

Moniulotteiset aikasarjat

Pentti Saikkonen

Kevät 2010

Sisältö

1.	Johdanto	1
1.1.	Taustaa	1
1.2.	Stokastinen prosessi	2
2.	Stationaariset prosessit	4
2.1.	Määritelmiä	4
2.2.	Lineaarinen prosessi	5
2.3.	VAR(p)–prosessi	7
2.4.	VARMA(p,q)–prosessi	16
3.	Stationaaristen prosessien ensimmäisten ja toisten momenttien estimointi	17
4.	Stationaarisen VAR(p)–mallin teoriaa	21
4.1.	Parametrien rajoittamaton estimointi	21
4.2.	Estimointi lineaarisin rajoittein	26
4.3.	Hypoteesien testaus	28
4.4.	Mallinvalinta ja valitun mallin sopivuuden tutkiminen	30
4.5.	Grangerin kausaalisuus	36
4.6.	Impulssivasteanalyysi	40
5.	Integroituneet ja yhteisintegroituneet prosessit	44
5.1.	Integroituneet prosessit	44
5.2.	Yhteisintegroituneet prosessit	47
5.3.	Yhteisintegroitunut VAR(p)–prosessi ja virheenkorjausmalli	49
6.	I(1)–prosessien asymptotiikkaa	55
7.	Yhteisintegroituneiden VAR(p)–mallien teoriaa	60
7.1.	Uskottavuusfunktio ja SU–estimointi	60
7.2.	Yhteisintegroituvuusasteen testaus	67
7.3.	SU–estimaattorin asymptoottinen jakauma	72
7.4.	Hypoteesien testaus	77
7.5.	Empiirinen esimerkki	81
	Liite A: Matriisilaskentaa	85
	Liite B: Asymptoottisia tuloksia	93

1 Johdanto

1.1 Taustaa

Kurssilla tarkastellaan moniulotteisten aikasarjojen kuvaamiseen ja ennustamiseen käytettäviä tilastollisia malleja ja niiden parametrien estimointia ja testausta. *Moniulotteinen tai vektoriarvoinen aikasarja* on havaintoaineisto, jossa on vektoriarvoisia havaintoja peräkkäisinä ajankohtina eli

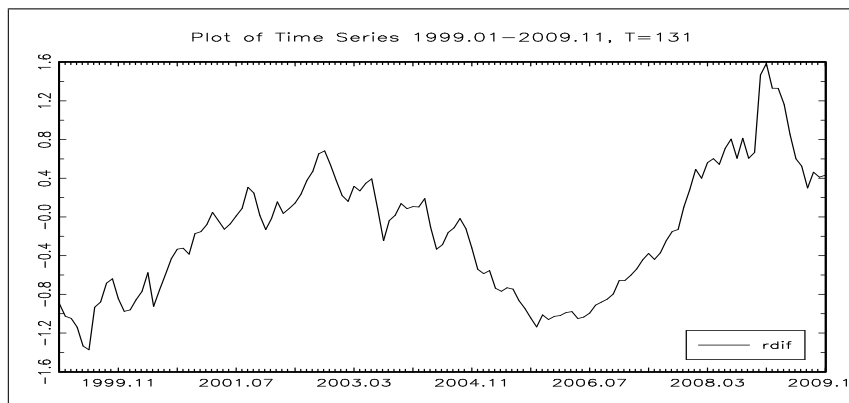
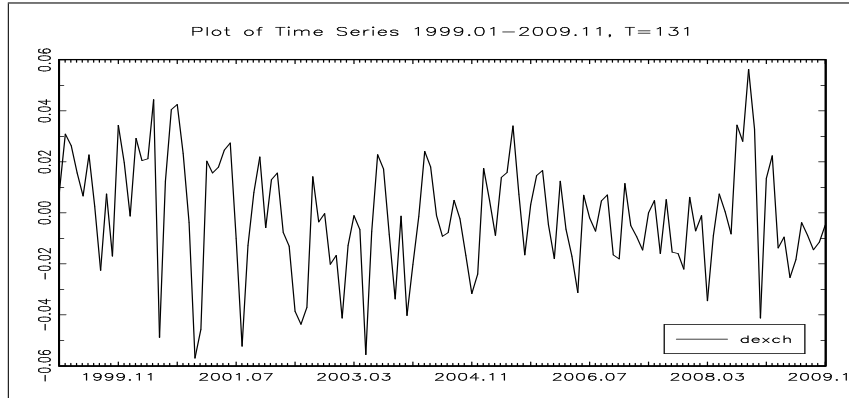
$$y_t = (y_{1t}, \dots, y_{nt}), \quad t = 1, \dots, T,$$

jossa t on ajankohta, n tarkasteltavien muuttujien lukumäärä ja T havaintojen lukumäärä.¹ Tällä kurssilla oletetaan, että havainnot ovat ajallisesti yhtä kaukana toisistaan tai niitä voidaan käsitellä sellaisina. Jos $n = 1$, on kysymyksessä *yksiulotteinen tai reaaliarvoinen aikasarja*. Kurssilla oletetaan perustiedot yksiulotteisten aikasarjojen mallintamisesta.

Esimerkki moniulotteisesta aikasarjasta saadaan tarkastelemalla vaikkapa usean pörssikurssin päivittäistä kehitystä Helsingin pörsissä vuonna 2009. Myöhemmin tarkastellaan kaksiulotteista esimerkkiä, jossa toisena komponenttina on euron ja Yhdysvaltojen dollarin välisen kuukausittaisen vaihtokurssin (euro/U.S. dollari) muutos ($dexch_t$) ajanjaksona 1999I - 2009XI. Taloustieteessä esitetyn ns. kattamattoman korkopariteettiteorian mukaan vaihtokurssin odotetun muutoksen pitäisi vastata vertailuvaluuttojen korkoeroa (euroalue/US), jota mitataan euroalueen ja Yhdysvaltojen 10 vuoden valtioiden obligaatioiden keskikorkojen erolla ($rdif_t$). Vaihtokurssin odotettu muutos ajankohtana t määritellään muuttujan $dexch_{t+1}$ ehdollisena odotusarvona ehdolla ”kaikki ajankohtana t käytävissä oleva relevantti informaatio”. Jos tämän odotuksen korvikkeena käytetään ajankohtana t tiedossa olevaa muutosta $dexch_t$, voidaan vaihtokurssin muutosta ja korkoeroa kuvaavien aikasarjojen $dexch_t$ ja $rdif_t$ odottaa olevan yhteydessä toisiinsa. Kuviossa 1.1 esitetään näistä muuttujista koostuva aineisto graafisesti. Keskeinen mielenkiinto tässä kuten moniulotteisissa aikasarja-aineistoissa yleensäkin kohdistuu tarkasteltavien muuttujien välisiin riippuvuuksiin, jotka eivät välttämättä ilmene samanaikaisina riippuvuuksina.

Kuten yksiulotteisessa tapauksessakin, vaikuttaa moniulotteisten aikasarja-aineistojen analysointiin ja siinä käytettäviin tekniikoihin se, ovatko aikasarjat stationaarisia vai ”trendimäisesti” epästationaarisia. Seuraavassa tarkastellaan tähän liittyviä yleisiä kysymyksiä moniulotteisten aikasarjojen tapauksessa.

¹Tällä kurssilla käytetään yleensä samaa merkintää satunnaismuuttujista ja niiden havaituista arvoista. Ero on ymmärrettävä asiayhteydestä. Vektorit tulkitaan matriisilaskuissa aina pystyvektoreiksi ja niistä voidaan käyttää myös matriisilaskennan merkintää eli esimerkiksi $y_t = [y_{1t} \cdots y_{nt}]'$, jossa pilkku osoittaa vektorin (samoin kuin myöhemmin matriisiin) transponointia.



Kuvio 1.1. Euron ja Yhdysvaltojen dollarin välisen kuukausittaisen vaihtokurssin (euro/U.S. dollari) muutos (ylhäällä) ja euroalueen ja Yhdysvaltojen 10 vuoden valtioiden obligaatioiden keskiporkkojen ero (alhaalla) ajanjaksona 1999I - 2009XI.

1.2 Stokastinen prosessi

Aikasarja voidaan tulkita stokastisen prosessin realisaatioksi eli satunnaisilmiön havaituksi kehityskulkuksi ajassa. *Stokastinen prosessi* tai usein vain prosessi voidaan tällä kurssilla määritellä joukoksi satunnaisvektoreita (sv) tai yksiulotteisissa tapauksessa satunnaismuuttujia (sm):

$$\{y_t, t \in \mathbb{Z}\} \quad (\text{diskreetti aika}) \quad \text{tai} \quad \{y_t, t \in \mathbb{R}\} \quad (\text{jatkuva aika}).$$

Erityisesti jatkuva-aikaisessa tapauksessa käytetään myös merkintää $y(t)$. Aikamuuttuja t voi kuulua myös johonkin \mathbb{Z} :n tai \mathbb{R} :n osajoukkoon kuten \mathbb{N} tai $\{0 \leq t \leq 1\}$.

Jatkossa merkitään yleensä lyhyesti $\{y_t\}$ tai vain y_t . Merkintä $\{y_t\}$ korostaa sitä, että stokastinen prosessi on (yleisesti ottaen) ääretönulotteinen satunnaismuuttuja. Esimerkiksi reaaliarvoinen diskreettiaikainen prosessi $\{y_t\} = \{\dots, y_{-1}, y_0, y_1, \dots\}$ saa

arvoja joukossa $\mathbb{R}^\infty = \mathbb{R} \times \mathbb{R} \times \dots$. Jatkuva-aikainen prosessi voi puolestaan saada arvoja esimerkiksi joukossa $C[0, 1]$ eli kaikkien välillä $[0, 1]$ määriteltyjen jatkuvien reaaliarvoisten funktioiden joukossa. Käytännössä havaittu aikasarja on vain yksi ääretönulotteisen satunnaismuuttujan saama arvo tai yleensä vain osa sen saamasta arvosta (esim. $y_1, \dots, y_T, T < \infty$).

Kun stokastinen prosessi tulkitaan ääretönulotteisena satunnaismuuttujana, voidaan kysyä mitä tarkoitetaan sen todennäköisyysjakaumalla, joka on määriteltävä esimerkiksi joukossa \mathbb{R}^∞ tai $C[0, 1]$. Tätä jakaumaa tarvitaan todennäköisyysteoreettisissa tarkasteluissa, kun tutkitaan sellaisten tapahtumien todennäköisyyksiä, joita ei voida ilmaista käyttäen todennäköisyyksiä

$$P \{y_{t_1} \in A_1, \dots, y_{t_m} \in A_m\}, \quad 1 \leq m < \infty, \quad (1.1)$$

eli ns. *äärellisulotteisia jakaumia*. Esimerkki tällaisesta tapahtumasta on vaikkapa $\{0 \leq y_t \leq 1 : \forall t \geq 100\}$.

Kysymys stokastisen prosessin todennäköisyysjakaumasta kuuluu kehittyneen todennäköisyyslaskennan piiriin. Keskeinen tulos stokastisten prosessien teoriasta sanoo hieman yksinkertaistaen, että prosessin (ääretönulotteinen) jakauma määräytyy yksikäsitteisesti äärellisulotteisista jakaumista (1.1). Käänteinen tulos eli prosessin äärellisulotteisten jakaumien (1.1) määräytyminen yksikäsitteisesti ääretönulotteisista jakaumista on ilmeinen. Prosessin jakauma voidaan siis karakterisoida käyttäen äärellisulotteisia jakaumia. Tästä seuraa esimerkiksi, että prosessi $\{y_t\}$ on normaalin täsmälleen silloin, kun sen kaikki äärellisulotteiset jakaumat ovat normaalisia eli sv $(y_{t_1}, \dots, y_{t_m})$ noudattaa multinormaalijakaumaa kaikilla t_1, \dots, t_m ja kaikilla $m \geq 1$ ($m \in \mathbb{Z}$, mutta jatkuva-aikaisessa tapauksessa t_1, \dots, t_m eivät välttämättä ole kokonaislukuja).

Stokastisen prosessin ensimmäiset ja toiset momentit määritellään seuraavasti (olettaen, että ne ovat äärellisinä olemassa). Prosessin y_t *odotusarvo(funktio)* on

$$\mu_t = \mathbf{E}(y_t) \quad (n \times 1)$$

ja *kovarianssifunktio*

$$\Gamma_{s,t} = \mathbf{Cov}(y_s, y_t) = \mathbf{E}[(y_s - \mu_s)(y_t - \mu_t)'] \quad (n \times n).$$

Valitsemalla viimeksi mainitussa $s = t$ saadaan prosessin *kovarianssmatriisi* $\Sigma_t = \Gamma_{t,t}$ eli

$$\Sigma_t = \mathbf{Cov}(y_t) = \mathbf{E}[(y_t - \mu_t)(y_t - \mu_t)'] \quad (n \times n).$$

Jatkuva-aikaisessa tapauksessa alaindeksit s ja t voivat saada mitä reaalityyppisiä arvoja tahansa. Käytännössä rajoitutaan usein näiden momenttien selvittämiseen (normaalissa tapauksessa muuta selvitettävää ei olekaan). Teoreettisissa tarkasteluissa oletukset korkeampiasteisista momenteista ovat kuitenkin tavallisia. Ellei toisin mainita, tulkitaan merkintöjen kuten $\mu_t = \mathbf{E}(y_t)$ ja $\Sigma_t = \mathbf{Cov}(y_t, y_t)$ sisältävän automaattisesti oletuksen, että kyseiset momentit ovat äärellisinä olemassa.

