-teoreeman yleistys. Luvussa 3 esitettyä identifiointirajoitusten testaamista koskevia huomioita samoin kuin luvussa 5 esitettyä estimaattoria en ole aiemmin havainnut esitetyksi. Kaikki liitteissä esitetyt todistukset ovat omiani, joten niihin on suhtauduttava varauksellisesti.

2. LINEAARINEN SIMULTAANIYHTÄLÖMALLI

2.1. Johdanto ja merkintöjä

Tarkastellaan lineaarista yhtälöryhmää, jossa on M yhtälöä. Oletetaan, että yhtälöryhmässä on M kappaletta muuttujia $y_1,\ldots,y_M$, joiden saamat arvot määräytyvät tarkasteltavassa yhtälöryhmässä ja riippuvat sekä toisistaan että yhtälöryhmässä olevien K:n muuttujan $x_1,\ldots,x_K$ saamista arvoista. Lisäksi y_i-muuttujien arvoihin vaikuttavat stokastiset virhetermit. Muuttujia y_i kutsutaan endogeenisiksi ja muuttujia x_i eksogeenisiksi muuttujiksi ja yhtälöryhmää malliksi. Jos oletetaan, että mallin muuttujista on T yhteistä havaintoa, voidaan mallin rakennemuoto esittää seuraavasti:

$$(2.1) \qquad Y^*\Gamma = X\beta + U$$

Y^* on endogeenisten muuttujien T x M-havaintomatriisi, X vastaavasti eksogeenisten muuttujien T x K-havaintomatriisi. Γ ja β ovat tuntemattomien parametrien muodostamia kerroinmatriiseja, joiden dimensiot ovat M x M ja K x M. Γ:n oletetaan olevan ei-singulaarinen. U on M:n T x 1-virhetermivektorin $u_1,\ldots,u_M$ muodostama tuntematon T x M -matriisi. Virhetermivektoreista oletetaan, että

$$(2.2) \qquad Eu_i = 0 \qquad i = 1,\ldots,M \text{ , ja}$$

$$(2.3) \qquad Eu_iu_j' = \sigma_{ij}I \qquad i,j = 1,\ldots,M \text{ .}$$

Koska virhetermit ovat satunnaismuuttujia, ovat myös muuttujat y_i satunnaismuuttujia. Mallin redusoituun muotoon päästään kertomalla yhtälö (1) puolittain oikealta puolelta Γ:n käänteismatriisilla:

$$(2.4) \qquad Y^* = X\beta\Gamma^{-1} + U\Gamma^{-1}$$

$$= X\Pi^* + V^*$$

Π* on tuntemattomien parametrien muodostama K x M-matriisi ja V* on virhetermimatriisi, joka koostuu vektoreista $v_1,\ldots,v_M$. Voidaan osoittaa[1], että

$$(2.5) \qquad Ev_i = 0 \qquad i = 1,\ldots,M \text{ , ja}$$

$$(2.6) \qquad Ev_i v_j = \omega_{ij} I \qquad i,j = 1,\ldots,M \text{ .}$$

Rakennemuodon ja redusoidun muodon välillä ovat siis yhteydet

$$(2.7) \qquad \beta\Gamma^{-1} = \Pi^*$$

$$(2.8) \qquad U\Gamma^{-1} = V^*$$

Tässä tutkimuksessa ovat kiinnostuksen kohteena ns. rajoitetun informaation menetelmät, jolloin malli (1) estimoidaan yhtälö kerrallaan. Olkoon tyypillinen tarkasteltava yhtälö

---

1) kts. esim. Goldberger, A.S.: Econometric Theory, s. 301

(2.9) $y = Y\gamma + X_1\beta_1 + u$

jossa y on varsinainen selitettävä muuttuja. Y on selittäjinä olevien endogeenisten muuttujien havaintomatriisi, X_1 on yhtälössä mukana olevien eksogeenisten muuttujien havaintomatriisi ja u on virhetermivektori. γ ja β_1 ovat tuntemattomia parametrivektoreita.

Voidaan olettaa, ilman että se vaikuttaa myöhempien tulosten yleisyyteen, että y on muuttuja y_1, että Y on muuttujien $y_2,\ldots,y_M$ havaintomatriisi, dimensioiltaan T x (M-1), ja että X_1 on muuttujien $x_1,\ldots,x_{K_1}$ T x K_1 havaintomatriisi. Tällöin γ on (M-1) x 1-vektori. u on muotoa T x 1. Yhtälössä ovat siis mukana kaikki endogeeniset muuttujat, mutta eksogeenisistä muuttujista puuttuvat muuttujat $x_{K_1+1},\ldots,x_K$, joiden muodostamaa T x $(K-K_1)$-havaintomatriisia merkitään X_2:lla. Sitä vastaava $(K-K_1)$ x 1-parametrivektori β_2 on siis 0-vektori. Tällöin ovat voimassa ositukset

(2.10) $Y^* = [y \; Y]$ ja

(2.11) $X = [X_1 \; X_2]$

Muutkin tässä tutkimuksessa käytettävät ositukset noudattelevat jompaa kumpaa tai molempia edellä olevia osituksia.

Esimerkiksi V* voidaan osittaa kaavaa (10) vastaavasti

(2.12) $V^* = [v \; V]$

Π* voidaan osittaa kaavaa (10) vastaavasti

(2.13) $\Pi^* = [\pi \; \Pi]$

kaavaa (11) vastaavasti

(2.14) $\Pi^* = \begin{bmatrix} \Pi_1^* \\ \\ \Pi_2^* \end{bmatrix}$;

tai molempia vastaavasti

(2.15) $\Pi^* = \begin{bmatrix} \pi_1 & \Pi_1 \\ \pi_2 & \Pi_2 \end{bmatrix}$

Lisäksi otetaan käyttöön merkintä

(2.16) $\gamma^* = \begin{bmatrix} 1 \\ -\gamma \end{bmatrix}$

γ* on siis M x 1-vektori. Yhtälö (9) voidaan nyt esittää muodossa

(2.17) $Y^*\gamma^* = X_1\beta_1 + u$

u:n varianssia merkitään täst'edes σ_u^2:lla.

2.2. Parametrien estimoinnissa tarvittava lisäinformaatio

2.2.1. Yleistä

Estimaattori on havaintojen funktio, jonka kustakin otoksesta saaman arvon toivotaan olevan "lähellä" perusjoukon vastaavan parametrin arvoa. Tässä tutkittavassa tapauksessa, kuten yleensäkin, havainnot, ainakin osa niistä, ovat satunnaismuuttujien saamia arvoja. Tällöin estimaattori voidaan tulkita satunnaismuuttujien funktioksi ja se on siis itsekin satunnaismuuttuja.

Estimaattorin hyvyyttä tarkastellaan usein mm. harhattomuuden ja konsistenssiuden käsitteiden avulla. Estimaattoria sanotaan harhattomaksi, jos sen odotusarvo on ko. parametrin arvo, ja konsistentiksi, jos todennäköisyys sille, että estimaattorin arvo poikkeaa ko. parametrin arvosta, lähestyy nollaa, kun havaintojen lukumäärä lähestyy ääretöntä. Estimaattorin ominaisuuksien tarkastelussa ovat tärkeitä myös estimaattorin varianssi ja asymptoottinen varianssi, joka on varianssin raja-arvo, kun havaintojen lukumäärä lähestyy ääretöntä.

Luvussa 2.1. esitettyjen tietojen avulla voitaisiin parametrit γ, β_1 ja Π^* estimoida, mutta tiedot eivät ole riittäviä estimaattorien edellämainittujen ominaisuuksien tutkimiseen. Estimaattoreiden ominaisuuksien kannalta ovat ratkaisevia tiedot tai oletukset selittävien muuttujien ja virhetermien välisistä suhteista.

Asian selventämiseksi tehdään eräitä väliaikaisia olettamuksia yhtälöstä (9). Merkitään $Z = [Y\, X_1]$ ja $\alpha = [\begin{smallmatrix}\gamma \\ \beta_1\end{smallmatrix}]$. Tällöin Z on $T \times (M-1+K_1)$ ja α on $(M-1+K_1) \times 1$-matriisi ja yhtälö (9) saa muodon

(2.18) $y = Z\alpha + u$

Oletetaan (väliaikaisesti), että

(2.19) $E(u|Z) = \mu_z$, ja

(2.20) $E(uu'|Z) = \sigma^2 W_z$,

eli virhetermien ehdollisen odotusarvon ja ehdollisen varianssin ehdolla Z oletetaan riippuvan Z:sta, eli virhetermit ja selittävät muuttujat eivät ole riippumattomia.

Tarkastellaan nyt y:n ja μ_z:n suhteen ehdollisesti lineaarisia estimaattoreita ehdolla Z[1]. Tällaiset estimaattorit ovat muotoa

(2.21) $\hat{\alpha} = C'y + D'\mu_z$

jossa C ja D ovat $T \times (M-1+K_1)$-matriiseja. Todistusliitteessä on osoitettu, että estimaattorien ehdollinen harhattomuus ehdolla Z edellyttää, että

1) Ilman ehdollistamista estimaattorit olisivat yleensä stokastisia funktioita, eivätkä lineaarisia. Vrt. Goldberger, m.t., s. 268

(2.22) $C = -D$,

ja että näistä ehdollisesti harhattomista estimaattoreista pienin ehdollinen varianssi on estimaattorilla

(2.23) $a = (Z'W_z^{-1}Z)^{-1}Z'W_z^{-1}y - (Z'W_z^{-1}Z)^{-1}Z'W_z^{-1}\mu_z$

Jälkimmäinen tulos on yleistys Aitkenin yleistetystä Gauss-Markov-teoreemasta[1]; olettamalla Z ei-stokastiseksi ja μ_z nollaksi, voidaan ehdollistamisesta luopua ja tulokseksi saadaan Aitkenin esittämä teoreema.

Estimaattoria (23) voitaisiin pitää varsin hyvänä estimaattorina, oletettavasti sen asymptoottisetkin ominaisuudet olisivat hyviä. Huomattakoon, että harhattomuus ei vaadi tietoja W_z:sta, vaan ainoastaan μ_z:sta. Yleensä μ_z:sta ei kuitenkaan ole tietoa muulloin kuin silloin, kun se voidaan olettaa nollaksi. Simultaaniyhtälömallien tapauksessa ei tällaista oletusta kuitenkaan voida tehdä, koska se on epälooginen endogeenisten muuttujien kohdalla; sen sijaan eksogeenisten muuttujien kohdalla olettamus voidaan tehdä ja yleensä tehdäänkin.