2 Stationaariset prosessit

2.1 Määritelmiä

Stationaarisuus määritellään moniulotteisessa tapauksessa samalla tavalla kuin yksiulotteisessa tapauksessa. Stokastinen prosessi y_t ($n \times 1$) on siten (vahvasti) *stationaarinen*, jos

$$\mathbb{P}\{y_{t_1} \in A_1, \dots, y_{t_m} \in A_m\} = \mathbb{P}\{y_{t_1+k} \in A_1, \dots, y_{t_m+k} \in A_m\}$$

kaikilla A_1, \dots, A_m , t_1, \dots, t_m , k ja $m \geq 1$ (jatkuva-aikaisessa tapauksessa vain m :n tarvitsee olla kokonaisluku). Toisin sanoen, stationaarisen prosessin todennäköisyysstruktuuri on aikainvariantti.

Stationaarisuuden määritelmästä seuraa erityisesti, että sv:t y_t ovat samoin jakautuneita kaikilla t ja myös, että sv:t (y_t, y_{t+h}) ja (y_s, y_{s+h}) ovat samoin jakautuneita kaikilla t ja s ja jokaisella (kiinteällä) h . Näiden ominaisuuksien lisäksi (vahvasti) stationaarisilla prosesseilla on se hyödyllinen ominaisuus, että ne säilyttävät luonteensa funktiomuunnoksissa. Erityisesti siis, jos z_t on (vahvasti) stationaarinen diskreetti-aikainen prosessi ($t \in \mathbb{Z}$), niin samoin on $y_t = g(z_{t+h}, \dots, z_{t-k})$ ($h, k \geq 0$) millä tahansa ”siistillä” (eli esim. jatkuvalla tai yleisemmin ns. mitallisella) funktiolla g . Lisäksi h :n tai k :n tai molempien paikalla voi olla ∞ . Kun h ja k ovat äärellisiä seuraa tämä ominaisuus jokseenkin suoraan (vahvan) stationaarisuuden määritelmästä.

Prosessi y_t ($n \times 1$) on *heikosti stationaarinen* tai *kovarianssistationaarinen*, jos sen toiset (ja siten myös ensimmäiset) momentit ovat äärellisinä olemassa ja ne ovat aikainvariantteja eli

$$\mathbb{E}(y_t) = \mu \quad \text{ja} \quad \text{Cov}(y_t, y_{t+k}) = \Gamma_k \quad \text{kaikilla } t \text{ ja } k.$$

(Jatkuva-aikaisessa tapauksessa $t \in \mathbb{R}$ ja myös $k \in \mathbb{R}$).

Vahvasta stationaarisuudesta ja toisten momenttien äärellisyydestä seuraa ilmeisesti heikko stationaarisuus. Käänteinen tulos pitää paikkansa normaalilla prosesseilla, mutta *ei yleisesti*. Huomaa myös, että heikosti stationaariset prosessit eivät yleisesti säilytä luonnettaan funktiomuunnoksissa. Jatkossa pelkkä ’stationaarisuus’ viittaa aina vahvaan stationaarisuuteen (tämä puhetapa ei ole yleisesti käytössä).

Yksinkertaisin esimerkki (diskreettiaikaisesta) stationaarisesta prosessista on jono riippumattomia samoin jakautuneita satunnaisvektoreita (stationaarisuus voidaan todeta suoraan määritelmästä). Jos y_t ($n \times 1$) on tällainen jono ja lisäksi $\mathbb{E}(y_t) = \mu$ ja $\text{Cov}(y_t) = \Sigma$ ovat äärellisinä olemassa, merkitään $y_t \sim \text{iid}(\mu, \Sigma)$ ($t \in \mathbb{Z}$). Tällaista prosessia sanotaan (vahvaksi) *valkoiseksi kohinaksi*. Jos y_t on lisäksi normaalisti jakautunut, merkitään $y_t \sim \text{nid}(\mu, \Sigma)$. Jos edellä oletetaan riippumattomuuden asemesta (lievemmin) korreloimattomuus eli $\text{Cov}(y_s, y_t) = 0$, $t \neq s$, on y_t heikosti stationaarinen eli ns. *heikko valkoinen kohina*.

Yksiulotteisessa tapauksessa kovarianssifunktio on symmetrinen viipymän k suhteen. Moniulotteisessa tapauksessa näin ei ole. Sen sijaan pätee

$$\begin{aligned}\Gamma_{-k} &= \mathbf{E} [(y_t - \mu)(y_{t-k} - \mu)'] \\ &= (\mathbf{E} [(y_{t-k} - \mu)(y_t - \mu)'])' \\ &= \Gamma_k' .\end{aligned}$$

Kuten yksiulotteisessa tapauksessakin, riittää siis tuntea Γ_k vain, kun $k \geq 0$ (tai $k \leq 0$).

Kovarianssifunktiosta saadaan prosessin komponenttien autokovarianssifunktiot ja ristikovarianssifunktiot. Olkoon $y_t = (y_{1t}, \dots, y_{nt})$ ja $\Gamma_k = [\gamma_{ab,k}]$ ($a, b = 1, \dots, n$). Tällöin prosessin y_{at} *autokorrelaatiofunktio* on

$$\rho_{a,k} = \gamma_{a,k} / \gamma_{a,0}, \quad a = 1, \dots, n$$

ja prosessien y_{at} ja y_{bt} välinen *ristikorrelaatiofunktio* on

$$\rho_{ab,k} = \gamma_{ab,k} / \sqrt{\gamma_{a,0}\gamma_{b,0}} \quad a, b = 1, \dots, n.$$

Edellä $\gamma_{a,k} = \gamma_{aa,k}$ ja $\gamma_{ab,k}$ ($a \neq b$) ovat vastaavasti *autokovarianssi-* ja *ristikovarianssifunktioita*.

Käytännön kannalta stationaariseen prosessiin voidaan yleensä liittää oletus

$$\Gamma_k \rightarrow 0, \quad \text{kun } |k| \rightarrow \infty, \quad (2.1)$$

jossa suppeneminen voidaan tulkita tapahtuvaksi alkioittain (matriisiarvoisten jonojen suppenemista on tarkasteltu Liitteessä A.1). Tällöin y_t ja y_{t+k} ovat lähes korreloimattomia, kun $|k|$ on ”suuri” eli, kun y_t ja y_{t+k} ovat ajallisesti ”kaukana” toisistaan. (Huomaa kuitenkin, että ehto (2.1) ei seuraa stationaarisuudesta.) Jos ehto (2.1) pätee, riittää y_t :n ensimmäisten ja toisten momenttien estimoimiseksi (olenaisesti) estimoida vain äärellinen määrä parametreja (eli μ ja Γ_k , $k = -K, \dots, K$, ”suurella” K :n arvolla).

Ellei toisin mainita, oletetaan tarkasteltavat prosessit seuraavassa diskreettiaikaisiksi.

2.2 Lineaarinen prosessi

Tarkastellaan aluksi yksiulotteista (kausaalista) lineaarista prosessia²

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad \varepsilon_t \sim \text{iid}(0, \omega^2), \quad t \in \mathbb{Z}, \quad (2.2)$$

²Lineaarinen prosessi voidaan määritellä myös yleisemmin sallimalla yhtälössä (2.2) indeksille j arvot $j < 0$. Tällaista lineaarista prosessia kutsutaan toisinaan ei-kausaaliseksi.

jossa $\psi_0 = 1$ ja

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty. \quad (2.3)$$

Yhtälön (2.2) oikealla puolella oleva ääretön summalauseke voidaan tulkita osasummien $\sum_{j=0}^N \psi_j \varepsilon_{t-j}$ jonon kvadraattiseksi raja-arvoksi ($N \rightarrow \infty$). Stokastisten prosessien (kehittyneestä) teoriasta tiedetään, että ehdon (2.3) voimassa ollessa tämä kvadraattinen raja-arvo on olemassa ja lisäksi, että raja-arvolla on äärellinen toinen momentti. Yhtälössä (2.2) olevasta iid-prosessista käytetään hieman tilanteesta riippuen nimitystä *virhe*, *sokki* tai *innovaatio*. Vahvan valkoisen kohinan asemasta voitaisiin ε_t olettaa monissa tarkasteluissa lievemmin heikoksi valkoiseksi kohinaksi (samoin sen alempana esitettävä moniulotteinen vastine).

Koska prosessi $\{\varepsilon_t\}$ on (vahvasti) stationaarinen pätee sama myös prosessille $\{y_t\}$. Heikko stationaarisuus nähdään seuraavassa esitettävän moniulotteisen yleistyksen yhteydessä.

Moniulotteisessa tapauksessa $y_t = (y_{1t}, \dots, y_{nt})$ määritellään (kausaalinen) *lineaarinen prosessi* yhtälöllä

$$y_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}, \quad \varepsilon_t \sim \text{iid}(0, \Omega), \quad t \in \mathbb{Z}, \quad (2.4)$$

jossa $\Psi_0 = I_n$ ($n \times n$ yksikkömatriisi) ja kerroinmatriisien Ψ_j täytyy toteuttaa sopiva yhtälön (2.3) vastine. Käyttäen (Euklidisen) matriisnormin käsitettä tämä ehto voidaan esittää kätevästi muodossa

$$\sum_{j=0}^{\infty} \|\Psi_j\|^2 < \infty. \quad (2.5)$$

Normi $\|\Psi_j\|$ määritellään yksinkertaisesti muodostamalla matriisin Ψ_j alkioista vektori ja soveltamalla tavallista \mathbb{R}^{n^2} :n (Euklidista) vektorinormia. Määritelmä on esitetty yksityiskohtaisemmin Liitteessä A.1, missä todetaan myös, että ehto (2.5) on yhtäpitävää sen kanssa, että ehdon (2.3) vastine pätee matriisijonon Ψ_j kaikille alkioille.

Yhtälön (2.4) oikea puoli voidaan tulkita komponenteittain kvadraattisina raja-arvoina aivan kuten yksiulotteisessa tapauksessa (2.2). Prosessin $\{y_t\}$ (vahva) stationaarisuus seuraa prosessin $\{\varepsilon_t\}$ vastaavasta ominaisuudesta kuten yksiulotteisessa tapauksessakin. Heikko stationaarisuus todetaan laskelmalla

$$\mathbf{E}(y_t) = \sum_{j=0}^{\infty} \mathbf{E}(\Psi_j \varepsilon_{t-j}) = \sum_{j=0}^{\infty} \Psi_j \mathbf{E}(\varepsilon_{t-j}) = 0$$

ja

$$\begin{aligned}
\text{Cov}(y_t, y_{t+k}) &= \mathbf{E}(y_t y'_{t+k}) \\
&= \mathbf{E}\left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \Psi_i \varepsilon_{t-i} \varepsilon'_{t+k-j} \Psi'_j\right) \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \Psi_i \mathbf{E}(\varepsilon_{t-i} \varepsilon'_{t+k-j}) \Psi'_j \\
&= \sum_{j=0}^{\infty} \Psi_j \Omega \Psi'_{j+k}.
\end{aligned}$$

Laskelmissa on käytetty odotusarvon lineaarisuusominaisuutta ja innovaatioiden ε_t riippumattomuutta, josta seuraa $\mathbf{E}(\varepsilon_{t-i} \varepsilon'_{t+k-j}) = 0$, kun $j \neq k + i$, ja $\mathbf{E}(\varepsilon_t \varepsilon'_t) = \Omega$.³ Koska matriisinormin ominaisuuksien perusteella $\|\Psi_j \Omega \Psi'_j\| \leq \|\Omega\| \|\Psi_j\|^2$, nähdään kolmioepäyhtälön avulla, että $\text{Cov}(y_t)$ on ehdon (2.5) voimassa ollessa äärellinen (ks. Liite A.1), joten y_t on heikosti stationaarinen.

Edellä esitetyt odotusarvoa $\mathbf{E}(y_t)$ ja kovarianssia $\text{Cov}(y_t, y_{t+k})$ koskevat laskelmat pätevät myös, kun innovaatiot ovat vain korreloimattomia (eli $\text{Cov}(\varepsilon_t, \varepsilon_{t+k}) = 0$, $k \neq 0$), joten heikko stationaarisuus pätee tällöinkin. Huomaa myös, että tarkastelemalla y_t :n asemesta prosessia $y_t - \mu$ voidaan ottaa huomioon nolasta poikkeava odotusarvo.

2.3 VAR(p)–prosessi

Astetta p olevan yksiulotteisen autoregressiivisen prosessin eli $AR(p)$ –prosessin tai $AR(p)$ –mallin lähtökohtana on yhtälö

$$y_t = a_1 y_{t-1} + \cdots + a_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \omega^2) \quad (t \in \mathbb{Z}). \quad (2.6)$$

Toisin sanoen, havaittavan prosessin y_t arvo ajankohtana t riippuu lineaarisesti sen omista viipymistä y_{t-1}, \dots, y_{t-p} ja ei-havaittavasta virhetermistä ε_t aivan kuten lineaarisessa regressiomallissa.

Moniulotteisessa tapauksessa $y_t = (y_{1t}, \dots, y_{nt})$ otetaan lähtökohdaksi yhtälön (2.6) yleistys

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \Omega) \quad (t \in \mathbb{Z}), \quad (2.7)$$

jossa A_1, \dots, A_p ovat $n \times n$ kerroinmatriiseja ja ε_t ($n \times 1$) on ei-havaittava virhetermi (tai sokki tai innovaatio). Ellei toisin mainita, oletetaan kovarianssimatriisi Ω positiivisesti definiitiksi (ja siten epäsingulaariseksi eli kääntyväksi).⁴ Tässä yhteydessä

³Lineaarisuusominaisuuden käyttäminen tässä eli odotusarvon ja äärettömän summan järjestyksen vaihtamisen luvallisuus ei ole itsestään selvää, vaan perustuu äärettömän summan (2.4) kvadraattisen konvergenssin tapaan stokastisten prosessien (kehittyneeseen) teoriaan.

⁴Symmetrinen matriisi A ($k \times k$) on positiivisesti definiitti, jos $x'Ax > 0$ kaikilla nolasta poikkeavilla vektoreilla x ($k \times 1$). Positiivisesti definiitti matriisi on tunnetusti epäsingulaarinen. Symmetristä matriisia A sanotaan positiivisesti semidefiniitiksi, jos $x'Ax \geq 0$ kaikilla x ($k \times 1$).

puhutaan $VAR(p)$ -prosessista tai $VAR(p)$ -mallista ('V' viittaa sanaan vektori). Ideana on siis selittää jokaista vektorin y_t komponenttia sen omilla viipymillä ja myös y_t :n muiden komponenttien viipymillä lineaarisen regressiomallin tapaan. Moniulotteisessa tapauksessa voidaan siten tarkastella myös muuttujien välisiä viivästettyjä riippuvuuksia kuten esimerkiksi sitä, miten y_{2t} vaikuttaa y_{1t} :n tulevaan kehitykseen ja miten y_{1t} vaikuttaa y_{2t} :n tulevaan kehitykseen.

Jotta yhtälö (2.7) määritteli stationaarisen prosessin (tai edes stokastisen prosessin), täytyy kerroinmatriisien A_1, \dots, A_p toteuttaa sopiva ehto. Tämän ehdon esittämiseksi on kätevää tarkastella yhtälöön (2.7) läheisesti liittyvää yhtälöä

$$\mathbf{y}_t = \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \quad (t \in \mathbb{Z}), \quad (2.8)$$

jossa

$$\mathbf{A} = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_n & 0 & \cdots & 0 & 0 \\ 0 & I_n & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_n & 0 \end{bmatrix} \quad (np \times np).$$

$$\mathbf{y}_t = [y'_t \cdots y'_{t-p+1}]' \quad (np \times 1) \quad \text{ja} \quad \boldsymbol{\varepsilon}_t = [\varepsilon'_t \ 0' \cdots 0']' \quad (np \times 1).$$

Nähdään helposti, että yhtälö (2.8) pätee jos ja vain jos yhtälö (2.7) pätee. Oletuksesta $\varepsilon_t \sim \text{iid}(0, \boldsymbol{\Omega})$ seuraa, että

$$\boldsymbol{\varepsilon}_t \sim \text{iid}(0, \boldsymbol{\Omega}), \quad \boldsymbol{\Omega} = \text{diag}[\Omega \ 0 \cdots 0],$$

jossa $\text{diag}[\Omega \ 0 \cdots 0]$ on lohkodeagonaalinen matriisi diagonaaleilla osoitetut $(n \times n)$ matriisit. Kovarianssimatriisi $\boldsymbol{\Omega}$ on siis singulaarinen (ei-kääntyvä) silloinkin, kun Ω on positiivisesti definiitti. Tästä huolimatta voidaan prosessia (2.8) käyttää apuvälineenä tutkittaessa $VAR(p)$ -prosessin (2.7) ominaisuuksia.

$VAR(p)$ -prosessin stationaarisuusehto. Seuraavassa todetaan, että yhtälö (2.8) määrittelee stationaarisen prosessin, jos matriisin \mathbf{A} kaikki ominaisarvot ovat itseisarvoltaan ykköstä pienempiä. On helppo todeta, että tämä ehto voidaan esittää yhtäpitävästi käyttäen determinanttiehtoa

$$\det(I_{np} - \mathbf{A}z) \neq 0, \quad |z| \leq 1 \quad (z \in \mathbb{C}), \quad (2.9)$$

jossa $|z|$ on kompleksiluvun z itseisarvo (ks. Liite A.2).⁵ Jos yhtälön (2.8) määrittelemä prosessi \mathbf{y}_t on stationaarinen, niin samoin on yhtälön (2.7) määrittelemä prosessi y_t (y_t on lineaarinen muunnos \mathbf{y}_t :stä). Kun edellä mainittua ehtoa sovelletaan

⁵Jos $z \in \mathbb{C}$ eli $z = x + iy$ ($i^2 = -1$), niin $|z| = \sqrt{x^2 + y^2}$ on z :n itseisarvo.

yksiulotteiseen AR-prosessiin (2.6) tapauksessa $p = 1$, saadaan tuttu AR(1)-prosessin riittävä stationaarisuusehto $|a_1| < 1$.

Tarkastellaan ensin yhtälöä (2.8) tapauksessa $t \geq 1$, jolloin prosessi \mathbf{y}_t on hyvin määritelty, kun alkuarvo \mathbf{y}_0 spesifioidaan. Peräkkäisillä sijoituksilla saadaan

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{A}\mathbf{y}_0 + \boldsymbol{\varepsilon}_1 \\ \mathbf{y}_2 &= \mathbf{A}\mathbf{y}_1 + \boldsymbol{\varepsilon}_2 \\ &= \mathbf{A}^2\mathbf{y}_0 + \mathbf{A}\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 \\ &\vdots \\ \mathbf{y}_t &= \mathbf{A}^t\mathbf{y}_0 + \sum_{j=0}^{t-1} \mathbf{A}^j\boldsymbol{\varepsilon}_{t-j}. \end{aligned}$$

”Sopivan” alkuarvon määrittelemiseksi todetaan, että ehdon (2.9) voimassa ollessa $\mathbf{A}^j \rightarrow 0$ geometrisesti, kun $j \rightarrow \infty$, mistä seuraa $\sum_{j=0}^{\infty} \|\mathbf{A}^j\|^2 < \infty$ (ks. Liite A.3). Edellisessä jaksossa todetun nojalla voidaan siten määritellä lineaarinen prosessi

$$\mathbf{y}_t^* = \sum_{j=0}^{\infty} \mathbf{A}^j\boldsymbol{\varepsilon}_{t-j}, \quad t \in \mathbb{Z},$$

joka on sekä vahvasti että heikosti stationaarinen (eli vahvasti stationaarinen ja toiset momentit äärellisinä olemassa). Valitsemalla alkuarvoksi $\mathbf{y}_0 = \mathbf{y}_0^*$ saadaan lineaarinen prosessi

$$\mathbf{y}_t = \mathbf{A}^t \sum_{j=0}^{\infty} \mathbf{A}^j\boldsymbol{\varepsilon}_{-j} + \sum_{j=0}^{t-1} \mathbf{A}^j\boldsymbol{\varepsilon}_{t-j} = \sum_{j=0}^{\infty} \mathbf{A}^j\boldsymbol{\varepsilon}_{t-j}, \quad t \geq 1.$$

Toisin sanoen, $\{\mathbf{y}_t, t \geq 1\}$ on stationaarinen. Lisäksi on selvää, että esitetty ratkaisu voidaan laajentaa kaikille $t \in \mathbb{Z}$ eli ehdon (2.9) voimassa ollessa yhtälö (2.8) määrittelee stationaarisen prosessin

$$\mathbf{y}_t = \sum_{j=0}^{\infty} \mathbf{A}^j\boldsymbol{\varepsilon}_{t-j}, \quad (t \in \mathbb{Z}). \quad (2.10)$$

Määrittelemällä matriisi $J = [I_n : 0 : \dots : 0]$ ($n \times np$) nähdään, että

$$\mathbf{y}_t = J\mathbf{y}_t \quad \text{ja} \quad \boldsymbol{\varepsilon}_t = J'\boldsymbol{\varepsilon}_t,$$

joten yhtälöstä (2.10) seuraa

$$\mathbf{y}_t = \sum_{j=0}^{\infty} \Psi_j\boldsymbol{\varepsilon}_{t-j}, \quad (t \in \mathbb{Z}), \quad (2.11)$$

jossa $\Psi_j = J\mathbf{A}^jJ'$ ja siten $\Psi_0 = I_n$. Koska $\mathbf{A}^j \rightarrow 0$ geometrisesti, kun $j \rightarrow \infty$, pätee vastaava ominaisuus myös matriiseille Ψ_j , joten erityisesti $\sum_{j=0}^{\infty} \|\Psi_j\|^2 < \infty$.

Jos prosessia y_t tarkastellaan rajoittuen ajankohtiin $t \geq 1$ ja sallien alkuarvojen y_{-p+1}, \dots, y_0 epästationaarisuus, on $\{y_t, t \geq 1\}$ epästationaarinen. Edellä esitetystä nähdään, että alkuarvojen vaikutus häviää kuitenkin t :n kasvaessa ja prosessi ”stationarisoituu”. Tässä yhteydessä puhutaan usein *stabiilisuudesta* tai *asymptoottisesta stationaarisuudesta* ja ehtoa (2.9) kutsutaan *stabiilisuusehdoksi*. Myöhemmin esitettävät estimointi- ja testaustulokset soveltuvat myös tällaisten prosessien analysointiin.

Stationaarisuusehdon vaihtoehtoinen muotoilu. Määritellään mihin tahansa (mahdollisesti vektoriarvoiseen) jonoon x_t liittyvä viivästysoperaattori \mathbf{B} yhtälöllä $\mathbf{B}x_t = x_{t-1}$ ja induktiivisesti $\mathbf{B}^k x_t = x_{t-k}$ ($\mathbf{B}^0 x_t = x_t$ ja $\mathbf{B}^{-k} x_t = x_{t+k}$). Tällöin VAR(p)-prosessi (2.7) voidaan kirjoittaa lyhyesti

$$A(\mathbf{B})y_t = \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \Omega) \quad (t \in \mathbb{Z}), \quad (2.12)$$

jossa $A(\mathbf{B}) = I_n - A_1\mathbf{B} - \dots - A_p\mathbf{B}^p$. Voidaan osoittaa, että stationaarisuusehto (2.9) on yhtäpitävä seuraavan ehdon kanssa (perustelu jätetään tehtäväksi):

$$\det A(z) \neq 0, \quad |z| \leq 1. \quad (2.13)$$

Tässä $A(z)$ on esimerkki polynomimatriisista eli matriisista, jonka alkiot ovat polynomeja (vastaavasti yhtälössä (2.12), jossa argumenttina on viivästysoperaattori \mathbf{B}). Vaihtoehtoisesti $A(z)$ voidaan tulkita matriisipolynomiksi eli polynomiksi, jonka kertoimet ovat matriiseja. Polynomimatriiseilla laskutoimitukset määritellään kuten tavanomaisilla reaalilukumatriiseilla (ks. Liite A.3). Yhtälö (2.12) voidaan siten ratkaista formaalisti kertomalla vasemmalta käänteismatriisilla $A(\mathbf{B})^{-1}$. Ratkaisuksi täytyy tietenkin tulla (2.11) eli lyhyesti ilmaistuna

$$y_t = \Psi(\mathbf{B})\varepsilon_t \quad (t \in \mathbb{Z}),$$

jossa $\Psi(\mathbf{B}) = \sum_{j=0}^{\infty} \Psi_j \mathbf{B}^j = A(\mathbf{B})^{-1}$. Stationaarisuusehtoa (2.13) voidaan havainnollistaa käyttämällä tulosta $A(\mathbf{B})^{-1} = [1/\det(A(\mathbf{B}))] A(\mathbf{B})^*$, jossa $A(\mathbf{B})^*$ on (äärelisasteinen) matriisipolynomi (ks. Liite A.4). Ehdon (2.13) riittävyys stationaarisuudelle nähdään siten käyttäen (tunnetuksi oletettua) yksiulotteisen AR(p)-prosessin stationaarisuusehtoa.