Tässä tutkimuksessa oletetaan, että muuttujat $x_1,\ldots,x_K$ ovat ei-stokastisia ja virhetermien odotusarvojen ei oleteta riippuvan niistä, eli

(2.24) $E(U|X) = E(U) = 0$

1) Vrt. Goldberger, m.t., s. 233.

ja

(2.25) $E(V^{*}|X) = E(V^{*}) = 0$.

Yhtälön (9) estimointia varten voidaan yhtälöstä (24) havaita, että

(2.26) $E(u|X_1) = 0$.

Lisäksi oletetaan, että $x_1 \equiv 1$, eli myös vakio on selittävänä muuttujana.

Seuraavaksi tarkastellaan parametrien Π^*, β_1 ja γ estimointia käytettävissä olevan informaation avulla.

2.2.2. Redusoidun muodon kertoimien estimointi

Parametrimatriisi Π^* voidaan estimoida yhtälöstä (4)

(4) $Y^{*} = X\Pi^{*} + V^{*}$

tavallisella pienimmän neliösumman menetelmällä (tästä eteenpäin PNS-menetelmällä). Yhtälö (4) sisältää M yhtälöä; estimointi tapahtuu yhtälö kerrallaan. Voidaan osoittaa, että Π^*:n estimaattori P^*:n on Π^*:n paras harhaton lineaarinen estimaattori[1],

1) kts. esim. Goldberger, m.t., s. 246-248.

ts. kaikista Π^*:n harhattomista estimaattoreista sillä on pienin varianssi.

$$(2.27) \quad P^* = (X'X)^{-1}X'Y^*$$

P^* voidaan osittaa samoin kuin Π^*:

$$(2.28) \quad P^* = [p \; P] = \begin{bmatrix} P_1^* \\ P_2^* \end{bmatrix} = \begin{bmatrix} p_1 & P_1 \\ p_2 & P_2 \end{bmatrix}$$

2.2.3. Eksogeenisten muuttujien kertoimien estimointi

Koska X on oletettu ei-stokastiseksi, voidaan μ_z kaavasta (19) merkitä seuraavasti

$$(2.29) \quad \mu_z = E(u|Z) = E(u|Y) = \mu_y$$

Se on kuitenkin edelleen tuntematon, eikä β_1:n estimaattorina voida käyttää kaavan (23) estimaattorin a β_1:tä vastaavaa osaa. Liitteessä on todistettu, että jos voitaisiin olettaa μ_y:n olevan lineaarinen Y:n suhteen, eli muotoa Yk, jossa k on tuntematon (M-1) x 1-vektori, saataisiin PNS-menetelmällä β_1:lle ehdollisesti harhaton estimaattori ehdolla Y. Koska aiemmin on oletettu, että $W_z = E(uu'|Z) = \sigma_u^2 I$, olisi saatu estimaattori ehdollisesti paras harhaton lineaarinen estimaattori ehdolla Y; lineaarinen tässä tapauksessa y:n ja μ_y:n suhteen.

β_1:n estimointi ei kuitenkaan muodosta oleellista ongelmaa. Kun tarkastellaan kaavaa (17):

(2.17) $Y^*\gamma^* = X_1\beta_1 + u$

havaitaan, että kaikki klassisen lineaarisen regressioanalyysin olettamukset ovat voimassa[1], kun tulkitaan $Y^*\gamma^*$ selitettäväksi muuttujaksi[2].

Täten β_1:n paras harhaton lineaarinen (Y^*c^*:n suhteen) estimaattori on PNS-estimaattori:

(2.30) $b_1 = (X_1'X_1)^{-1}X_1'Y^*\gamma^*$

Vaikeutena on edelleen se, että γ^*:ä ei tunneta ja täten b_1:tä ei voida kokonaan laskea. Käytännössä γ^* korvataan jollain estimaattorillaan; saatu β_1:n estimaattori on tällöin paras harhaton estimaattori ehdolla $\hat{\gamma}^*$. Tarkasteltaessa yleisimpiä simultaanimallien estimaattoreita, kuten kaksivaiheista pienimmän neliösumman menetelmää (täst'edes 2VPNS) ja rajoitetun informaation maximum likelihood-menetelmää (RIML), havaitaan, että niissä β_1:n estimaattori voidaan esittää muodossa (2.30) siten, että γ^*:n paikalla on sen estimaattori.

2.2.4. Endogeenisten muuttujien kertoimien estimointi

Sopivan estimaattorin löytäminen γ^*:lle on tarkasteltavan

1) vrt. Goldberger, m.t., s. 162.

2) Tätä tulkintaa on käyttänyt esim. Chow (1964).

yhtälön estimoinnin vaikein vaihe. Koska $E(u \mid Y)$ ei ole tunnettu, ei kaavassa (23) esitettyä estimaattoria voida käyttää. Virhetermin ehdollisesta odotusarvosta tiedetään kuitenkin sen verran, että se ei ole nolla; jos γ^* (eli varsinaisesti siis γ) estimoitaisiin PNS-menetelmällä yhdessä β_1:n kanssa, saataisiin ainakin γ:n kohdalla ehdollisesti harhaisia estimaatteja.

γ^*:n estimointia varten tarvitaan siis lisää informaatiota. Ainoa toistaiseksi käyttämätön informaatio on tieto siitä, että $\beta_2 = 0$; $Y^*\gamma^*$ ei riipu X_2:sta, eli muuttujista $x_{K_1+1},\ldots,x_K$.[1] Voidaan helposti löytää useitakin muuttujia, joista $Y^*\gamma^*$ ei riipu: kaikki sellaiset muuttujat, joista mikään muuttujista y_i, $i = 1,\ldots,M$, ei ole riippuvainen, ovat varmasti sellaisia, että jokainen endogeenisten muuttujien y_i lineaarikombinaatio on niistä riippumaton. Informaatio siitä, että β_2 on nollavektori, on kuitenkin tulkittavissa toisin. Mallin endogeenisistä muuttujista jokin tai jotkut riippuvat muuttujista $x_{K_1+1},\ldots,x_K$, mutta endogeenisten muuttujien tietty lineaarikombinaatio $Y^*\gamma^*$ ei riipu yhdestäkään muuttujista $x_{K+1},\ldots,x_K$. Tämä havaitaan ottamalla odotusarvot toisaalta yhtälöstä (4) ja toisaalta yhtälöstä (17):

$$(2.31) \qquad E(Y^*) = E(X\Pi^*) + E(V^*) = X\Pi^* = X_1\Pi_1^* + X_2\Pi_2^*$$

$$(2.32) \qquad E(Y^*\gamma^*) = [E(Y^*)]\gamma^* = E(X_1\beta_1) + E(u) = X_1\beta_1$$

1) Muunlaisesta lisäinformaatiosta kts. esim. Fisher (1959).

Tieto siitä, että $\beta_2 = 0$, on ratkaiseva γ^*:n estimoinnin kannalta. Kyseessä on itse asiassa $K - K_1$ erillistä tietoa; onko tämä määrä riittävä, riippuu siitä, montako endogeenistä muuttujaa tarkasteltavassa yhtälössä on. Tätä ns. identifiointiongelmaa tarkastellaan myöhemmin.

Yhtälöistä (31) ja (32) saadaan tunnetut yhteydet Π^*:n, γ^*:n β:n ($= \begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}$) välille:

(2.33) $\Pi_1^*\gamma^* = \beta_1$

(2.34) $\Pi_2^*\gamma^* = \beta_2 = 0$

Tarkastellaan seuraavaksi mielivaltaista endogeenisten muuttujien lineaarikombinaatiota $Y^*\alpha$, jossa α on $M \times 1$-vektori. Oletetaan muuttuja $Y^*\alpha$ selitetyksi muuttujilla $x_1,\ldots,x_K$, ja yhtälö estimoiduksi PNS-menetelmällä. Merkitään kerroinestimaattoria b_α:lla:

(2.35) $b_\alpha = (X'X)^{-1}X'Y^*\alpha$

Estimaattori voidaan osittaa kuten β:kin:

(2.36) $b_\alpha = \begin{bmatrix} b_{\alpha 1} \\ b_{\alpha 2} \end{bmatrix}$

Koska PNS-estimaattori on selitettävän muuttujan suhteen lineaarinen, havaitaan, että[1])

(2.37) $b_\alpha = P^*\alpha$,

jossa P* on aiemmin määritelty Π*:n estimaattori. Yhtälö (2.37) voidaan nyt osittaa seuraavasti:

(2.38) $P_1^*\alpha = b_{\alpha 1}$

(2.39) $P_2^*\alpha = b_{\alpha 2}$

Yhtälöiden (2.33) ja (2.34) kanssa analogiset yhtälöt olisivat voimassa, jos $b_{\alpha 2}$ olisi nollavektori. Tuntuisikin luonnolliselta pyrkiä löytämään sellainen α, joka samalla olisi γ^*:n estimaattori, että $b_{\alpha 2}$ olisi nollavektori tai "mahdollisimman lähellä" sitä. Jatkossa osoitetaankin, että tavallisimmat simultaanimalleissa käytetyt rajoitetun informaation estimaattorit, 2VPNS ja RIML, voidaan johtaa suoraan tästä tavoitteesta. Huomattakoon, että triviaaliratkaisu $\alpha = 0$ ei kelpaa, sillä merkinnän (2.16) mukaan γ^*:n ylin elementti on ykkönen.

1) Näitä PNS-estimaattoreiden lineaarikombinaatioita olen tarkastellut tutkimuksessani "Rakennuskustannusten ennustemalli", kansantaloustieteen pro gradu-työ Helsingin yliopistossa, toukokuu 1975.

3. ESTIMAATTORIEN JOHTAMINEN JA VERTAILU

3.1. Johdanto

Tarkasteltava yhtälö on siis

(3.1) $Y^*\gamma^* = X_1\beta_1 + u$ eli

(3.2) $y - Y\gamma = X_1\beta_1 + u$.

Ongelmana on γ^*:n eli varsinaisesti γ:n estimoiminen; β_1:n estimointi ei muodosta sen jälkeen ongelmaa. γ:n estimointia varten on käytettävissä $K - K_1$ kappaletta informaatiota siitä, että muuttujien $y_1,\ldots,y_M$ lineaarikombinaatio $Y^*\gamma^*$ ei ole riippuvainen muuttujista $x_{K_1+1},\ldots,x_K$, vaikka jokaisesta x-muuttujasta on ainakin yksi y-muuttuja riippuvainen.

Yhtälö (3.1) on identifioituva jos redusoidun muodon kerroinmatriisista Π^* voidaan johtaa yhtälön (3.1) parametrit yksikäsitteisesti. Välttämätön ja riittävä ehto tälle on, että matriisin Π_2^* asteluku on $M-1$. Tämä on ns. rank-ehto[1).