Yhtälön (2.11) yhteydessä todettiin, että $\Psi_j = J\mathbf{A}^j J'$. Vaihtoehtoinen esitys matriiseille Ψ_j saadaan käyttäen identiteettiä $\Psi(\mathbf{B})A(\mathbf{B}) = I_n$ eli

$$\begin{aligned} I_n &= (\Psi_0 + \Psi_1\mathbf{B} + \Psi_2\mathbf{B}^2 + \dots)(I_n - A_1\mathbf{B} - \dots - A_p\mathbf{B}^p) \\ &= \Psi_0 + (\Psi_1 - \Psi_0 A_1)\mathbf{B} + (\Psi_2 - \Psi_1 A_1 - \Psi_0 A_2)\mathbf{B}^2 + \dots \end{aligned}$$

Tästä seuraa

$$\begin{aligned} I_n &= \Psi_0 \\ 0 &= \Psi_1 - \Psi_0 A_1 \\ 0 &= \Psi_2 - \Psi_1 A_1 - \Psi_0 A_2 \\ &\vdots \end{aligned}$$

mistä nähdään edelleen, että matriisit Ψ_j voidaan laskea rekursiivisesti yhtälöistä

$$\begin{aligned}\Psi_0 &= I_n \\ \Psi_j &= \sum_{i=1}^j \Psi_{j-i} A_i, \quad j = 1, 2, \dots,\end{aligned}$$

jossa $A_i = 0$, kun $i > p$.

Impulssivasteet. Kuten aiemmin jo mainittiin, voidaan VAR(p)-malli ajatella usean lineaarisen regressiomallin systeeminä, jossa selittävät muuttujat ovat selitettävien muuttujien viivästettyjä arvoja. Tavanomaiseen regressiomalliin verrattuna selittävät muuttujat on siten valittu erityisellä tavalla. Ei-havaittavien muuttujien ε_t tulkinta jäännös- tai virhetermeinä on kuitenkin pitkälti samanlainen kuin regressiomalleissa. Regressiomallien tapaan voidaan kerroinmatriisien A_1, \dots, A_p alkioiden tulkita kuvaavan selittävien muuttujien y_{t-1}, \dots, y_{t-p} komponenttien vaikutusta selitettäviin muuttujiin eli y_t :n komponentteihin. Näiden kerroinmatriisien alkioiden tulkinnallinen hyöty on kuitenkin varsin rajallinen, sillä yleensä ollaan kiinnostuneita muuttujien välisistä dynaamisista vaikutuksista eli esimerkiksi siitä, miten y_{1t} :ssä tapahtuva muutos vaikuttaa kaikkien y_t :n komponenttien tulevaan kehitykseen. Tällaisten dynaamisten vaikutusten selvittäminen suoraan kerroinmatriisien A_1, \dots, A_p alkioiden avulla on hankalaa. Tämän näkee havainnollisesti jo tapauksessa $n = 2$ kaaviokuvasta, johon on merkitty nuolilla se, miten muuttujat y_{1t} ja y_{2t} vaikuttavat tuleviin arvoihinsa $y_{1,t+j}$ ja $y_{2,t+j}$, $j = 1, 2, \dots$. Vaikka kysymyksessä on vain kaksiulotteinen tapaus, tulee kuviosta varsin sekavan näköinen.

Vaihtoehtoinen tulkinta perustuu ajatukseen, jossa virhetermin ε_t komponentit tulkitaan prosessiin y_t vaikuttaviksi impulsseiksi tai sokeiksi. Tällöin kiinnostava kysymys on, miten tällaisen yhden yksikön kokoisen impulssin vaikutus ilmenee prosessin y_t komponenteissa ja niiden tulevissa arvoissa. Tähän kysymykseen palataan yksityiskohtaisemmin myöhemmin. Yhtälön (2.11) perusteella on kuitenkin helppo uskoa, että matriisin Ψ_j alkiot kuvaavat tällaisen ajankohtana t_0 tapahtuneen impulssin vaikutusta prosessin y_t komponenttien saamiin arvoihin ajankohtana $t_0 + j$. Tästä syystä matriisin Ψ_j alkioita sanotaan *impulssivasteiksi*.

VAR(p)-prosessin odotusarvo ja kovarianssifunktio. Koska stationaarisella VAR(p)-prosessilla on lineaarinen esitys (2.11), voidaan sen odotusarvo ja kovarianssifunktio johtaa jaksossa 2.2 esitetyllä tavalla. Tulokseksi saadaan

$$E(y_t) = 0 \quad \text{ja} \quad \Gamma_k = \sum_{j=0}^{\infty} \Psi_j \Omega \Psi_{j+k}' = \Gamma_{-k}', \quad k \geq 0,$$

jossa Ψ_j on kuten yhtälössä (2.11). Koska $\Psi_j \rightarrow 0$ geometrisesti, kun $j \rightarrow \infty$, vaimenee myös kovarianssifunktio Γ_k geometrisesti nollaan viipymän k kasvaessa rajatta.

Vaimeneminen on sitä hitaampaa mitä lähempänä ykköstä yhtälössä (2.8) määritellyn matriisin \mathbf{A} suurin ominaisarvo on itseisarvoltaan (vrt. yhtälöt (2.10) ja (2.11) ja huomaa, että tapauksessa $n = p = 1$ matriisi \mathbf{A} on AR-kerroin ja \mathbf{A}^k on k . autokorrelaatiokerroin). Kovarianssifunktio voidaan selvittää myös seuraavalla vaihtoehdoisella tavalla.

Kertomalla yhtälö (2.7) oikealta (transponoidulla) vektorilla y'_{t-k} ja ottamalla odotusarvo puolittain saadusta yhtälöstä saadaan

$$\mathbf{E}(y_t y'_{t-k}) = A_1 \mathbf{E}(y_{t-1} y'_{t-k}) + \cdots + A_p \mathbf{E}(y_{t-p} y'_{t-k}) + \mathbf{E}(\varepsilon_t y'_{t-k}).$$

Kun $k = 0$, nähdään yhtälöstä (2.11), että $\mathbf{E}(\varepsilon_t y'_t) = \mathbf{E}(\varepsilon_t \varepsilon'_t) = \Omega$, joten ottamalla huomioon tulos $\mathbf{E}(y_t) = 0$, tulee edellä johdetusta yhtälöstä

$$\Gamma_0 = A_1 \Gamma_1 + \cdots + A_p \Gamma_p + \Omega.$$

Kun $k > 0$, nähdään vastaavasti, että $\mathbf{E}(\varepsilon_t y'_{t-k}) = 0$ ja edelleen, että

$$\Gamma_{-k} = A_1 \Gamma_{1-k} + \cdots + A_p \Gamma_{p-k}, \quad k > 0,$$

tai yhtäpitävästi

$$\Gamma'_k = A_1 \Gamma'_{k-1} + \cdots + A_p \Gamma'_{k-p}, \quad k > 0.$$

Kuten yksiulotteisessa tapauksessakin, sanotaan viimeksi mainittuja yhtälöitä *Yule-Walker -yhtälöiksi*. Niiden avulla voidaan kerroinmatriisit A_1, \dots, A_p lausua kovarianssimatriisien $\Gamma_0, \dots, \Gamma_{p-1}$ funktiona (perustelu jätetään tehtäväksi). Lisäksi niiden avulla voidaan laskea kovarianssimatriisit $\Gamma_k, k \geq p$, kun kerroinmatriisit A_1, \dots, A_p on annettu ja kovarianssimatriisit $\Gamma_0, \dots, \Gamma_{p-1}$ tunnetaan. Koko kovarianssifunktion Γ_k selvittämiseksi (kerroinmatriisien A_1, \dots, A_p funktiona) on siten vielä selvitettävä kovarianssimatriisien $\Gamma_0, \dots, \Gamma_{p-1}$ riippuvuus kerroinmatriiseista A_1, \dots, A_p ja (kuten ilmeistä) kovarianssimatriisista Ω .

Koska $\mathbf{E}(y_t) = 0$ ja $\Gamma_{-k} = \Gamma'_k$, pätee yhtälössä (2.8) määritellylle (stationaarille) prosessille \mathbf{y}_t

$$\begin{aligned} \text{Cov}(\mathbf{y}_t) &= \mathbf{E} \left(\begin{bmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{bmatrix} \begin{bmatrix} y'_t & \cdots & y'_{t-p+1} \end{bmatrix} \right) \\ &= \begin{bmatrix} \Gamma_0 & \Gamma_{-1} & \cdots & \Gamma_{-p+1} \\ \Gamma_1 & \Gamma_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \Gamma_{-1} \\ \Gamma_{p-1} & \cdots & \Gamma_1 & \Gamma_0 \end{bmatrix} := \mathbf{\Gamma}_0. \end{aligned}$$

Koska $E(\varepsilon_t \mathbf{y}'_{t-k}) = 0$, $k > 0$, on $E(\varepsilon_t \mathbf{y}'_{t-1}) = 0$, joten yhtälöstä (2.8) saadaan matriisille Γ_0 yhtälö

$$\begin{aligned}\Gamma_0 &= E \left[(\mathbf{A} \mathbf{y}_{t-1} + \varepsilon_t) (\mathbf{A} \mathbf{y}_{t-1} + \varepsilon_t)' \right] \\ &= \mathbf{A} E(\mathbf{y}_{t-1} \mathbf{y}'_{t-1}) \mathbf{A}' + E(\varepsilon_t \varepsilon_t') \\ &= \mathbf{A} \Gamma_0 \mathbf{A}' + \Omega,\end{aligned}$$

jossa $\Omega = \text{Cov}(\varepsilon_t)$ kuten aikaisemminkin.

Matriisi Γ_0 voidaan ratkaista matriisien \mathbf{A} ja Ω funktiona edellä johdetusta yhtälöstä. Ratkaisukaavan esittämisessä tarvitaan ns. $\text{vec}(\cdot)$ -operaattoria ja Kroneckerin tuloa. Edellinen muodostaa matriisista vektorin asettamalla sarakkeet alekkain alkaen ensimmäisestä (ks. Liite A.1). Dimensioista riippumatta matriisien A ja B Kroneckerin tulo $A \otimes B$ määritellään yhtälöllä $A \otimes B = [a_{ij}B]$, jossa a_{ij} on matriisin A tyypillinen alkio (ks. Liite A.5). Jättäen perustelut tehtäväksi todetaan, että mille tahansa matriiseille A , B ja C pätee $\text{vec}(A + B) = \text{vec}(A) + \text{vec}(B)$ ja $\text{vec}(ABC) = (C' \otimes A) \text{vec}(B)$ (olettaen summa ja tulo määritellyiksi). Edellä johdetusta yhtälöstä saadaan siten

$$\text{vec}(\Gamma_0) = (\mathbf{A} \otimes \mathbf{A}) \text{vec}(\Gamma_0) + \text{vec}(\Omega)$$

eli

$$(I_{n^2 p^2} - \mathbf{A} \otimes \mathbf{A}) \text{vec}(\Gamma_0) = \text{vec}(\Omega),$$

josta voidaan ratkaista⁶

$$\text{vec}(\Gamma_0) = (I_{n^2 p^2} - \mathbf{A} \otimes \mathbf{A})^{-1} \text{vec}(\Omega).$$

Kun matriisit A_1, \dots, A_p ja Ω tunnetaan, tunnetaan tämän yhtälön oikea puoli ja matriisin Γ_0 määritelmän nojalla siten myös matriisit $\Gamma_0, \dots, \Gamma_{p-1}$.

VAR(p)-prosessin ennustaminen. Tarkastellaan VAR(p)-prosessin ennustamista teoreettisessa tilanteessa, jossa prosessin parametrien arvot tunnetaan. Käytännössä joudutaan tietenkin käyttämään estimaatteja.

Todennäköisyyslaskennasta tiedetään, että $sm:n$ Y $sv:iin$ X perustuva ”optimaalinen” ennuste on $E(Y|X)$ eli $Y:n$ ehdollinen odotusarvo ehdolla X (ks. Liite A.6, jossa motivoidaan myös merkintä $E(Y|X)$ ja esitetään tarvittavat ehdollisen odotusarvon ominaisuudet). Tässä ”optimaalinen” viittaa ennusteeseen, joka minimoi ennusteen keskineliövirheen. Jos Y on vektori, pätee sama tulos, kun kriteerinä on keskineliövirhematriisin minimointi. Tällä tarkoitetaan seuraavaa. Olkoon $\tilde{Y} = g(X)$

⁶Käänteismatriisin olemassaolo seuraa, sillä stationaarisuusoletuksen mukaan matriisin \mathbf{A} ominaisarvot ovat itseisarvoltaan ykköstä pienempiä ja matriisin $\mathbf{A} \otimes \mathbf{A}$ ominaisarvojen voidaan puolestaan osoittaa olevan $\mathbf{A}:n$ ominaisarvojen tuloja, joten nekin ovat itseisarvoltaan ykköstä pienempiä. Tästä seuraa erityisesti, että matriisi $I_{n^2 p^2} - \mathbf{A} \otimes \mathbf{A}$ on epäsingulaarinen ja mainittu (yksikäsitteinen) ratkaisu on olemassa.

mikä tahansa sv:n Y ennuste, joka on sv:n X funktio. Tällöin

$$\text{MSE}(\tilde{Y}) = \mathbf{E}[(Y - \tilde{Y})(Y - \tilde{Y})']$$

on ennusteen \tilde{Y} keskineliövirhematriisi ($= \mathbf{E}[(Y - \tilde{Y})^2]$ reaaliarvoisessa tapauksessa) ja (perustelu jätetään tehtäväksi)

$$a' \text{MSE}(\tilde{Y}) a \geq a' \text{MSE}(\mathbf{E}(Y|X)) a$$

kaikilla dimensioltaan sopivilla vektoreilla a . Toisin sanoen, matriisi $\text{MSE}(\tilde{Y}) - \text{MSE}(\mathbf{E}(Y|X))$ on positiivisesti semidefiniitti eli symbolein $\text{MSE}(\tilde{Y}) - \text{MSE}(\mathbf{E}(Y|X)) \geq 0$. Valitsemalla $a = (0, \dots, 0, 1, 0, \dots, 0)$ nähdään, että ennusteen \tilde{Y} jokaisen komponentin keskineliövirhe on vähintään yhtäsuuri kuin ennusteella $\mathbf{E}(Y|X)$.

Sovelletaan nyt edellä esitettyä VAR(p)-prosessiin (2.7). Muodostetaan ensin ennuste muuttujalle y_{t+1} , kun prosessin koko edeltävä historia y_t, y_{t-1}, \dots tunnetaan.⁷ Prosessin määrittely-yhtälöstä (2.7) nähdään, että

$$\mathbf{E}_t(y_{t+1}) := \mathbf{E}(y_{t+1}|y_{t-j}, j \geq 0) = \mathbf{E}(y_{t+1}|y_t, \dots, y_{t-p+1}) =: y_t(1).$$

Yhtälöstä (2.11) ja virheiden riippumattomuudesta nähdään lisäksi, että $\varepsilon_t \perp\!\!\!\perp y_{t-j}$, $j > 0$. Käyttäen yleisesti merkintää $\mathbf{E}_t(\cdot) = \mathbf{E}(\cdot|y_{t-j}, j \geq 0)$ saadaan siten $\mathbf{E}_t(\varepsilon_{t+j}) = \mathbf{E}(\varepsilon_{t+j}) = 0$ ($j > 0$) ja

$$\begin{aligned} y_t(1) &= A_1 \mathbf{E}_t(y_t) + \dots + A_p \mathbf{E}_t(y_{t+1-p}) + \mathbf{E}_t(\varepsilon_{t+1}) \\ &= A_1 y_t + \dots + A_p y_{t+1-p}. \end{aligned}$$

Yleisemmin, kun $h \geq 1$, saadaan

$$\mathbf{E}_t(y_{t+h}) = A_1 \mathbf{E}_t(y_{t+h-1}) + \dots + A_p \mathbf{E}_t(y_{t+h-p})$$

eli toisin merkinnöin

$$y_t(h) = A_1 y_t(h-1) + \dots + A_p y_t(h-p),$$

jossa $y_t(h-j) = \mathbf{E}_t(y_{t+h-j})$ ja $y_t(h-j) = y_{t+h-j}$, kun $h \leq j$. Toisin sanoen, $y_t(h-j)$ on prosessin havaittu arvo y_{t+h-j} , kun sellainen on ennustejankohtana t tiedossa, ja muulloin ajankohtana t tehty ennuste $\mathbf{E}_t(y_{t+h-j})$.

Käyttäen mallin lineaarista esitystä (2.11) saadaan ennusteelle vaihtoehtoinen esitys

$$y_t(h) = \mathbf{E}_t(y_{t+h}) = \sum_{j=0}^{\infty} \Psi_j \mathbf{E}_t(\varepsilon_{t+h-j}) = \sum_{j=h}^{\infty} \Psi_j \varepsilon_{t+h-j},$$

⁷Vaihtoehtoisesti voitaisiin tarkastella tulevien arvojen y_{T+h} , $h \geq 1$, ennustamista, kun käytävissä on havaittu aikasarja y_1, \dots, y_T . VAR(p)-prosessin tapauksessa näillä kahdella tavalla ei ole eroa, kun $T > p$.

josta voidaan johtaa ennustevirheelle yhtälö

$$y_{t+h} - y_t(h) = \sum_{j=0}^{h-1} \Psi_j \varepsilon_{t+h-j} \quad (\Psi_0 = I_n).$$

Tästä nähdään, että ennustevirheen odotusarvo on nolla eli ennuste on harhaton ja kovarianssimatriisi eli keskineliövirhematriisi on $\sum_{j=0}^{h-1} \Psi_j \Omega \Psi_j'$. Normaaliossa $\varepsilon_t \sim \text{nid}(0, \Omega)$ saadaan ennustevirheelle siten tulos

$$y_{t+h} - y_t(h) \sim \mathbf{N} \left(0, \sum_{j=0}^{h-1} \Psi_j \Omega \Psi_j' \right),$$

jonka avulla voidaan muodostaa luottamusvälejä ennustettavan y_{t+h} komponenteille. Jos $\sigma_i^2(h)$ on matriisin $\sum_{j=0}^{h-1} \Psi_j \Omega \Psi_j'$ i . diagonaali-alkio, niin ilmeisin merkinnöin

$$(y_{i,t+h} - y_{i,t}(h)) / \sigma_i(h) \sim \mathbf{N}(0, 1)$$

ja

$$\begin{aligned} 0.95 &= \mathbf{P} \left\{ -1.96 \leq \frac{y_{i,t+h} - y_{i,t}(h)}{\sigma_i(h)} \leq 1.96 \right\} \\ &= \mathbf{P} \{ y_{i,t}(h) - 1.96\sigma_i(h) \leq y_{i,t+h} \leq y_{i,t}(h) + 1.96\sigma_i(h) \}. \end{aligned}$$

Toisin sanoen, 95%:n luottamusväli ennustettavan muuttujan y_{t+h} i :nnelle komponentille on $y_{i,t}(h) \pm 1.96\sigma_i(h)$.

Käytännössä tuntemattomat parametrit korvataan estimaateilla, jolloin edellä esitetty pätee vain likimäärisesti. Huomaa myös, että esitettyä luottamusväliä johdettaessa indeksit i ja h oletettiin kiinnitettyiksi. Tarkasteltaessa useita tällaisia luottamusvälejä joudutaan ”tavanomaisiin” yhdistettyjen luottamusvälien ongelmiin.

VAR(p)–prosessi, kun $\mathbf{E}(y_t) \neq 0$. Käytännössä on yleensä epärealistista olettaa, että havaitut aikasarjat tuottaneen prosessin odotusarvo olisi nolla. Kaikki VAR(p)–prosessista edellä esitetyt tulokset voidaan yleistää helposti tapaukseen $\mathbf{E}(y_t) \neq 0$. Prosessin y_t paikalle asetetaan vain sen keskistetty version $y_t - \mu$, jolloin stationaarisuusehdon voimassa ollessa $\mu = \mathbf{E}(y_t)$. Koska $A(\mathbf{B})(y_t - \mu) = A(\mathbf{B})y_t - A(1)\mu$, voidaan vaihtoehtoisesti tarkastella prosessia

$$y_t = \nu + A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \Omega) \quad (t \in \mathbb{Z}),$$

jossa $\nu = A(1)\mu$ ja muut merkinnät ovat kuten aikaisemmin. Kuten stationaarisuusehdosta (2.13) nähdään, on matriisi $A(1)$ epäsingulaarinen, joten odotusarvon μ ja vakiotermin ν välinen suhde voidaan kirjoittaa myös $\mu = A(1)^{-1}\nu$.

2.4 VARMA(p,q)–prosessi

Yksiulotteisessa tapauksessa AR(p)–prosessi voidaan yleistää ARMA(p,q)–prosessiksi lisäämällä yhtälön (2.6) oikealle puolelle innovaatioiden $\varepsilon_{t-1}, \dots, \varepsilon_{t-q}$ lineaarikombinaatio. Vastaava moniulotteinen yleistys on

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t - M_1 \varepsilon_{t-1} - \dots - M_q \varepsilon_{t-q}, \quad \varepsilon_t \sim \text{iid}(0, \Omega) \quad (t \in \mathbb{Z}), \quad (2.14)$$

jossa M_1, \dots, M_q ovat $n \times n$ kerroinmatriiseja ja muut merkinnät ovat kuten yhtälössä (2.7). Tässä yhteydessä käytetään nimitystä *VARMA(p,q)–prosessi* tai *VARMA(p,q)–malli*.

Samaan tapaan kuin yksiulotteisessa tapauksessa voidaan todeta, että VARMA(p,q)–prosessin riittävä stationaarisuusehto on (2.9) tai vaihtoehtoisesti (2.13) eli sama kuin VAR(p)–prosessilla. Käyttäen viivästysoperaattoria voidaan yhtälö (2.14) kirjoittaa

$$A(\mathbf{B}) y_t = M(\mathbf{B}) \varepsilon_t,$$

jossa $A(\mathbf{B}) = I_n - A_1 \mathbf{B} - \dots - A_p \mathbf{B}^p$ ja $M(\mathbf{B}) = I_n - M_1 \mathbf{B} - \dots - M_q \mathbf{B}^q$. Stationaarisuusehdon voimassa ollessa voidaan tässä yhtälössä kertoa vasemmalta käänteismatriisilla $A(\mathbf{B})^{-1}$ ja saada ratkaisuksi hyvin määritelty lineaarinen prosessi

$$y_t = \Psi(\mathbf{B}) \varepsilon_t,$$

jossa nyt $\Psi(\mathbf{B}) = \sum_{j=0}^{\infty} \Psi_j \mathbf{B}^j = A(\mathbf{B})^{-1} M(\mathbf{B})$ ja $\Psi_0 = I_n$. Kuten VAR(p)–prosessin tapauksessakin, $\Psi_j \rightarrow 0$ geometrisesti, kun $j \rightarrow \infty$, ja siten myös VARMA(p,q)–prosessin kovarianssifunktio vaimenee geometrisesti nollaan viipymän kasvaessa rajatta.

VARMA(p,q)–prosessien teoria ja soveltaminen on huomattavasti hankalampaa kuin VAR(p)–prosessien. Tämä johtuu pitkälti polynomimatriisien ominaisuuksista, joiden vuoksi VARMA(p,q)–prosessien rakenneteoria ja siihen liittyvä mallinvalinta tulee monimutkaisemmaksi kuin VAR(p)–prosesseilla. Tämä näkyy erityisesti yksikäsitteisyys- tai identifioituvuusehdossa, joka yksiulotteisessa tapauksessa $n = 1$ vaatii, että (nyt) *polynomeilla* $A(z)$ ja $M(z)$ ei ole yhteisiä tekijöitä ja ainakin toinen korkeimman asteen kertoimista on nolasta poikkeava.

Seuraava esimerkki havainnollistaa miksi yksikäsitteisyyskysymys hankaloituu moniulotteisessa tapauksessa $n > 1$. Tarkastellaan kaksiulotteista VARMA(1,1)–prosessia

$$\begin{bmatrix} 1 & a\mathbf{B} \\ 0 & 1 \end{bmatrix} y_t = \begin{bmatrix} 1 & m\mathbf{B} \\ 0 & 1 \end{bmatrix} \varepsilon_t, \quad a \neq m. \quad (2.15)$$

Ensimmäisen asteen polynomimatriisit $A(\mathbf{B})$ ja $M(\mathbf{B})$ näyttäisivät toteuttavan yksiulotteisen ARMA(1,1)–prosessin yksikäsitteisyys ehdon kaltaisen ehdon. Kertomalla yhtälö (2.15) vasemmalta käänteismatriisilla

$$A(\mathbf{B})^{-1} = \begin{bmatrix} 1 & -a\mathbf{B} \\ 0 & 1 \end{bmatrix}$$

saadaan prosessin lineaarinen esitys

$$y_t = \begin{bmatrix} 1 & (m-a)\mathbf{B} \\ 0 & 1 \end{bmatrix} \varepsilon_t,$$

joka on vektoriarvoinen MA(1)–prosessi eli lyhyesti VMA(1)–prosessi (vastaavasti puhutaan MA(q)–prosessista, jossa $q = \infty$ on myös mahdollinen). Kertomalla yhtälö (2.15) vasemmalta käänteismatriisilla $M(\mathbf{B})^{-1}$ saadaan vastaavasti

$$\begin{bmatrix} 1 & (a-m)\mathbf{B} \\ 0 & 1 \end{bmatrix} y_t = \varepsilon_t$$

eli VAR(1)–prosessi. VARMA(1,1)–prosessi (2.15) voidaan siis lausua myös VAR(1)– ja VMA(1)–prosessina. Jos tällaisessa tilanteessa sovitetaan aineistoon VARMA(1,1)–prosessi, joutuu estimointimenetelmä hankaluuksiin, koska löytyy kolme eri mallia, jotka sopivat aineistoon yhtä hyvin. Ongelma ratkeaa tietysti, jos prosessin rakenteesta on tietoa niin, että estimoitavaksi voidaan valita joko VAR(1)– tai VMA(1)–malli. Yleensä tällaista tietoa ei kuitenkaan ole ja monimutkaisemmissa tapauksissa ongelman tunnistaminenkaan ei onnistu yhtä helposti kuin tässä tarkastellussa yksinkertaisessa esimerkissä.