Koska Π^* on tuntematon, on ns. järjestysehto (order condition[2)) käytännössä tärkeämpi; sen mukaan välttämätön ehto (3.1):n identifioituvuudelle on se, että mallissa mukana olevien, mutta yhtälöstä (3.1) puuttuvien eksogeenisten muuttujien lukumäärä on vähintään yhtä suuri kuin yhtälössä (3.1) selittäjinä olevien

---

1) kts. esim. Goldberger, mt. s.316

2) Goldberger, mt. s.316

endogeenisten muuttujien lukumäärä, eli että $K - K_1 \geq M - 1$.

Kun $K - K_1 = M - 1$, yhtälöä (3.1) sanotaan täsmälleen identifioituvaksi. On huomattava, että tässä tilanteessa informaatiota on "lukumääräisesti" yhtä paljon kuin jos virhetermin ehdollinen odotusarvo kaavassa (1.19) tunnettaisiin; ehdollisessa odotusarvossa on kysymys $M - 1 + K_1$:n muuttujan vaikutuksesta virhetermiin, nyt taas on käytettävissä $M - 1$ riippumattomuustietoa sen lisäksi, että aikaisemmin oletettiin, että K_1 muuttujaa $x_1, \ldots, x_{K_1}$ eivät vaikuta virhetermiin. Myöhemmin osoitetaan, että informaation määrä on tällöin juuri riittävä: γ^*:n estimaattori c^* voidaan johtaa redusoidun muodon kerroinmatriisin estimaattorista P^*[1].

Kun $K - K_1 > M - 1$, yhtälöä sanotaan yli-identifioituvaksi. Tällöin informaatiota on tavallaan liikaa. Tämä informaatio on siinä mielessä ristiriitaista, että γ:n estimaattoria c ei suoraan voida johtaa P^*:stä. Tarvitaan siis jotain kriteeriä tai periaatetta tämän liian suuren informaatiomäärän käyttämiseksi, tavallaan painottamiseksi, mielekkäällä tavalla; myöhemmin osoitetaan, miten eri kriteereillä päädytään eri estimaattoreihin.

Ellei erikseen toisin sanota, tarkastellaan tilannetta, jossa $K < T$, eli mallissa on eksogeenisiä muuttujia vähemmän kuin havaintoja. Luvun 3.5. lopussa vertaillaan estimaattoreita

1) Tällöin on oletettu, että P^* on "täyttä astetta", eli $\text{rank}(P^*) = \min(K,M)$.

tilanteesssa K = T, ja luvussa 5 tarkastellaan tilannetta K > T.

3.2. Täsmälleen identifioituva yhtälö

Kuten luvussa 2.2.4. todettiin, luonnolliselta tuntuva γ^*:n estimaattori α on sellainen, että yhtälössä

$$(1.39) \quad P_2^* \alpha = b_{\alpha 2}$$

olisi $b_{\alpha 2}$ nollavektori, tai "lähellä" sitä. Merkitään täst'edes α:n sijasta c^*:llä γ^*:n estimaattoria. Yhtälö (1.39) saa muodon

$$(3.3) \quad P_2^* c^* = b_{c*2}$$

c^* voidaan osittaa seuraavasti:

$$(3.4) \quad c^* = \begin{bmatrix} 1 \\ -c \end{bmatrix}$$

Tällöin (3.3) voidaan osittaa:

$$(3.5) \quad p_2 = P_2 c = b_{c*2}$$

Havaitaan, että

$$(3.6) \quad b_{c*2} = 0 <=> P_2 c = p_2$$

Täsmälleen identifioituvassa yhtälössä P_2^* on $(M - 1) \times M$-matriisi, joten P_2 on $(M - 1) \times (M - 1)$-matriisi. Se on yleensä ei-singulaarinen, joten b_{c*2} saadaan nollaksi asettamalla

(3.7) $$c = P_2^{-1} p_2$$

Käyttämällä ositetun matriisin käänteismatriisin kaavaa[1] c voidaan esittää muodossa

(3.8) $$c = [D^{-1}X_2'M_1Y]^{-1}D^{-1}X_2'M_1y$$

$$= (X_2'M_1Y)^{-1}DD^{-1}X_2'M_1y$$

$$= (X_2'M_1Y)^{-1}X_2'M_1y$$

jossa

(3.9) $$M_1 = I - X_1(X_1'X_1)^{-1}X_1'$$

(3.10) $$D = X_2'M_1X_2$$

Täsmälleen identifioituvassa tapauksessa, kun muuttujaa $Y^*c^* = y - Yc$ selitetään muuttujilla $x_1, \ldots, x_K$ ja yhtälö estimoidaan PNS-menetelmällä, saa muuttujien $x_{K_1+1}, \ldots, x_K$ kerroinvektorin β_2 estimaattori b_{c*2} vapaasti määräytyessään arvon nolla, kuten yhtälöistä (3.6) ja (3.7) havaitaan.

---

1) kts. Goldberger, mt., s.174

β_1:n estimaattori on

$$(3.11) \quad b_1 = (X_1'X_1)^{-1}X_1'(y-Yc)$$

$$= (X_1'X_1)^{-1}X_1'[I - Y(X_2'M_1Y)^{-1}X_2'M_1]y$$

3.3. Yli-identifioituva yhtälö

3.3.1. Lisäkriteerejä

Yli-identifioituvassa tapauksessa ei c:tä voida ratkaista kuten yhtälössä (3.7), sillä P_2 on singulaarinen, koska se ei ole edes neliömatriisi. Yleensä ei ole olemassa sellaista vektoria c, että b_{c*2} olisi nollavektori yhtälössä (3.5). Tällöin täytyy pyrkiä siihen, että löytyisi c, jolle b_{c*2} olisi "mahdollisimman lähellä" nollavektoria. Seuraavassa on esitetty kolme ehkä helpoimmin mieleen tulevaa kriteeriä:

1° Minimoidaan β_2:n kerroinestimaattien neliösumma, eli valitaan c siten, että $b_{c*2}'b_{c*2}$ minimoituu.
Saadulla γ:n estimaattorilla on eräitä epätyydyttäviä ominaisuuksia, kuten myöhemmin huomataan.

2° Valitaan c siten, että selitettäessä muuttujaa y - Yc ensin muuttujilla $x_1,...,x_{K_1}$ ja lisättäessä sen jälkeen selittäjiksi $x_{K_1+1},...,x_K$, jälkimmäisten muuttujien mukaantulon aiheuttama selitetyn varianssin absoluuttinen lisäys on mahdollisimman pieni, kun yhtälö estimoidaan PNS-menetelmällä. Toisin sanoen, minimoidaan X_2:n mukaantulon aiheuttama selitetyn varianssin absoluuttinen lisäys. Osoittautuu, että tällöin päädytään kaksivaiheiseen PNS-estimaattoriin (2VPNS).

3^o Minimoidaan X_2:n mukaantulon aiheuttaman selitetyn varianssin lisäyksen suhde jäännösvarianssiin. Saatu estimaattori minimoi myös todennäköisyyden sille, että selitettäessä y - Yc:tä sekä muuttujille $x_1,\ldots,x_{K_1}$ että muuttujille $x_{K_1+1},\ldots,x_K$ käyttäen PNS-menetelmää, β_2:n estimaattori poikkeaa nollasta, kun virhetermin oletetaan noudattavan normaalijakaumaa. Osoittautuu, että tällöin päädytään rajoitetun informaation maximum likelihood-estimaattoriin (RIML).

3.3.2. Kerroinestimaattien neliösumman minimointi

Neliösumman minimointi tapahtuu normaalisti: sen ensimmäinen derivaatta c:n suhteen asetetaan nollaksi. Neliösumma on kaavan (3.5.) perusteella

$$(3.12)\quad b'_{c*2}b_{c*2} = (p_2-P_2c)'(p_2-P_2c) = p_2'p_2 - 2p_2'P_2c + c'P_2'P_2c$$

Derivoidaan neliösumma c:n suhteen:

$$(3.13)\quad \frac{\partial b'_{c*2}b_{c*2}}{\partial c} = -2P_2'p_2 + 2P_2'P_2c\ ,$$

joka asetettuna nollaksi antaa "normaaliyhtälöt"

(3.14) $P_2'P_2c = P_2'p_2$

josta saadaan ratkaisu

(3.15) $c = (P_2'P_2)^{-1}P_2'p_2$

Saatu ratkaisu on minimi, sillä

(3.16) $\frac{\partial^2 b'_{c*2} b_{c*2}}{\partial c^2} = 2P_2'P_2$

joka on positiivisesti definiitti[1].

Saatu estimaattori voitaisiin luultavasti osoittaa konsistentiksi[2]. Estimaattori on kuitenkin sikäli epätyydyttävä, että se riippuu muuttujien $x_{K_1+1},\ldots,x_K$ mittayksiköistä.

P_2:n elementithän ovat näitä muuttujia vastaavia redusoidun muodon kerroinestimaattoreita ja niiden saamat arvot riippuvat muuttujien mittayksiköistä. Esimerkiksi muuttujan havaintojen kertominen vakiolla aiheuttaa sen, että vastaava kerroinestimaatti tulee jaetuksi vakiolla. Tällöin kerroinestimaattien neliösumma on suuresti riippuvainen muuttujien $x_{K_1+1},\ldots,x_K$ mittayksiköistä. Mahdollinen korjauskeino olisi kyseiden eksogeenisten muuttujien normeeraaminen jakamalla ne keskihajonnoillaan.

1) vrt. Goldberger, s. 36, kaava 7.7. P_2 on $(K - K_1) \times (M - 1)$-matriisi, ja rank $(P_2) = M - 1$.

2) tekemällä eräitä lisäolettamuksia. vrt. Goldberger, mt., s. 301

Kaava (3.15) saa auki kirjoitettuna muodon

(3.17) $c = (Y'M_1X_2D^{-1}D^{-1}X_2'M_1Y)^{-1}Y'M_1X_2D^{-1}D^{-1}X_2'M_1Y$

β_1:n estimaattori saadaan jälleen kaavasta (1.30).

3.3.3. Absoluuttisen lisäselityksen minimointi

Otetaan aluksi käyttöön eräitä lyhentäviä merkintöjä. Selitetään muuttujaa Y^*c^* PNS-menetelmää käyttäen muuttujilla $x_1,\ldots,x_{K_1}$ ja $x_{K_1+1},\ldots,x_K$. Merkitään kokonaisneliösummaa K:lla, eli

(3.18) $K = c^{*}'Y'Yc^{*} = (y-Yc)'(y-Yc)$

Merkitään selitettyä neliösummaa S(X):llä

(3.19) $S(X) = c^{*}'Y'X(X'X)^{-1}X'Yc^{*}$

Merkitään jäänösneliösummaa J(X):llä

(3.20) $J(X) = c^{*}'Y'[I - X(X'X)^{-1}X']Yc^{*}$

Kun Y^*c^*:ä selitetään vain muuttujilla $x_1,\ldots,x_{K_1}$, merkitään selitettyä neliösummaa $S(X_1)$:llä jne. Koska $x_1 \equiv 1$, tiedetään, että

$$(3.21) \quad K = S(X) + J(X)$$

$$= S(X_1) + J(X_1)$$

Merkitään lisäselitystä $S(X) - S(X_1)$ $L(X_2)$:lla;

$$(3.22) \quad L(X_2) = S(X) - S(X_1)$$

$L(X_2)$ on auki kirjoitettuna

$$(3.23) \quad L(X_2) = c^{*\prime}Y^{*\prime}M_1X_2D^{-1}X_2'M_1Y^*c^*$$

Minimoidaan nyt (3.23) c^*:n suhteen. Menettelytapa on sama kuin luvussa 3.3.2. Tulokseksi saadaan

$$(3.24) \quad c = (Y'HY)^{-1}Y'Hy$$

jossa

$$(3.25) \quad H = M_1X_2D^{-1}X_2'M_1$$

Liitteessä on johdettu kaava (3.24) ja todistettu, että tulos on minimi. β_1:n estimaattori on tällöin

$$(3.26) \quad b_1 = (X_1'X_1)^{-1}X_1'[I - Y(Y'HY)^{-1}Y'H]y$$

1) vrt. Goldberger, s. 176-177.

Liitteessä on osoitettu, että b_1 ja c ovat itse asiassa β_1:n ja γ:n kaksivaiheisen pienimmän neliösumman estimaattorit. Täten niiden ominaisuudet ovat hyvin tunnettuja. Ominaisuuksia käsitellään luvussa 3.5.

3.3.4. Suhteellisen lisäselityksen minimointi

Edellisessä luvussa minimoitiin absoluuttinen lisäselitys $L(X_2)$, nyt minimoidaan lisäselityksen suhde jäännösneliösummaan, eli lauseke $\frac{L(X_2)}{J(X)}$. Minimointi tapahtuu c:n suhteen. Minimoitava lauseke on auki kirjoitettuna seuraava:

$$(3.27)\qquad \frac{L(X_2)}{J(X)} = \frac{c^{*\prime}Y^{*\prime}HY^{*}c^{*}}{c^{*\prime}Y^{*\prime}MY^{*}c^{*}},$$

jossa

$$(3.28)\qquad M = I - X(X'X)^{-1}X'$$

Voidaan osoittaa[1], että (3.27) saa minimin, kun ratkaistaan yhtälö

$$(3.29)\qquad (H - \lambda M)c^{*} = 0$$

ja valitaan pienintä ominaisarvoa λ vastaava ominaisvektori c^{*}_{λ}. Tämän jälkeen käytetään ns. normalisointisääntöä, eli

1) kts. esim. Wonnacott-Wonnacott, Econometrics, s. 376-378.

tietoa siitä, että c*:n ylimmän alkion on oltava ykkönen. Ilman tätä sääntöä c* saataisiin vain vakiokerrointa vaille. Huomattakoon, että lausekkeen (3.27) minimiarvo on juuri pienin λ[1].

β_1:n estimaattori on jälleen muotoa

$$(3.30) \quad b_1 = (X_1'X_1)^{-1}X_1'Y^*c^*$$

Seuraavaksi osoitetaan, että saadut estimaattorit minimoivat todennäköisyyden sille, että β_2:n estimaattori b_{c*2} poikkeaisi merkitsevästi nollasta, kun Y*c*:tä ajatellaan selitetyn kaikilla x_i-muuttujilla i = 1,...,K.

Oletetaan, että virhetermit noudattavat normaalijakaumaa, ja että $E\,u = 0$ ja $E\,uu' = \sigma_u^2 I$. Tällöin sen hypoteesin vallitessa, että $\beta_2 = 0$, noudattaa testisuure[2]

$$(3.31) \quad \frac{T-K}{K-K_1} \cdot \frac{Q_2}{J(X)}$$

F-jakaumaa vapausastein $\frac{K-K_1}{T-K}$, jossa

$$(3.32) \quad Q_2 = b_2'Db_2 = c^{*\prime}Y^{*\prime}HY^*c^*$$

Suuret testisuureen arvot antavat aihetta epäillä hypoteesin $\beta_2 = 0$ oikeellisuutta. On selvää, että estimaattori, joka

1) vrt. Wonnacott-Wonnacott, mt., s.378

2) vrt. Goldberger, mt., s.177

minimoi lausekkeen (3.27), minimoi myös lausekkeen (3.31).

Liitteessä on todistettu, että edellä esitetyt β_1:n ja γ:n estimaattorit ovat itse asiassa rajoitetun informaation maximum likelihood-estimaattoreita. Näiden estimaattoreiden ominaisuudet ovat tunnetut. Niitä käsitellään luvussa 3.5.

3.3.5. Identifiointirajoitusten testaaminen

Tähän asti on oletettu, että muuttuja $Y^*\gamma^*$ on riippumaton muuttujista $x_{K_1+1},\dots,x_K$. Tätä oletusta voidaan myös testata. Ekonometrian oppikirjoissa[1] on yleisesti esitetty käytettäväksi testisuuretta $(\hat{\lambda}-1)$, joka itse asiassa on

$$(3.33) \qquad \hat{\lambda}-1 = \lambda \,,$$

jossa λ on pienin ominaisarvo yhtälöstä (3.29) kuten liitteessä on todistettu. $(\hat{\lambda}-1)$:n voidaan osoittaa noudattavan asymptoottisesti jakaumaa $\chi^2_{K-K_1-M+1}$, jos identifiointirajoitus on oikea.

Kuten luvussa 3.3.4. osoitettiin, noudattaa testisuure

$$(3.31) \qquad \frac{T-K}{K-K_1}\,\lambda = \frac{T-K}{K-K_1}\,(\hat{\lambda}-1)$$

F-jakaumaa vapausastein $\frac{K-K_1}{T-K}$, jos identifiointirajoitus on

1) kts. esim. Goldberger, m.t., s. 346.

oikea. Jos testisuure joutuu hylkäämisalueelle, tiedetään, että jokaiselle lineaarikombinaatiolle $Y^*\alpha$, jossa α:n ylin alkio on ykkönen, testisuure joutuu hylkäämisalueelle: λ saatiin minimoimalla lauseke, joka samalla minimoi ko. testisuureen. Jos minimikin on liian suuri, ovat varmasti kaikki muutkin testisuureen saamat arvot liian suuria. Identifiointirajoitus $\beta_2 = 0$ voidaan silloin hylätä ja vaihtoehtoinen hypoteesi $\beta_2 \neq 0$ astuu voimaan.

Basmann[1] on johtanut vastaavan testisuureen 2VPNS-estimaattoria käytettäessä. Hänen mukaansa testisuure

$$(3.32) \qquad \frac{T-K}{K-K_1-M+1} \cdot \frac{c^{*\prime}Y^{*\prime}HY^*c^*}{c^{*\prime}Y^{*\prime}MY^*c^*}$$

jossa c^* on γ^*:n 2VPNS-estimaattori, noudattaa approksimatiivisesti F-jakaumaa vapausasteilla $K-K_1-M+1$ ja $T-K$. Testisuure voidaan esittää neliösummalyhenteiden avulla muodossa

$$(3.33) \qquad \frac{T-K}{K-K_1-M+1} \cdot \frac{L(X_2)_{2VPNS}}{J(X)_{2VPNS}}$$

jossa alaviite 2VPNS viittaa siihen, että c^* on 2VPNS-estimaattori.

Theilin[2] mukaan Basmannin menetelmä on Monte Carlo-kokeissa osoittautunut χ^2-menetelmää paremmaksi.

---

1) Basmann (1960).

2) Theil: Principles of Econometrics, s. 508.

2VPNS-menetelmän yhteydessä voitaisiin myös ajatella ehdollista F-testiä: ehdolla, että γ^*:n 2VPNS-estimaattori c^* on oikea, testataan, poikkeaako muuttujia $x_{K_1+1},\ldots,x_K$ vastaavan kerroinvektorin β_2 estimaattori b_{c^*2} merkitsevästi nollasta, kun muuttujaa Y^*c^* selitetaan muuttujilla $x_1,\ldots,x_K$, ja yhtälö estimoidaan PNS-menetelmällä. Testisuure

$$(3.34) \qquad \frac{T-K}{K-K_1} \cdot \frac{L(X_2)_{2VPNS}}{J(X)_{2VPNS}}$$

noudattaa F-jakaumaa parametrein $K-K_1$ ja $T-K$, kun identifiointirajoitukset ovat voimassa.

Kaikissa eo. testimenettelyissä on oletettu virhetermin noudattavan normaalijakaumaa.

3.4. k-luokan estimaattorit

Ennen kuin siirrytään estimaattoreiden vertailuun, laajennetaan tarkasteltavien estimaattoreiden ryhmää yksiparametrisella estimaattoriperheellä, k-luokan estimaattoreilla. Tämän Theilin kehittämän estimaatoriryhmän normaaliyhtälöt voidaan esittää muodossa[1)]

$$(3.35) \qquad (Y'Y-k\hat{V}'\hat{V})c_k + Y'X_1b_k = Y'y-k\hat{V}'y$$

$$(3.36) \qquad X_1'Yc_k + X_1'X_1b_k = X_1'y$$

1) vrt. Goldberger, m.t., s. 336.

jossa

(3.37) $\hat{V}'y = Y'[I - X(X'X)^{-1}X']y = Y'My$

(3.38) $\hat{V}'\hat{V} = Y'MY$

(3.39) $b_k = (X_1'X_1)^{-1}X_1'(y-Yc_k)$

(3.35) voidaan esittää seuraavasti:

(3.40) $Y'[I - kI + kX(X'X)^{-1}X']Yc_k + Y'X_1b_k = Y'[I - kI + kX(X'X)^{-1}X']y$

Sijoittamalla tähän (3.39) saadaan

(3.41) $Y'[I - kI + kX(X'X)^{-1}X' - X_1(X_1'X_1)^{-1}X_1']Yc_k$

$= Y'[I - kI + kX(X'X)^{-1}X' - X_1(X_1'X_1)^{-1}X_1']y$

Tästä saadaan edelleen

(3.42) $Y'\{(1-k)[I-X_1(X_1'X_1)^{-1}X_1'] + k[X(X'X)^{-1}X'-X_1(X_1'X_1)^{-1}X_1']\}Yc_k$

$= Y'\{(1-k)[I-X_1(X_1'X_1)^{-1}X_1'] + k[X(X'X)^{-1}X'-X_1(X_1'X_1)^{-1}X_1']\}y$

Toisaalta

(3.43) $X'(X'X)^{-1}X' - X_1(X_1'X_1)^{-1}X_1' = M_1X_2D^{-1}X_2'M_1$ 1)

$= H$ 2)

1) vrt. Goldberger, m.t., s. 174 ja 176.