VARMA(p,q)–prosesseille on kehitetty varsin täydellinen yksikäsitteisyysteoria, johon perustuen mallinvalintakin voidaan hoitaa käytännössä ns. mallinvalintakriteerejä käyttäen. Myös (normaalijakaumaan perustuva) suurimman uskottavuuden estimointiteoria on kehitetty. Näiden kysymysten tarkastelu sivuutetaan kuitenkin tällä kurssilla, koska niiden opiskelu on selvästi ”haastavampaa” kuin VAR(p)–prosesseilla. Esitettyjä VAR(p)–prosesseihin perustuvia tilastollisia menetelmiä voidaan perustella approksimaatioina, sillä käännettävyysehdon

$$\det(M(z)) \neq 0, \quad |z| \leq 1,$$

voimassa ollessa, VARMA(p,q)–prosessi voidaan kirjoittaa

$$M(\mathbf{B})^{-1} A(\mathbf{B}) y_t = \varepsilon_t,$$

jossa $M(\mathbf{B})^{-1} A(\mathbf{B}) = \sum_{j=0}^{\infty} \Pi_j \mathbf{B}^j$ ja kerroinmatriisit Π_j konvergoivat nolnaan geometrisesti, kun $j \rightarrow \infty$ ($\Pi_0 = I_n$). VARMA(p,q)–prosessia voidaan siten approksimoida VAR(h)–prosessilla, kun aste h on ”tarpeeksi suuri”.

3 Stationaarisen prosessin ensimmäisten ja toisten momenttien estimointi

Otoskeskiarvon tarkentuvuus. Olkoon $y_t = (y_{1t}, \dots, y_{nt})$ stationaarinen prosessi odotusarvona $E(y_t) = \mu$ ja kovarianssifunktiona $\text{Cov}(y_t, y_{t+k}) = \Gamma_k = [\gamma_{ij,k}]$. Oletetaan, että (vrt. ehto (2.1))

$$\sum_{k=-\infty}^{\infty} \|\Gamma_k\| < \infty. \quad (3.1)$$

Tarkastellaan otoskeskiarvoa $\bar{y} = T^{-1} \sum_{t=1}^T y_t$ odotusarvon μ estimaattorina. Harhattomuus on ilmeisen, sillä

$$\mathbf{E}(\bar{y}) = T^{-1} \sum_{t=1}^T \mathbf{E}(y_t) = \mu.$$

Kun merkitään $\bar{y} = (\bar{y}_1, \dots, \bar{y}_n)$ ja $\mu = (\mu_1, \dots, \mu_n)$ saadaan lisäksi

$$\begin{aligned} \text{Var}(\bar{y}_i) &= \mathbf{E} \left[\left(T^{-1} \sum_{t=1}^T (y_{it} - \mu_i) \right)^2 \right] \\ &= T^{-2} \sum_{t=1}^T \sum_{s=1}^T \mathbf{E} [(y_{it} - \mu_i)(y_{is} - \mu_i)] \\ &= T^{-1} \sum_{t-s=-T}^T (T - |t-s|) \gamma_{i,t-s} \\ &= T^{-1} \sum_{k=-T}^T \left(1 - \frac{|k|}{T} \right) \gamma_{i,k}. \end{aligned}$$

Tässä kolmannen yhtälön voi todeta huomaamalla, että edeltävä kaksoissumma on matriisin $[\gamma_{i,t-s}]_{t,s=1,\dots,T}$ alkioiden summa. Kolmioepäyhtälön ja oletuksen (3.1) avulla saadaan tulos

$$\text{Var}(\bar{y}_i) = \mathbf{E} [(\bar{y}_i - \mu_i)^2] \leq T^{-1} \sum_{k=-T}^T \left(1 - \frac{|k|}{T} \right) |\gamma_{i,k}| \rightarrow 0, \quad \text{kun } T \rightarrow \infty.$$

Yhdistämällä edellä todettu voidaan päätellä (Markovin epäyhtälöä käyttäen), että otoskeskiarvo \bar{y}_i on odotusarvon μ_i tarkentuva estimaattori ja edelleen, että sama pätee otoskeskiarvolle \bar{y} . Toisin sanoen, on todettu stokastinen konvergenssi

$$\bar{y} \xrightarrow{p} \mu. \quad (3.2)$$

Oletus (3.1) on voimassa erityisesti VARMA(p,q)-prosesseille, koska tällöin $\Gamma_k \rightarrow 0$ geometrisesti, kun $|k| \rightarrow \infty$.

Otoskovarianssifunktion tarkentuvuus. Tarkastellaan nyt otoskovarianssifunktiota

$$C_k = [c_{ij,k}] = (T-k)^{-1} \sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})', \quad k = 0, \dots, T-1,$$

$C_k = C'_{-k}$, $k = -T+1, \dots, -1$. Kiinnitetään indeksit i, j ja $k \geq 0$ ja määritellään yksinkertaisuuden vuoksi muuttuja $x_t = y_{it}y_{j,t+k}$, joka on keskeinen osa estimaattoria

$c_{ij,k}$. Oletetaan $\mathbf{E}(y_{it}^4) < \infty$ ($i = 1, \dots, n$), jolloin x_t on stationaarinen ja sillä on äärellinen toinen momentti. Jos oletetaan, että x_t :n autokovarianssifunktio toteuttaa ehdon (3.1) vastineen, saadaan edellä otoskeskiarvolle todetun perusteella tulos

$$\bar{x} = (T - k)^{-1} \sum_{t=1}^T y_{it} y_{j,t+k} \xrightarrow{p} \mathbf{E}(y_{it} y_{j,t+k}),$$

jossa kovarianssin määritelmän perusteella

$$\mathbf{E}(y_{it} y_{j,t+k}) = \gamma_{ij,k} + \mathbf{E}(y_{it}) \mathbf{E}(y_{j,t+k}) = \gamma_{ij,k} + \mu_i \mu_j.$$

Olettaen, että edellä tehdyt oletukset pätevät kaikilla i, j ja $k \geq 0$ voidaan estimaattorin \bar{x} tarkentuvuudesta ja tuloksesta (3.2) johtaa edelleen tulokset $c_{ij,k} \xrightarrow{p} \gamma_{ij,k}$ ja

$$C_k \xrightarrow{p} \Gamma_k. \quad (3.3)$$

Yksityiskohdat, joissa tarvitaan lisäksi stokastisen konvergenssin jatkuvan kuvauksen lausetta (ks. Liite B, Lause B.1), jätetään tehtäväksi.

Voidaan siis todeta, että kohtuullisin oletuksin otoskovarianssifunktio on tarkentuva. (Huomaa kuitenkin, että edellä k on kiinteä.) Tarvittavat oletukset on helppo tarkistaa erikoistapauksessa $y_t \sim \text{iid}(\mu, \Sigma)$ ja sen jälkeen edelleen VMA(q)-prosessin tapauksessa ($q < \infty$). Yleisillä VARMA(p, q)-prosesseilla tarkentuvuustulos (3.3) voidaan perustella käyttäen approksimaatiota VARMA(p, q) \approx VMA(h), kun h on ”suuri”. Yksityiskohdat ovat kuitenkin hieman teknisiä ja sivuutetaan.

Asymptoottinen normaalisuus. Edellä esitetyt tulokset havainnollistavat sitä, että stationaarisilla prosesseilla suurten lukujen lait pätevät ”varsin yleisin oletuksin”. Seuraavassa yritetään havainnollistaa sitä, että sama pätee keskeiselle raja-arvolauseelle.

Tarkastellaan yksinkertaisuuden vuoksi (n -ulotteista) lineaarista prosessia

$$y_t = \mu + \Psi(\mathbf{B}) \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \Omega), \quad (3.4)$$

jossa $\Psi(\mathbf{B}) = \sum_{j=0}^{\infty} \Psi_j \mathbf{B}^j$ ja $\Psi_0 = I_n$. Oletetaan lisäksi vahvemmin kuin jaksossa 2.2 eli

$$\sum_{j=1}^{\infty} j \|\Psi_j\| < \infty.$$

Tällöin voidaan kirjoittaa $\Psi(\mathbf{B}) = \Psi(1) + \Delta G(\mathbf{B})$, jossa $\Delta = 1 - \mathbf{B}$ on differensointioperaattori ja $G(\mathbf{B}) = \sum_{j=0}^{\infty} G_j \mathbf{B}^j$ toteuttaa ehdon $\sum_{j=0}^{\infty} \|G_j\| < \infty$ ja siten $\sum_{j=0}^{\infty} \|G_j\|^2 < \infty$ (ks. Liite A. 4). Soveltamalla tätä yhtälössä (3.4) saadaan

$$y_t - \mu = \Psi(1) \varepsilon_t + \Delta u_t, \quad u_t = G(\mathbf{B}) \varepsilon_t.$$

Edellä todetun nojalla u_t on lineaarinen prosessi ja sen komponenteille pätee $\mathbf{E}(u_{it}^2) < \infty$ (ks. jakso 2.2). Otoskeskiarvolle \bar{y} saadaan siten

$$\begin{aligned}\sqrt{T}(\bar{y} - \mu) &= \Psi(1) \frac{1}{\sqrt{T}} \sum_{t=1}^T \varepsilon_t + \frac{1}{\sqrt{T}} \sum_{t=1}^T (u_t - u_{t-1}) \\ &= \Psi(1) \frac{1}{\sqrt{T}} \sum_{t=1}^T \varepsilon_t + \frac{1}{\sqrt{T}} (u_T - u_0).\end{aligned}$$

Koska $\varepsilon_t \sim \text{iid}(0, \Omega)$, niin tavanomaisen (moniulotteisen) keskeisen raja-arvolauseen nojalla pätee jakaumakonvergenssi $T^{-1/2} \sum_{t=1}^T \varepsilon_t \xrightarrow{d} \mathbf{N}(0, \Omega)$. Toisaalta, koska $\mathbf{E}(u_{it}) = 0$ ja $\mathbf{E}(u_{it}^2) \leq c < \infty$ kaikilla $i = 1, \dots, n$ ja kaikilla $t \in \mathbb{Z}$, saadaan Markovin epäyhtälöä käyttäen $T^{-1/2} u_T \xrightarrow{p} 0$ ja $T^{-1/2} u_0 \xrightarrow{p} 0$ eli $T^{-1/2}(u_T - u_0) \xrightarrow{p} 0$. Käyttäen tunnettua jakaumakonvergenssia ja stokastista konvergenssia koskevaa jatkuvan kuvauksen lausetta (ks. Liite B, Lause B.3) voidaan kaiken kaikkiaan päätellä, että sv:lla $\sqrt{T}(\bar{y} - \mu)$ ja $\Psi(1) T^{-1/2} \sum_{t=1}^T \varepsilon_t$ on sama asymptoottinen jakauma ja

$$\sqrt{T}(\bar{y} - \mu) \xrightarrow{d} \mathbf{N}(0, \Psi(1) \Omega \Psi(1)') . \quad (3.5)$$

Otoskeskiarvon keskeinen raja-arvolause voidaan osoittaa myös ilman lineaarisuusoletusta (3.4), mutta yhtä kätevää perustelua kuin edellä ei ole helppo esittää. Olennainen vaatimus useissa vastaavissa tuloksissa on, että sv:t y_t ja y_{t-k} ”tulevat riippumattomiksi”, kun $k \rightarrow \infty$. Lineaarisen prosessin (3.4) tapauksessa y_{t-k} riippuu virhetermeistä $\varepsilon_{t-k}, \varepsilon_{t-k-1}, \dots$ ja y_t riippuu näistä (ja siten y_{t-k} :sta) vain muunnosten $\Psi_k \varepsilon_{t-k}, \Psi_{k+1} \varepsilon_{t-k-1}, \dots$ kautta. Koska $\Psi_k \rightarrow 0$, kun $k \rightarrow \infty$, ”häviää” sv:ien y_t ja y_{t-k} riippuvuus k :n kasvaessa.