2) vrt. kaava (3.25).

joten (3.42) saa muodon

$$(3.44) \quad Y'[(1-k)M_1 + kH]Yc_k = Y'[(1-k)M_1 + kH]y$$

jossa $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$ eli c_k on

$$(3.45) \quad c_k = \{Y'[(1-k)M_1+kH]Y\}^{-1}Y'[(1-k)M_1+kH]y$$

(3.44) on auki kirjoitettuna

$$(3.46) \quad (1-k)Y'M_1Yc_k + kY'HYc_k$$

$$= (1-k)Y'M_1y + kY'Hy$$

Toisaalta kaavaan (3.44) päästään minimoimalla c_k:n suhteen lauseke

$$(3.47) \quad (1-k)(y-Yc_k)'M_1(y-Yc_k) + k(y-Yc_k)'H(y-Yc_k)$$

$$\Longleftrightarrow (1-k)J(X_1) + kL(X_2)$$

kuten liitteessä on todistettu. Täten k-luokan estimaattori saadaan minimoimalla (3.46) c_k:n suhteen ja laskemalla b_k kaavasta (3.39).

3.5. Estimaattorien vertailu

3.5.1. Estimaattorien toimenkuvaukset

Seuraavassa pyritään vertailemaan toisiinsa erilaisia γ :n estimaattoreita c . Kaikki β_1:n estimaattorithan osoittautuivat olevan tavallisia PNS-estimaattoreita, kun muuttujaa $y - Yc$ selitetään muuttujilla $x_1,\ldots,x_{K_1}$. Tästä seuraa, että onpa c saatu millä tahansa menetelmistä 2VPNS, RIML, k-luokan estimaattorimenetelmä tai PNS, jossa siis c ja b_1 on estimoitu yht'aikaa, on voimassa neliösummahajoitelma:

(3.48) $K = S(X_1) + J(X_1)$ (kts. luku 3.3.3)

Samoin on voimassa, että jos $(y - Yc)$:tä olisi PNS-menetelmällä selitetty kaikilla x_i-muuttujilla $i = 1,\ldots,K$, niin

(3.49) $K = S(X) + J(X)$

Toisaalta kaavan (3.22) mukaan

(3.22) $L(X_2) = S(X) - S(X_1)$

jolloin kaavoista (3.47), (3.48) ja (3.22) päästään kaavaan

(3.50) $J(X_1) - L(X_2) = J(X)$

Kaikki neliösummat kaavassa (3.49) ovat c*:n (eli c:n) funktioita. Estimaattorin toimenkuvaukseksi kutsutaan tästä eteenpäin jonkin neliösumman minimoimista c:n suhteen. PNS-estimaattorin toimenkuvaus on luonnollisesti

$$(3.51) \quad \text{PNS:} \ \min_{c} J(X_1)$$

Luvun 3.3.3. mukaan 2VPNS:n toimenkuvaus on [1)]

$$(3.52) \quad \text{2VPNS:} \ \min_{c} L(X_2)$$

Luvun 3.3.4. mukaan RIML:n toimenkuvaus on

$$(3.53) \quad \text{RIML:} \ \min_{c} \frac{L(X_2)}{J(X)}$$

Luvun 3.5 mukaan k-luokan estimaattorin toimenkuvaus on

$$(3.54) \quad \text{k-luokka:} \ \min_{c} (1-k)J(X_1) + kL(X_2)$$

eli k-luokan estimaattorin toimenkuvaus on painotettu keskiarvo PNS:n ja 2VPNS:n toimenkuvauksista[2)], painoina (1 - k) ja k. Tästä ei seuraa, että k-luokan estimaattori olisi painotettu keskiarvo PNS- ja 2VPNS-estimaattoreista.

---

1) Teoksessa Wonnacott-Wonnacott, m.t., mainitaan s. 381, että 2VPNS minimoi jäännösvarianssien erotuksen $J(X_1) - J(X)$; tämähän on yhtä kuin $L(X_2)$.

2) Teoksessa Wonnacott-Wonnacott, m.t., osoitetaan s. 379, että k-luokan estimaattori voidaan tulkita instrumenttimuuttujamenetelmäksi, jossa instrumenttimuuttuja on painotettu keskiarvo PNS:n ja 2VPNS:n instrumenttimuuttujista.

Yhdistettynä estimaattorien toimenkuvaukset ovat siis seuraavat:

(3.55)

Estimaattori	minimoi c:n suhteen	tavoitefunktion
PNS		$J(X_1)$
2VPNS		$L(X_2)$
RIML		$\dfrac{L(X_2)}{J(X)}$
k-luokan estimaattori		$(1-k)J(X_1)+k\,L(X_2)$

Kaikki tavoitefunktiot ovat kvadraattisia c:n funktioita. Kolme ensimmäistä tavoitefunktiota voi saada vain ei-negatiivisia arvoja. Huomattakoon, että RIML:iä ei voida laskea, kun $J(X) = 0$.

Yllä olevat tavoitefunktiot voidaan luonnollisesti esittää useissa eri muodoissa käyttämällä hyväksi yhtälöitä (3.21) ja (3.22).

3.5.2. Estimaattorien vertailu

Estimaattorien vertailu tapahtuu pääasiassa taulukon (3.55) avulla. Aluksi vertaillaan estimaattorien keskinäisiä suhteita eri tilanteissa ja lopuksi yleisesti. Eri tilanteissa havaintojen, poissa- ja mukana olevien eksogeenisten muuttujien sekä endogeenisten muuttujien lukumäärät vaihtelevat. Muistettakoon, että tutkimuksessa on oletettu mallien kaikkien endogeenisten muuttujien olevan mukana tarkastellussa yhtälössä. Oletus ei sinänsä ole tärkeä eikä vähennä todistusten yleisyyttä. Jos kaikki endogeeniset muuttujat eivät ole mukana, on luvun $M-1$ paikalle seuraavassa tarkastelussa asetettava yhtälössä selittäjinä olevien endogeenisten muuttujien lukumäärä.

1^{o} Tarkastellaan täsmälleen identifioituvaa yhtälöä, eli $K-K_1 = M-1$, ja oletetaan, että $K<T$. Luvun 3.2. perusteella tiedetään, että muuttujia $x_{K_1+1},\ldots,x_K$ vastaavat regressiokertoimet voidaan saada nolliksi. Täten luvussa 3.2. esitetty ratkaisu minimoi sekä 2VPNS:n että RIML:n tavoitefunktion, eli nämä estimaattorit yhtyvät luvussa 3.2. esitettyyn estimaattoriin, jota kutsutaan epäsuoraksi PNS-estimaattoriksi.

Kuten tunnettua, 2VPNS-estimaattori on k-luokan estimaattori k:n avulla 1, ja RIML on k-luokan estimaattori k:n arvolla $\lambda'+1$. Tässä tapauksessa λ' saa arvon 0, joten 2VPNS, RIML ja k-luokan estimaattori k = 1 yhtyvät.

2^o Tarkastellaan yli-identifioituvaa yhtälöä, eli $K - K_1 > M - 1$ jossa edelleen $K < T$.

Estimoidun yhtälön selityskykyä voidaan vertailla jäännösneliösumman $J(X_1)$ avulla. Tämä saa pienimmän arvonsa, kun käytetään PNS-estimaattoria. Seuraavaksi todistetaan, että 2VPNS:n jäännösneliösumma on pienempi kuin RIML:n jäännösneliösumma. Taulukosta (3.54) voidaan päätellä, että

$$(3.56) \quad \frac{L(X_2)_{RIML}}{J(X)_{RIML}} < \frac{L(X_2)_{2VPNS}}{J(X)_{2VPNS}}$$

ja että

$$(3.57) \quad L(X_2)_{RIML} > L(X_2)_{2VPNS} \, ,$$

josta seuraa

$$(3.58) \quad J(X)_{RIML} > J(X)_{2VPNS} \, ,$$

joka voidaan (3.49):n mukaan esittää muodossa

$$(3.59) \quad J(X_1)_{RIML} - L(X_2)_{RIML} > J(X_1)_{2VPNS} - L(X_2)_{2VPNS}$$

Tästä voidaan päätellä, että

$$(3.60) \quad J(X_1)_{RIML} - J(X_1)_{2VPNS} > L(X_2)_{RIML} - L(X_2)_{2VPNS} > 0 \, .$$

Yhtälöiden selityskyvystä voidaan siis todeta, että

(3.61) $J(X_1)_{PNS} < J(X_1)_{2VPNS} < J(X_1)_{RIML}$

3^o Seuraavaksi tarkastellaan tilannetta, jossa $K - K_1 > M - 1$ ja $K = T$. Tällöin $J(X) = 0$, eli $J(X_1) = L(X_2)$ kaikilla estimaattoreilla. Taulukosta (3.54) havaitaan, että

(3.62) $J(X) = 0 \Rightarrow$ RIML ei ole olemassa

(3.63) PNS = 2VPNS = k-luokan estimaattori kaikilla äärellisillä k:n arvoilla. RIML on tässä tapauksessa k-luokan estimaattori k:n arvolla ääretön.

4^o Lopuksi tarkastellaan tilannetta, jossa $K - K_1 = M - 1$ ja $K = T$. Kohdista 1^o ja 3^o voidaan päätellä, että

(3.64) RIML ei ole olemassa

(3.65) PNS = 2VPNS = k-luokan estimaattori kaikilla $|k| < \infty$

(3.66) $J(X_1)_{PNS} = J(X_1)_{2VPNS} = J(X_1)_k = 0$.

Eri estimaattorien toimenkuvaukset siis painottavat eri lailla yhtälön selityskykyä [eli termiä $J(X_1)$] ja rajoitusta $\Pi_2{*}\gamma{*} = 0$ [eli termiä $L(X_2)$]. k-luokan estimaattori painottaa selityskykyä painoa (1-k) ja identifiointirajoitusta painolla k.

Täten

- PNS painottaa vain selityskykyä,
- 2VPNS painottaa vain identifiointirajoitusta,
- RIML painottaa yli-identifioituvassa tapauksessa, samoin kuin ne k-luokan estimaattorit, joilla $k > 1$, selityskykyä negatiivisesti ja identifiointirajoitusta yli 1:n suuruisella painolla.

Huomattakoon, että RIML on k-luokan estimaattori, jota vastaava k-luokan estimaattorin tavoitefunktion arvo on nolla. Jos merkitään λ_i:llä RIML:n laskemisessa saatua i:nnettä ominaisarvoa, voidaan päätellä, että k-luokan estimaattorin tavoitefunktio saa arvon nolla kaikilla k_i:

(3.67) $$k_i = \lambda_i + 1$$

k-luokan estimaattorin tavoitefunktio voi saada myös negatiivisia arvoja. Tämä saattaa olla yhteydessä siihen, että tietyillä k:n arvoilla estimaattori "räjähtää".

2VPNS vaikuttaa monessa suhteessa miellyttävämmältä kuin RIML. Se on laskennallisesti yksinkertaisempi, selityskyky on yli-identifioituvassa tapauksessa parempi, se on olemassa tilanteissa 3^o ja 4^o eli tilanteissa, jossa $K = T$. Tällöin RIML ei ole olemassa. Jos RIML tulkitaan k-luokan estimaattoriksi, sen ominaisuus painottaa selityskykyä negatiivisesti vaikuttaa epäilyttävältä.

RIML vaikuttaa 2VPNS-estimaattoria käyttökelpoisemmalta ainoastaan identifiointirajoitusten testaamisessa.

Luvussa 3.3.2. esitetty estimaattori ei ole ollut vertailussa mukana. Tähän estimaattoriin palataan lyhyesti seuraavassa luvussa.

4. YHTEENVETO

Tutkimuksen tarkoituksena oli selvittää, miten teoreettisen mallin sisältämää, erilaisista oletuksista koostuvaa informaatiota voidaan käyttää hyväksi estimaattoreita muodostettaessa, ja miten eräät yleisesti tunnetut ja käytetyt estimaattorit tätä informaatiota käyttävät.

Aluksi tutkittiin virhetermien ja selittävien muuttujien keskinäistä riippuvuutta. Todettiin, että oleellista ei ole se, että virhetermit olisivat riippumattomia selittävistä muuttujista, vaan se, että riippuvuus on tunnettu. Luvussa 2 esitettiin kaava hyvälle estimaattorille, kun virhetermin ehdollinen odotusarvo ja ehdollinen varianssi tunnetaan.

Kun virhetermit oletetaan riippumattomiksi eksogeenisistä muuttujista, ei redusoidun muodon parametrien estimoinnissa ole vaikeuksia. Myös β_1:n estimointi on selväpiirteistä: tulkitaan Y^*c^*, jossa c^* on γ^*:n estimaattori, selitettäväksi muuttujaksi, ja estimoidaan β_1 tavallisella PNS-menetelmällä. Todettiin, että k-luokan estimaattorit, siis myös 2VPNS ja RIML, voidaan esittää tässä muodossa. Identifiointirajoitusten testaaminen voitiin myös palauttaa tavallisen klassisen regressioyhtälön kertoimia koskevien hypoteesien testaamiseen, kun Y^*c^* tulkittiin selitettäväksi muuttujaksi.

γ:n estimointia varten ei riittävää informaatiota virhetermien ja muuttujien y_i riippuvuuksista ole. Tämän sijasta käytettiin hyväksi olettamusta

(4.1) $\quad \Pi_2^{*}{}'\gamma^{*} = 0$

joka voidaan myös esittää muodossa

(4.2) $\quad \pi_2 = \Pi_2\gamma$

γ:n estimointi osoittautui olevan sellaisen vektorin c etsimistä, että yhtälö

(4.3) $\quad p_2 = P_2c$

olisi "mahdollisimman hyvin" voimassa, kun p_2 ja P_2 ovat π_2:n ja Π_2:n PNS-estimaattoreita. Antamalla erilaisia operationaalisia tulkintoja sanonnalle "mahdollisimman hyvin", päädytään erilaisiin estimaattoreihin.

γ:n estimointi on itse asiassa tavallista regressioanalyysiä, jossa redusoidun muodon estimaatit ovat havaintoja. Yhtälön (4.3) voidaan tulkita poikkeavan yhtälöstä (4.2) siinä mielessä, että muuttujien havainnoissa on mittausvirheitä. Tiedetään, että[1]

1) Vrt. Goldberger, mt., s. 174.

(4.4) $$p_2 = \pi_2 + D^{-1}X_2'M_1v$$

(4.5) $$P_2 = \Pi_2' + D^{-1}X_2'M_1V$$

josta seuraa

(4.6) $$p_2 = P_2\gamma + D^{-1}X'_2M_1(v-V\gamma) = P_2\gamma + \varepsilon$$

(4.7) $$E\varepsilon = 0$$

(4.8) $$E\varepsilon\varepsilon' = \sigma_\varepsilon^2 D^{-1} \qquad (D = X_2'M_1X_2)$$

Tällöin γ:n estimaattori c saadaan Aitkenin yleistetyllä pienimmän neliösumman menetelmällä[1]:

(4.9) $$c = (P_2'DP_2)^{-1}P_2'Dp_2$$

Sijoittamalla (4.9):ään p_2:n ja P_2:n lausekkeet ja sieventämällä saadaan

(4.10) $$c = [Y'M_1X_2D^{-1}X_2'M_1Y]^{-1}Y'M_1X_2D^{-1}X_2'M_1y$$

$$= (Y'HY)^{-1}Y'Hy$$

---

1) Vrt. Goldberger, mt., s. 232.

Havaitaan, että c on γ:n 2VPNS-estimaattori[1]. Se on konsistentti, vaikka virhetermit eivät olekaan riippumattomia selittävistä muuttujista (joiden havaintomatriisi on P_2).

Luvussa 3.3.2 johdettiin γ:lle estimaattori

$$(4.11) \quad c = (P_2'P_2)^{-1}P_2'p_2$$

joka on muuten samanlainen kuin kaavan (4.9) estimaattori, mutta se ei huomioi virhetermien kovarianssimatriisia. Jos virhetermit olisivat riippumattomia selittävistä muuttujista, jotka olisivat ei-stokastisia, olisi Aitkenin yleistetyn Gauss-Markov-teoreeman mukaan estimaattorilla (4.11) suurempi varianssi kuin estimaattorilla (4.9). Vaikka nämä olettamukset eivät ole voimassa, tuntuu (4.9) paremmalta estimaattorilta kuin (4.11).

Huomattakoon, että yhtälössä (4.6) on havaintoja $K-K_1$ kappaletta ja selittäviä muuttujia M-1 kappaletta. Vapausasteiden määrä $K-K_1-M+1$ mittaa samalla alkuperäisen tarkasteltavan yhtälön yli-identifioituvuuden määrää. Tästä havaitaan myös, että estimointi ei ole mahdollista ali-identifioituvassa tapauksessa.

---

1) Vrt. luku 3.3.3.

Intuitiivisen mielikuvan eri estimointimenetelmien toiminnasta eri tilanteissa saa tarkastelemalla estimaattoreiden toimenkuvauksia. Näiden toimenkuvausten avulla voidaan redusoidun muodon estimaatteja jossain määrin käyttää hyväksi myös rakennemuodon yhtälöitä spesifioitaessa. Ilmeiseltä virhespesifioinnilta esim. vaikuttaa tilanne, jossa rakennemuodon yhtälöstä on jätetty pois sellainen eksogeeninen muuttuja x_j, josta muuttuja y on redusoidun muodon estimaattien perusteella riippuvainen, mutta mikään mukaan otetuista endogeenisistä selittävistä muuttujista ei ole riippuvainen x_j:stä. Tällainen "poikkeava havainto" painottuu ilmeisesti merkittävästi kaikissa γ:n estimointimenetelmissä. Helpoimmin asia on ymmärrettävissä tarkastelemalla estimaattoria (4.11).

5. PARAMETRIEN ESTIMOINTI ALAMITTAISISTA OTOKSISTA

Aiemmissa luvuissa on oletettu, että mallissa on eksogeenisiä muuttujia vähemmän kuin havaintoja, eli $K < T$. Luvussa 3.5 tarkasteltiin estimaattoreita myös tilanteessa $K = T$. Tällöin matriisi X on ei-singulaarinen ja Π^*:n estimaattori P^* on

$$P^* = X^{-1}Y^* \tag{5.1}$$

Kun $K > T$, ei redusoidun muodon kerroinmatriisia voida estimoida, eikä yhtään luvussa 3 esitellyistä γ:n estimaattoreista myöskään voida laskea. Tällaisessa tilanteessa otosta voidaan kutsua alamittaiseksi (undersized).

Useita alamittaisten otosten tapaukseen sopivia estimaattoreita on jo kehitetty[1]; monet näistä ovat iteratiivisia. Seuraavassa johdetaan yksinkertainen γ:n estimaattori, joka sopii tilanteeseen $K > T$.

Oletetaan, että moniulotteinen stationaarinen stokastinen prosessi generoi muuttujien $x_1,\dots,x_K$ arvot (ts. hylätään olettamus muuttujien ei-stokastisuudesta). Virhetermien oletetaan olevan riippumattomia muuttujista x_i.

---

1) Kts. Dutta (1975).

Muuttujan $Y^*\gamma^*$ oletetaan edelleen olevan riippumaton muuttujista $x_{K_1+1},\ldots,x_K$. Riippumattomuudesta seuraa, että

(5.2) $\quad \mathrm{cov}(Y^*\gamma^*, x_i) = 0$ kaikilla $i = K_1+1,\ldots,K$

Tällöin myös

(5.3) $\quad E(X_2 - EX_2)'(Y^*\gamma^* - EY^*\gamma^*) = 0$

eli

(5.4) $\quad E(X_2 - EX_2)'(Y^*\gamma^* - X_1\Pi_1^*\gamma^* - X_2\Pi_2^*\gamma^*) = 0$

eli

(5.5) $\quad E(X_2 - EX_2)'(Y^* - X_1\Pi_1^*\gamma)^* = 0$

sillä edelleen on oletettu, että

(5.6) $\quad \Pi_2^*\gamma^* = 0$

EX_2 ja Π_1^* ovat tuntemattomia, mutta ne voidaan korvata estimaattoreillaan. EX_2:n estimaattorina käytetään otoskeskiarvoista muodostettua matriisia $\bar{X}_2$, jonka tyypillinen alkio on

(5.7) $\quad \{\bar{X}_2\}_{ij} = \frac{1}{T} \cdot x'_{K_1+j} \cdot a$,

jossa siis a on Tx1-vektori, jonka alkiot ovat ykkösiä.
Merkitään

(5.8) $Z_2 = X_2 - \bar{X}_2$

Π_1^*:n estimaattorissa käytetään $\tilde{P}_1^*$:tä, joka on[1)]

(5.9) $\tilde{P}_1^* = (X_1'X_1)^{-1}X_1'Y^*$

Tällöin $\tilde{P}_1^*$ voidaan osittaa samoin kuin Π_1^*:

(5.10) $\tilde{P}_1^* = [\tilde{p}_1 \tilde{P}_1]$

$\tilde{P}_1^*$ on yleensä harhainen estimaattori.