Tuloksen (3.5) perustelussa käytettyä tekniikkaa voidaan soveltaa myös otoskovarianssifunktion C_k tapauksessa, mutta algebralliset laskelmat tulevat ikäviksi. Koska saatava tulos on lisäksi monimutkainen erikoistapauksia (lähinnä VMA(q)- ja iid-tapausta) lukuun ottamatta, ei sitä käsitellä. Käytännössä kovarianssit korvataan yleensä vastaavilla korrelaatioilla eli käytetään otossuureita

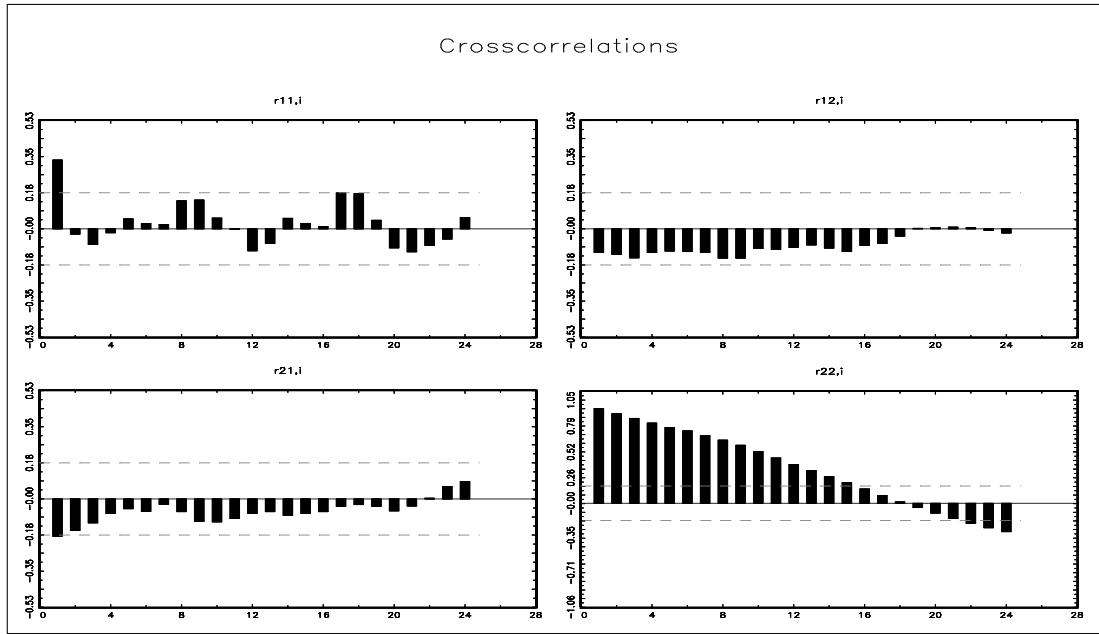
$$r_{ab,k} = \frac{c_{ab,k}}{\sqrt{c_{a,0}c_{b,0}}}, \quad c_{a,0} = c_{aa,0}.$$

Voidaan osoittaa, että tapauksessa $y_t \sim \text{iid}(\mu, \Sigma)$ pätee asymptoottinen tulos

$$r_{ab,k} \underset{as}{\sim} \mathbf{N}(0, 1/T) . \quad (3.6)$$

Lisäksi, jos $R_k = [r_{ab,k}]$, ovat R_k ja R_l asymptoottisesti riippumattomia, kun $k \neq l$. Tämä ei kuitenkaan päde yleisesti estimaattorin R_k alkioille.

Esimerkki 3.1. Kuviossa 3.1 nähdään Kuvion 1.1 valuuttakurssi/korkoaineiston kahdesta aikasarjasta lasketut auto- ja ristikorrelaatiofunktiot. Kuvioon on piirretty



Kuvio 3.1. Euron ja Yhdysvaltojen dollarin välisen kuukausittaisen vaihtokurssin muutoksen (y_{1t}) ja euroalueen ja Yhdysvaltojen 10 vuoden valtioiden obligaatioiden keskikorkojen eron (y_{2t}) estimoidut autokorrelaatio- ja ristikorrelaatiofunktiot ($t = 1999\text{I}, \dots, 2009\text{XI}$, $T = 131$).

myös kriittiset rajat $\pm 2/\sqrt{T}$, joiden sisään yksittäisen korrelaation pitäisi jäädä likimain 95%:n todennäköisyydellä, mikäli kysymyksessä olisi iid-prosessista saatu reaalisatio. Estimoidut autokorrelaatiot vahvistavat jo Kuvioista 1.1 nähtävän selvän autokorrelotuneisuuden. Estimoidut ristikorrelaatiot näyttävät sen sijaan pieniltä. Huomaa kuitenkin, että komponenttisarjojen autokorrelotuneisuuden vuoksi kuvion kriittisiä rajoja ei voida sovaltaa ristikorrelaatioestimaatteihin edellä kuvatulla tavalla.

4 Stationaarisen VAR-mallin teoriaa

4.1 Parametrien rajoittamaton estimointi

Mallin määrittely. Oletetaan, että havaittuun $n:n$ aikasarjan aineistoon voidaan soveltaa VAR(p)-mallia

$$y_t = \nu + A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t, \quad t = 1, \dots, T, \quad (4.1)$$

jossa alkuarvot y_{-p+1}, \dots, y_0 on myös havaittu ja kerroinmatriisit A_1, \dots, A_p ($n \times n$) toteuttavat stationaarisuus- tai stabiilisuusehdon (2.9) (tai yhtäpitävästi ehdon (2.13)). Jotta parametrien estimoinnissa ja testauksessa päästäisiin soveltamaan suurimman uskottavuuden (SU) teoriaa, liitetään virhetermiin normaalisuusoletus

$$\varepsilon_t \sim \text{nid}(0, \Omega), \quad (4.2)$$

jossa kovarianssimatriisi Ω on positiivisesti definiitti. Todettakoon kuitenkin, että esitettävät asymptoottiset tulokset pätevät yleisemmin oletuksella $\varepsilon_t \sim \text{iid}(0, \Omega)$.

Uskottavuusfunktion johto. Otetaan käyttöön merkinnät

$$x_t = [1 \ y'_{t-1} \ \cdots \ y'_{t-p}]' = [1 \ \mathbf{y}'_{t-1}]' \quad ((np+1) \times 1)$$

ja

$$\Pi = \begin{bmatrix} \pi'_1 \\ \vdots \\ \pi'_n \end{bmatrix} = [\nu : A_1 : \cdots : A_p] \quad (n \times (np+1)).$$

Tällöin yhtälö (4.1) voidaan kirjoittaa

$$y_t = \Pi x_t + \varepsilon_t, \quad t = 1, \dots, T. \quad (4.3)$$

Tämä on yhtälöittäin esitettynä ($\pi'_i x_t = x'_t \pi_i$)

$$y_{it} = x'_t \pi_i + \varepsilon_{it}, \quad i = 1, \dots, n, \quad t = 1, \dots, T, \quad (4.4)$$

mistä nähdään, että yhtälö (4.3) voidaan kirjoittaa

$$y_t = X'_t \pi + \varepsilon_t, \quad t = 1, \dots, T, \quad (4.5)$$

jossa $\pi = [\pi'_1 \ \cdots \ \pi'_n]'$ ($n(np+1) \times 1$) ja

$$X'_t = \text{diag}[x'_t \ \cdots \ x'_t] = I_n \otimes x'_t \quad (n \times n(np+1)).$$

Uskottavuusfunktion johtoa varten merkitään $\mathbf{Y}_t = [y'_{-p+1} \ \cdots \ y'_0 \ y'_1 \ \cdots \ y'_t]'$, $t = 0, 1, \dots, T$. Jättäen argumentit yksinkertaisuuden vuoksi pois merkitään satunnaisvektorin \mathbf{Y}_t yhteistiheysfunktioita symbolilla $f_{\mathbf{Y}_t}$. Tällöin koko aineistoa vastaavan sv:n \mathbf{Y}_T yhteistiheysfunktio on $f_{\mathbf{Y}_T}$ ja käyttäen ehdollisen tiheysfunktion kaavaa saadaan

$$f_{\mathbf{Y}_T} = f_{y_T | \mathbf{Y}_{T-1}} \cdot f_{\mathbf{Y}_{T-1}} = f_{y_T | \mathbf{Y}_{T-1}} \cdot f_{y_{T-1} | \mathbf{Y}_{T-2}} \cdot f_{\mathbf{Y}_{T-2}} = \cdots = \prod_{t=1}^T f_{y_t | \mathbf{Y}_{t-1}} \cdot f_{\mathbf{Y}_0}.$$

Jaksossa 2.3 esitetyistä VAR(p)-prosessin lineaarisesta esityksestä (2.11) seuraa riippumattomuus $y_{t-j} \perp\!\!\!\perp \varepsilon_t$, $j \geq 1$, joten erityisesti $\mathbf{Y}_{t-1} \perp\!\!\!\perp \varepsilon_t$. Tämän ja malliyhtälön (4.1) perusteella on selvää, että sv:n y_t ehdollinen jakauma ehdolla \mathbf{Y}_{t-1} riippuu vain \mathbf{Y}_{t-1} :n p :stä ensimmäisestä osavektorista y_{t-1}, \dots, y_{t-p} eli sv:sta \mathbf{y}_{t-1} . Käyttäen malliyhtälön (4.1) esitystä (4.5) saadaan siten⁸

$$y_t | \mathbf{Y}_{t-1} \sim y_t | \mathbf{y}_{t-1} \sim \mathbf{N}(X'_t \pi, \Omega), \quad t = 1, \dots, T.$$

⁸Tässä voi käyttää seuraavaa multinormaalijakaumaan liittyvää tulosta. Jos $Y = X + Z$, jossa $X \perp\!\!\!\perp Z$ ja $Z \sim \mathbf{N}(\mu, \Sigma)$, niin $Y | (X = x) \sim \mathbf{N}(\mu + x, \Sigma)$ (perustelu jätetään tehtäväksi).

Multinormaalijakauman tiheysfunktion lausekkeesta saadaan näin ollen

$$f_{y_t|\mathbf{Y}_{t-1}}(y_t|\mathbf{y}_{t-1}) = (2\pi)^{-1/2} \det(\Omega)^{-1/2} \exp\left\{-\frac{1}{2}(y_t - X_t'\pi)' \Omega^{-1} (y_t - X_t'\pi)\right\}. \quad (4.6)$$

Koko aineistoa vastaavan sv:n \mathbf{Y}_T tiheysfunktion $f_{\mathbf{Y}_T}$ selvittämiseksi on siten vielä selvitettävä alkuarvovektorin $\mathbf{Y}_0 = \mathbf{y}_0$ jakauma. Stationaarisessa tapauksessa oletuksesta (4.2) seuraa, että \mathbf{y}_0 on (normaalisti jakautuneiden sv:ien lineaarisena muunnoksena) normaalisti jakautunut. Odotusarvo saadaan tuloksesta $\mathbf{E}(y_t) = A(1)^{-1}\nu$ (ks. jakson 2.3 loppu) ja kovarianssimatriisi on $\mathbf{\Gamma}_0 = [\Gamma_{i-j}]$ ($i, j = 1, \dots, p$), joka voidaan lausua parametrien ν, A_1, \dots, A_p ja Ω funktiona (ks. jakso 2.3).

Seuraavassa alkuarvo \mathbf{y}_0 tulkitaan kiinteäksi vakioksi tai yhtäpitävästi ehdollistaa sen saaman arvon suhteen. Tällöin päädytään ns. *ehdolliseen uskottavuusfunktioon*, joka saadaan edellä todetun mukaan tulkitsemalla tulo $f_{y_1|\mathbf{Y}_0} \cdots f_{y_T|\mathbf{Y}_{T-1}}$ parametrien funktiona. Käyttäen yhtälöä (4.6) ja ottamalla logaritmi päädytään siten ehdolliseen log-uskottavuusfunktioon (parametrin riippumaton vakio jätetään pois)

$$l(\pi, \Omega) = -\frac{T}{2} \log \det(\Omega) - \frac{1}{2} \sum_{t=1}^T (y_t - X_t'\pi)' \Omega^{-1} (y_t - X_t'\pi). \quad (4.7)$$

On selvää, että suurilla havaintomäärillä alkuarvojen y_{-p+1}, \dots, y_0 huomiotta jättämisellä on mitätön vaikutus ja uskottavuusfunktioon perustuvien tilastollisten menetelmien asympotoottisiin ominaisuuksiin sillä ei ole mitään vaikutusta. Ellei toisin mainita, tarkoitetaan uskottavuusfunktioilla seuraavassa ehdollista uskottavuusfunktiota.