Yhtälön (5.5) vasemman puolen paikalle tulee nyt vastaava "otossuure"

(5.11) $Z_2'(Y^* - X_1\tilde{P}_1^*)\gamma^* = d$

joka $(K-K_1)$x1-vektori. Se ei ole nollavektori, mutta sen voisi olettaa olevan "lähellä" nollavektoria.
Vektori d voidaan osittaa seuraavasti:

(5.12) $d = Z_2'(y-X_1\tilde{p}_1) - Z_2'(Y-X_1\tilde{P}_1)\gamma$,

---

1) Oletetaan, että $K_1 < T$.

jossa on käytetty hyväksi Y^*:n, γ^*:n ja $\tilde{P}_1^*$:n osituksia.

(5.12) voidaan esittää muodossa

$$(5.13) \quad Z_2'(y-X_1\tilde{p}_1) = Z_2'(Y-X_1\tilde{P}_1)\gamma + d$$

Yhtälö (5.13) on eräänlainen regressioyhtälö.
Merkitään

$$(5.14) \quad s = Z_2'(y-X_1\tilde{p}_1)$$

$$(5.15) \quad R = Z_2'(Y-X_1\tilde{P}_1)$$

jossa s on $(K-K_1)x1$-vektori ja R on $(K-K_1)x(M-1)$ - matriisi. Tällöin (5.13) saa muodon

$$(5.16) \quad s = R\gamma + d$$

Virhetermistä voidaan päätellä, että

(5.17) Ed on tuntematon ja

(5.18) E(d-Ed)(d-Ed)' on tuntematon.

γ:lle saadaan nyt estimaattori PNS-menetelmällä:

$$(5.19) \quad \hat{\gamma} = (R'R)^{-1}R's$$

Huolimatta siitä, että $\hat{\gamma}$ on harhainen eikä sillä liene mitään minimivarianssiin vivahtavia ominaisuuksia, se on ilmeisesti konsistentti. Olen yrittänyt osoittaa konsistenssiuden liitteessä. Konsistenssin mielekkyys tässä tapauksessa tuntuu kuitenkin kyseenalaiselta, koska siinä havaintojen määrän lähestyessä ääretöntä on myös $K < T$.

Monimutkaisista merkinnöistä huolimatta estimaattorin $\hat{\gamma}$ taustalla oleva ajatus on yksinkertainen: valitaan sellainen $\hat{\gamma}$, että otoskovarianssit muuttujan $(y-Y\hat{\gamma})$ ja muuttujien $x_{K+1},\ldots,x_K$ välillä ovat "mahdollisimman lähellä" nollaa.[1] Kriteeriksi on valittu se, että otoskovarianssien neliösumma minimoituu. Muuttujan $(y-Y\hat{\gamma})$ otoskovarianssi muuttujan x_j kanssa on hajotettu eri muuttujien y_i otoskovarianssseiksi muuttujan x_j kanssa.

Muuttujien y_i otoskovarianssia x_j:n kanssa ei kuitenkaan ole laskettu y_i:n otoskeskiarvon avulla, vaan sen sijaan on käytetty termiä

(5.20) $$(X_1'X_1)^{-1}X_1'y_i$$

Regressioyhtälössä (5.16) muuttujan s havainto j on ristitulo

(5.12) $$\{s\}_j = (x_j-\bar{x}_j)'[y_1-(X_1'X_1)^{-1}X_1'y_1]$$

ja matriisin R tyypillinen alkio on

(5.22) $\{R\}_{ji} = (x_j-\bar{x}_j)'[y_{i+1}-(X_1'X_1)^{-1}X_1'y_{i+1}]$

Tästä havaitaan, että alkiot ja samalla myös estimaattori $\hat{\gamma}$ riippuu muuttujien x_j, $j=K_1+1,\ldots,K$ mittayksiköistä. Tämä vika voidaan poistaa jakamalla muuttujien x_j arvot ko. muuttujan otoskeskihajonnalla.

Huomattakoon, että eo. estimaattori voidaan laskea kaikissa tapauksissa, joissa $K - K_1 > M - 1$ ja $K_1 < T$. Todella suurissa malleissa on vapausasteita $K - K_1 - M + 1$ luultavasti hyvin runsaasti. Käytännössä tämä lienee usein näennäistä, koska osa havainnoista on luultavasti lähellä nollavektoria, ts. mikään alkuperäisen yhtälön endogeenisistä muuttujista ei korreloi joidenkin muuttujien x_j kanssa.

Kun γ on estimoitu, voidaan β_1:n estimaattorina jälleen käyttää estimaattoria

(5.23) $b_1 = (X_1'X_1)^{-1}X_1'(y - Y\hat{\gamma})$

TODISTUSLIITE

Luku 2.2.1.

Tarkastellaan yhtälöä

(1) $y = Z\alpha + u$,jossa

(2) $E(u|Z) = \mu_z$

(3) $E[(u-Eu)(u-Eu)'|Z] = \sigma_u^2 W_z$, jossa W_z on positiivisesti definiitti.

Tarkastellaan α:n estimaattoreita $\hat{\alpha}$, jotka ovat lineaarisia y:n ja μ_z:n suhteen, eli

(4) $\hat{\alpha} = C'y + D'\mu_z$

Väite: a) jotta $\hat{\alpha}$ olisi harhaton ehdolla Z, täytyy olla

(5) $D' = -C'$ ja

(6) $C'Z = I$

b) ehdollisesti harhattomista estimaattoreista on pienin ehdollinen varianssi estimaattorilla

(7) $a = \tilde{C}'(y-\mu_z) = (Z'W_z^{-1}Z)^{-1}Z'W_z^{-1}(y-\mu_z)$

Todistus:

(8) $E(\hat{\alpha}|Z) = E[C'(Z\alpha + u)|Z] + ED'\mu_z$

$= C'Z\alpha + C'E(u|Z) + D'\mu_z$

$= C'Z\alpha + C'\mu_z + D'\mu_z$

Jotta $E(\hat{\alpha}|Z) = \alpha$, täytyy väitteen a)-kohdan olla voimassa.

Ehdollisesti harhattomille estimaattoreille on voimassa

(9) $[(\hat{\alpha} - \alpha)|Z] = C'u - C'\mu_z = C'(u - \mu_z)$

(10) $E[(\hat{\alpha} - \alpha)(\hat{\alpha} - \alpha)'|Z] = C'\sigma_u^2 W_z C = \sigma_u^2 C'W_z C$

Voidaan todeta, että estimaattori $\tilde{C}'(y - \mu_z)$ kaavassa (7) on ehdollisesti harhaton ja sen ehdollinen varianssi on

$$(11) \quad (\Sigma_{aa}|Z) = \sigma_u^2 \tilde{C}'W_z\tilde{C}$$
$$= \sigma_u^2 (Z'W_z^{-1}Z)^{-1}Z'W_z^{-1}W_zW_z^{-1}Z(Z'W_z^{-1}Z)^{-1}$$
$$= \sigma_u^2 (Z'W_z^{-1}Z)^{-1}$$

Kaikki ehdollisesti harhattomat y:n ja μ_z:n suhteen lineaariset estimaattorit voidaan esittää muodossa

$$(12) \quad \hat{a} = (\tilde{C}' + F')(y - \mu_z)$$

jossa ehdollisesta harhattomuudesta edelleen seuraa

$$(13) \quad F'Z = 0$$

Tällöin $\hat{a}$:n ehdollinen varianssi on

$$(14) \quad (\Sigma_{\hat{a}\hat{a}}|Z) = \sigma_u^2 |(\tilde{C}' + F')W_z(\tilde{C} + F)|$$
$$= \sigma_u^2 (\tilde{C}'W_z\tilde{C} + F'W_z\tilde{C} + \tilde{C}'W_zF + F'W_zF)$$
$$= (\Sigma_{aa}|Z) + \sigma_u^2 F'W_zF$$

sillä

$$(15) \quad \tilde{C}'W_zF = (Z'W_z^{-1}Z)^{-1}Z'W_z^{-1}W_zF$$
$$= (Z'W_z^{-1}Z)^{-1}Z'F$$
$$= (Z'W_z^{-1}Z)^{-1}0$$
$$= 0$$

ja samoin myös $F'W_z\tilde{C} = 0$.

Koska $F'W_zF$ on ei-negatiivisesti definiitti, on b)-kohta todistettu.

Todistus on täysin analoginen Goldbergerin esittämän Aitkenin yleistetyn Gauss-Markov-teoreeman todistuksen kanssa (kts. Goldberger, mt., s. 232 - 233). Aitkenin teoreema saadaan erikoistapauksena tässä esitetystä teoreemasta.

Luku 2.2.3.

Väite: Jos edellisessä todistuksessa olisi kaavassa (2) ollut

(1) $E(u|Z) = E(u|Y) = Yk$,

saataisiin PNS-menetelmällä α:n X_1:tä vastaavalle osalle β_1 ehdollisesti harhaton estimaattori

Todistus: Merkitään $m = \begin{bmatrix} k \\ 0 \end{bmatrix}$, jossa m on $(M-1+K_1)x1$-vektori.

Tällöin

(2) $Yk = Zm$

Edellisen todistuksen kaavasta (8) voidaan päätellä, että estimaattorin

(3) $a = (Z'Z)^{-1}Z'y$

ehdollinen harha

(4) $E(a|Z) = (Z'Z)^{-1}Z'Zm = m$,

jonka β_1:tä vastaava osa on nollavektori.

Luku 3.3.3.

Väite: $c = (Y'HY)^{-1}Y'Hy$ minimoi lausekkeen

(1) $L(X_2) = (y-Yc)'H(y-Yc)$

Todistus:

$$(2) \quad \frac{\partial L(X_2)}{\partial c} = 2Y'HYc - 2Y'Hy$$
$$= 0 \iff Y'HYc = Y'Hy$$
$$\iff c = (Y'HY)^{-1}Y'Hy$$

Saatu tulos on minimi, kun

(3) $\frac{\partial^2 L(X_2)}{\partial c^2} = 2Y'HY$ on positiivisesti definiitti, eli kun $c'y'HYc > 0$ kaikilla $c \neq 0$. Tämä taas pitää paikkansa, kun selitettäväksi muuttujaksi tässä tapauksessa tulkittavaa Yc:tä vastaava X_2:n mukanaan tuoma lisäselitys on >0: näin voidaan olettaa yleisesti olevan.

Väite: Estimaattori $c = (Y'HY)^{-1}Y'Hy$ on γ:n ja estimaattori $b_1 = (X_1'X_1')^{-1}X_1'(y-Yc)$ β_1:n 2VPNS-estimaattori.

Todistus:

Lähdetään liikkeelle Goldbergerin esittämästä 2VPNS-estimaattorin kaavasta[1]:

(1) $$\begin{pmatrix} Y'X(X'X)^{-1}X'Y & Y'X_1 \\ X_1'Y & X_1'X_1 \end{pmatrix}\begin{pmatrix} \hat{\gamma} \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} Y'X(X'X)^{-1}X'y \\ X_1'y \end{pmatrix}$$

Kirjoitetaan se auki:

(2) $Y'X(X'X)^{-1}X'Y\hat{\gamma} + Y'X_1\beta_1 = Y'X(X'X)^{-1}X'y$

(3) $X_1'Y\hat{\gamma} + X_1'X_1\beta_1 = X_1'y$

Ratkaistaan (3) β_1:n suhteen:

(4) $\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'(y-Y\hat{\gamma})$

Sijoitetaan (4) kaavaan (2):

(5) $Y'X(X'X)^{-1}X'Y\hat{\gamma} + Y'X_1(X_1'X_1)^{-1}X_1(y-Y\hat{\gamma}) = Y'X(X'X)^{-1}X'y$

eli

(6) $[Y'X(X'X)^{-1}X'Y - Y'X_1(X_1'X_1)^{-1}X_1'Y]\hat{\gamma} =$

$[Y'X(X'X)^{-1}X' - Y'X_1(X_1'X_1)^{-1}X_1']y$

Kaava (6) taas on[2]

---

1) Goldberger, mt., s. 332 kaava (6.17)

2) vrt. Goldberger, mt., s. 177

(7) $Y'M_1X_2D^{-1}X_2'M_1Y\hat{\gamma} = Y'M_1X_2D^{-1}X_2'M_1Y$

eli

(8) $Y'HY\hat{\gamma} = Y'Hy$

josta saadaan

(9) $\hat{\gamma} = (Y'HY)^{-1}Y'Hy = c$

Täten myös $\hat{\beta}_1$ on b_1 kuten kaavasta (4) havaitaan.

Luku 3.3.4.

Väite: Estimaattori, joka minimoi suhteen $\frac{L(X_2)}{J(X)}$, on rajoitetun informaation maximum likelihood -estimaattori RIML.

Todistus: Teoksessa Wonnacott - Wonnacott: Econometrics , on osoitettu s. 378 , että RIML minimoi suhteen

(1) $\frac{J(X_1)}{J(X)}$

Toisaalta

(2) $J(X_1) = J(X) + L(X_2)$,

joten (1) on

(3) $J(X_1)/J(X) = [J(X) + L(X_2)]/J(X) = 1 + L(X_2)/J(X)$

joten estimaattori, joka minimoi suhteen $L(X_2)/J(X)$ minimoi myös suhteen $J(X_1)/J(X)$, eli on RIML.

Luku 3.3.5.

Väite: $\hat{\lambda} - 1 = \lambda$ (kaava 3.33)

Todistus: Teoksessa Wonnacott - Wonnacott, Econometrics, on osoitettu s. 378, että $\hat{\lambda}$ on suhteen $J(X_1)/J(X)$ minimiarvo. Vastaavasti λ on suhteen $L(X_2)/J(X)$ minimiarvo. Edellisessä todistuksessa osoitettiin, että aina on

(1) $J(X_1)/J(X) = 1 + L(X_2)/J(X)$

Edellisessä todistuksessa osoitettiin myös, että RIML minimoi molemmat suhteet. Tällöin minimipisteet (c) ovat samoja, ja

(2) $\hat{\lambda} = 1 + \lambda$ eli $\hat{\lambda} - 1 = \lambda$

Luku 3.4.

Väite: k-luokan estimaattori minimoi lausekkeen $(1-k)J(X_1) + kL(X_2)$.

Todistus:

(1) $J(X_1) = (y-Yc)'M_1(y-Yc)$ ja

(2) $L(X_2) = (y-Yc)'H(y-Yc)$

Tällöin

$$(3) \quad \frac{\partial[(1-k)J(X_1) + kL(X_2)]}{\partial c} = 2(1-k)Y'M_1Yc - 2(1-k)Y'M_1y + 2kY'HYc - 2kY'Hy = 0 \iff$$

$$(1-k)Y'M_1Yc + kY'HYc = (1-k)Y'M_1y + kY'Hy \iff$$

$$Y'[(1-k)M_1 + kH]Yc = Y'[(1-k)M_1 + kH]y \iff$$

kaava 3.44 sivulla 31.

Kyseessä on minimi, jos lausekkeen (3) toinen derivaatta on positiivisesti definiitti. Toinen derivaatta on $Y'M_1Y + Y'HY$. Jälkimmäinen termi on jo aiemmin osoitettu positiivisesti definiitiksi. Tällöin koko lauseke on positiivisesti definiitti, jos $Y'M_1Y$ on p.d. Tämä taas pitää paikkansa, kun mikään Y:n lineaarikombinaatio Yc, $c \neq 0$, ei PNS-menetelmällä selity täysin muuttujilla $x_1,\ldots,x_{K_1}$: näin voidaan olettaa olevan.

Luku 5.

Väite: $\hat{\gamma} = (R'R)^{-1}R's$ on γ:n konsistentti estimaattori.

Todistus: Todistusta varten oletetaan, että

(1) plim $T^{-1}R'R$ on olemassa ja ei-singulaarinen

(2) plim $T^{-1}X'u$ = plim $T^{-1}X'v$ = plim $T^{-1}X'V = 0$

Todistuksen ajatuksena on saattaa $s = Z_2'(y - X_1\tilde{p}_1)$ muotoon, jossa yksi termi on $R\gamma$, osa termeistä päättyy $(\pi_2 - \Pi_2\gamma)$:aan ja loput termit sisältävät jonkin tuloista $X'u$, $X'v$ tai $X'V$ tai vastaavan eksogeenisten muuttujien ja virhetermien ristitulon. Tällöin voidaan päätellä, että plim $(R'R)^{-1}R's = \gamma$.

Koska $T \to \infty$, voidaan tarkastella tilannetta $T > K$. Tarkastellaan s:n jälkimmäistä osaa $y - X_1\tilde{p}_1$. Todetaan, että

(3) $y - X_1\tilde{p}_1 = Y\gamma + X_1\beta_1 + u - X_1\tilde{p}_1$

Sijoitetaan yhtälöön (3) hajoitelma

(4) $\beta_1 = \Pi_1^*\gamma^* = \pi_1 - \Pi_1\gamma$

jolloin saadaan

(5) $y - X_1\tilde{p}_1 = Y\gamma + X_1\pi_1 - X_1\Pi_1\gamma + u - X_1\tilde{p}_1$

$\tilde{p}_1$ voidaan esittää muodossa (6): (vrt. Goldberger, s.195)

(6) $\tilde{p}_1 = p_1 + (X_1'X_1)^{-1}X_1'X_2p_2 = p_1 + Bp_2$

Yhtälö (5) saadaan nyt muotoon

(7) $y - X_1\tilde{p}_1 = (Y - X_1\Pi_1)\gamma + X_1(\pi_1 - p_1) - X_1Bp_2$

Toisaalta tiedetään, että (vrt. Goldberger, s.164 ja 174)

(8) $\pi_1 - p_1 = (X_1'X_1)^{-1}X_1'(X_2D^{-1}X_2'M_1 - I)v$

$= Fv$

Samoin voidaan päätellä, että

(9) $\Pi_1 = P_1 + FV$

Tällöin saadaan yhtälö (7) muotoon

(10) $y - X_1\tilde{p}_1 = (Y - X_1P_1)\gamma - X_1FV\gamma + X_1Fv - X_1Bp_2$

Analogisesti kaavan (6) kanssa on voimassa

(11) $P_1 = \tilde{P}_1 - BP_2$

Tällöin voidaan (10) esittää muodossa (12)

(12) $y - X_1\tilde{p}_1 = (Y - X_1\tilde{P}_1)\gamma - X_1B(p_2 - P_2\gamma) + X_1F(v - V\gamma)$

Toisaalta (vrt. Goldberger, s. 175)

(13) $p_2 = \pi_2 + D^{-1}X_2'M_1v = \pi_2 + Av$ ja

(14) $P_2 = \Pi_2 + AV$

Tällöin (12) saa muodon

(15) $y - X_1\tilde{p}_1 = (Y - X_1\tilde{P}_1)\gamma - X_1B(\pi_2 - \Pi_2\gamma) + X_1(F-A)(v - V\gamma)$

Oikean puolen keskimmäinen termi on identifiointirajoituksen mukaan nolla. Havaitaan, että

(16) $$\begin{aligned} s = Z_2'(y - X_1\tilde{p}_1) &= Z_2'(Y - X_1\tilde{P}_1)\gamma + Z_2'X_1(F-A)(v - V\gamma) \\ &= R\gamma + Z_2'X_1(F-A)(v-V\gamma) \end{aligned}$$

jolloin

(17) $$\begin{aligned} \text{plim}\ \hat{\gamma} &= \text{plim}\ (T^{-1}R'R)^{-1}T^{-1}R's \\ &= (R'R)^{-1}R'R\gamma + \text{plim}(T^{-1}R'R)^{-1}\text{plim}T^{-1}R'Z_2'X_1(F-A)(v-V\gamma) \end{aligned}$$

Havaitaan, että yhtälön (17) oikean puolen jälkimmäinen termi sisältää eksogeenisten muuttujien ristituloja, sillä F päättyy matriisiin X_1' ja A päättyy myös matriisiin X_1'. Täten yhtälö (2) huomioon ottaen plim $\hat{\gamma} = (R'R)^{-1}R'R\gamma = \gamma$.

LÄHDELUETTELO

TEOKSET

Goldberger, A.S. : Econometric Theory, John Wiley & Sons, New York 1964

Theil, H. : Principles of Econometrics, North-Holland, Amsterdam 1971

Wonnacott, R.J. - Wonnacott, T.H. : Econometrics, John Wiley & Sons, New York 1970

ARTIKKELIT

Chow, G.C. : A Comparison of Alternative Estimators for Simultaneous Equations, Econometrica, Vol. 32 October 1964

Basmann, R.L. : On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics, Journal of the American Statistical Association, Vol. 55, December 1960

Fisher, F.M. : Generalization of the Rank and Order Conditions for Identifiability, Econometrica, Vol. 27, July 1959

Dutta, M. : The Iterative Instrumental Variables (IIV) Method of Estimation and Econometric Models Undersized and Nonlinear in Variables Only - A Note Econometric Society European Meeting, Helsinki 1976