Uskottavuusfunktion maksimointi. Maksimoidaan nyt uskottavuusfunktio tavantomaista derivointimenettelyä käyttäen. Derivoimalla saadaan yhtälöstä (4.7) (ks. Liite A.7)

$$\frac{\partial}{\partial \pi} l(\pi, \Omega) = \sum_{t=1}^T X_t \Omega^{-1} (y_t - X_t'\pi). \quad (4.8)$$

Matriisin X_t määritelmän ja Kroneckerin tulon ominaisuuksia käyttäen saadaan (ks. Liite A.5)

$$X_t \Omega^{-1} = (I_n \otimes x_t) (\Omega^{-1} \otimes 1) = \Omega^{-1} \otimes x_t = (\Omega^{-1} \otimes I_{np+1}) (I_n \otimes x_t),$$

jossa siis $(I_n \otimes x_t) = X_t$. Yhtälö (4.8) voidaan näin ollen kirjoittaa

$$\frac{\partial}{\partial \pi} l(\pi, \Omega) = (\Omega^{-1} \otimes I_{np+1}) \sum_{t=1}^T X_t (y_t - X_t'\pi). \quad (4.9)$$

Asettamalla $\partial l(\pi, \Omega) / \partial \pi = 0$ voidaan (epäsingulaarinen) tekijä $(\Omega^{-1} \otimes I_{np+1})$ supistaa pois, joten parametrin π uskottavuusyhtälöksi saadaan

$$\sum_{t=1}^T X_t X_t' \pi = \sum_{t=1}^T X_t y_t,$$

jonka yksikäsitteinen ratkaisu on

$$\hat{\pi} = \left(\sum_{t=1}^T X_t X_t' \right)^{-1} \sum_{t=1}^T X_t y_t$$

eli malliyhtälöstä (4.5) muodostettu pienimmän neliösumman (PNS) estimaattori. Viimeksi mainittu seikka tulee havainnollisemmaksi, kun yhtälöryhmä (4.5) kirjoitetaan matriisimuodossa

$$\begin{bmatrix} y_1 \\ \vdots \\ y_T \end{bmatrix} = \begin{bmatrix} X_1' \\ \vdots \\ X_T' \end{bmatrix} \pi + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{bmatrix}$$

tai lyhyesti

$$y = X' \pi + \varepsilon.$$

Suoraviivaisella matriisilaskulla nähdään, että $\hat{\pi}$ saadaan tavanomaisella PNS-kaavalla

$$\hat{\pi} = (X' X)^{-1} X' y.$$

Huomaa kuitenkin, että $\text{Cov}(\varepsilon) = I_T \otimes \Omega$, mikä poikkeaa tavanomaisesta lineaarisen mallin yhteydessä tehtävästä oletuksesta. Syy sille, että SU-menetelmä johtaa tästä huolimatta tavanomaiseen PNS-estimointiin piilee selittäjämatrisin X erityisessä rakenteessa. Myöhemmin nähdään, että vastaava tulos ei päde, kun matriisin X rakenne muuttuu. Suoralla (tehtäväksi jätettävällä) laskulla nähdään, että estimaattorivektorin $\hat{\pi}$ osavektorit $\hat{\pi}_1, \dots, \hat{\pi}_n$ saadaan soveltamalla PNS-menetelmää yhtälöryhmän (4.4) yksittäisiin yhtälöihin.

Tarkastellaan nyt kovarianssimatriisin Ω SU-estimointia. Tässä yhteydessä on kätevää johtaa käänteisematriisin $\Omega^{-1} =: \Phi$ SU-estimaattori, minkä jälkeen Ω :n SU-estimaattori saadaan SU-menetelmän invarianssiominaisuuden nojalla kääntämällä Φ :n SU-estimaattori. Käyttäen parametrimatriisia Φ ja lyhennysmerkintää $\varepsilon_t(\pi) = y_t - X_t' \pi$ voidaan log-uskottavuusfunktio kirjoittaa

$$l(\pi, \Phi) = \frac{T}{2} \log \det(\Phi) - \frac{1}{2} \sum_{t=1}^T \varepsilon_t(\pi)' \Phi \varepsilon_t(\pi).$$

Liitteessä A.7 esitettyjen derivointikaavojen mukaan on

$$\frac{\partial}{\partial \Phi} \varepsilon_t(\pi)' \Phi \varepsilon_t(\pi) = \varepsilon_t(\pi) \varepsilon_t(\pi)' \quad \text{ja} \quad \frac{\partial}{\partial \Phi} \log \det(\Phi) = \Phi^{-1} = \Omega,$$

joten

$$\frac{\partial}{\partial \Phi} l(\pi, \Phi) = \frac{T}{2} \Omega - \frac{1}{2} \sum_{t=1}^T \varepsilon_t(\pi) \varepsilon_t(\pi)'$$

Asettamalla oikea puoli nolaksi ja korvaamalla vektorissa $\varepsilon_t(\pi)$ parametri edellä johdetulla π :n uskottavuusyhtälön ratkaisulla $\hat{\pi}$ saadaan ratkaisu

$$\hat{\Omega} = \frac{1}{T} \sum_{t=1}^T (y_t - X_t' \hat{\pi})(y_t - X_t' \hat{\pi})'.$$

Koska $\hat{\pi}$ minimoi PNS-estimaattorina log-uskottavuusfunktion $l(\pi, \Omega)$ summalausekkeen (ks. (4.7)), on SU-menetelmän invarianssiominaisuuden nojalla selvää, että $\hat{\pi}$ ja $\hat{\Omega}$ ovat parametrien π ja Ω (ehdollisia) SU-estimaattoreita.⁹

SU-estimaattorien asymptoottiset ominaisuudet. SU-estimaattorien asymptoottisten ominaisuuksien selvittämisessä nojaututaan tavanomaiseen SU-teoriaan, jonka toimiminen perustuu siihen, että tavanomaiset raja-arvolauseet (eli suurten lukujen laki ja keskeinen raja-arvolause) pätevät tarkasteltavan VAR(p)-mallin tapauksessa. Oletetaan alkuarvot y_{-p+1}, \dots, y_0 yksinkertaisuuden vuoksi stationaarisiksi.

Estimaattorin $\hat{\pi}$ asymptoottisen jakauman johtamiseksi lasketaan ensin (ks. (4.9) ja Liite A.7)

$$\frac{\partial^2}{\partial \pi \partial \pi'} l(\pi, \Omega) = -(\Omega^{-1} \otimes I_{np+1}) \sum_{t=1}^T X_t X_t' = -\Omega^{-1} \otimes \sum_{t=1}^T x_t x_t',$$

jossa jälkimmäinen yhtälö perustuu matriisiin X_t esitykseen $X_t = I_n \otimes x_t$ ja Kroneckerin tulon ominaisuuksiin (ks. Liite A.5). Koska $x_t = [1 \ \mathbf{y}'_{t-1}]'$, on parametrin π Fisherin informaatiomatriisi siten

$$\mathcal{I}_{\pi\pi}(\pi, \Omega) := -\mathbf{E} \left[\frac{\partial^2}{\partial \pi \partial \pi'} l(\pi, \Omega) \right] = T(\Omega^{-1} \otimes \mathbf{\Gamma}_x),$$

jossa

$$\mathbf{\Gamma}_x = \mathbf{E}(x_t x_t') = \begin{bmatrix} 1 & \mathbf{E}(\mathbf{y}'_{t-1}) \\ \mathbf{E}(\mathbf{y}_{t-1}) & \mathbf{E}(\mathbf{y}_{t-1} \mathbf{y}'_{t-1}) \end{bmatrix}.$$

Todetaan seuraavaksi, että (ks. (4.8) ja Liite A.7)

$$\begin{aligned} \mathbf{E} \left[\frac{\partial^2}{\partial \pi \partial \omega_{ij}} l(\pi, \Omega) \right] &= \sum_{t=1}^T \mathbf{E} \left[X_t \left(\frac{\partial}{\partial \omega_{ij}} \Omega^{-1} \right) (y_t - X_t' \pi) \right] \\ &= \sum_{t=1}^T \mathbf{E} \left[X_t \left(\frac{\partial}{\partial \omega_{ij}} \Omega^{-1} \right) \right] \mathbf{E}(\varepsilon_t) \\ &= 0. \end{aligned}$$

⁹Mainittakoon, että edellä esitettyä menettelyä soveltaen voidaan johtaa SU-estimaattorit myös usean yhtälön regressiomallissa eli mallissa, jonka määrittelee yhtälö (4.3), kun X_t on kiinteä (tai voidaan tulkita kiinteäksi) ja virhetermi toteuttaa normaalisuusoletuksen (4.2). Erikoistapaus tästä on malli $Y_1, \dots, Y_T \sim \mathbf{N}(\mu, \Sigma)$ ||.

Tässä toinen yhtälö perustuu yhtälöön $\varepsilon_t = y_t - X_t' \pi$ ja siihen, että ε_t on riippumaton muuttujista y_{t-1}, y_{t-2}, \dots ja siten muuttujasta X_t (ks. X_t :n määritelmä yhtälössä (4.5)).

Edellä todetusta seuraa, että parametrien π ja Ω Fisherin informaationmatriisi on lohkodeagonaalinen eli parametrit π ja Ω ovat ortogonaaliset. Yleisestä SU-estimointiteoriasta voidaan siten päätellä, että estimaattorit $\hat{\pi}$ ja $\hat{\Omega}$ ovat asympotoottisesti riippumattomat ja estimaattorin $\hat{\pi}$ asympotoottinen jakauma on sama kuin tapauksessa, jossa parametrin Ω arvo tunnetaisiin. Toisin sanoen, $\hat{\pi}$ on asympotoottisesti normaalin odotusarvona π ja kovarianssimatriisina $\mathcal{I}_{\pi\pi}(\pi, \Omega)^{-1}$. Koska $(\Omega^{-1} \otimes \Gamma_x)^{-1} = \Omega \otimes \Gamma_x^{-1}$ saadaan matriisin $\mathcal{I}_{\pi\pi}(\pi, \Omega)$ määritelmän nojalla siten tulos

$$\hat{\pi} \underset{as}{\sim} \mathbf{N}(\pi, T^{-1}(\Omega \otimes \Gamma_x^{-1})).$$

Estimaattorin $\hat{\Omega}$ asympotoottista jakaumaa ei tarkastella. Myöhemmin vedotaan kuitenkin tarkentuvuuteen $\hat{\Omega} \xrightarrow{p} \Omega$.

Huom.: Edellä johdettu tulos pätee myös ilman oletusta stationaarisista alkuarvoista, kunhan matriisi $T(\Omega^{-1} \otimes \Gamma_x)$ korvataan Fisherin informaatiomatriisin yleisellä lausekkeella $\mathcal{I}_{\pi\pi}(\pi, \Omega) = \Omega^{-1} \otimes \sum_{t=1}^T \mathbf{E}(x_t x_t')$. Sama pätee seuraavassa esitettävään rajoitettuun estimointiin ja testimenetelmiin. On kuitenkin syytä huomata, että asympotoottisten jakauma-approksimaatioiden toimivuuteen vaikuttaa se, kuinka voimakasta prosessin korreloituneisuus on. Kun prosessin stabiilisuusehto (ks. jakso 2.3) on ”lähellä rikkoontua”, voivat asympotoottiset jakaumat olla pienillä havaintomäärillä harhaanjohtavia.

4.2 Estimointi lineaarisin rajoittein

Tarkastellaan edellisen jakson mallia, jossa parametreihin A_1, \dots, A_p (ja mahdollisesti myös ν) liitetään lineaarisia rajoitteita. Nämä rajoitteet on kätevä esittää käyttäen yhtälön (4.5) parametrivektoria π , jonka oletetaan toteuttavan

$$\pi = H\delta + a, \tag{4.10}$$

jossa H on tunnettu $n(np+1) \times m$ matriisi, a on tunnettu $n(np+1)$ vektori ja δ on tuntematon $m \times 1$ parametrivektori. Lisäksi oletetaan, että matriisi H on täyttä sarakeastetta eli $r(H) = m$

Usein lineaariset parametrirajoitteet esitetään vaihtoehtoisella tavalla käyttäen yhtälöä

$$R\pi = b, \tag{4.11}$$

jossa R ($q \times n(np+1)$) ja b ($q \times 1$) ovat tunnettuja ja $r(R) = q$. Lineaarialgebraa käyttäen rajoitteet (4.10) ja (4.11) voidaan todeta yhtäpitäviksi, kun $q + m = n(np+1)$ (perustelu jätetään tehtäväksi). Edellinen on kätevä tämän jakson estimointitarkasteluissa ja jälkimmäinen seuraavan jakson Waldin testissä.

Tyypillisin esimerkki edellä esitetystä rajoitteista rajoittaa osan parametrivektorin π komponenteista nolaksi. Jos esimerkiksi valitaan $H = [I_{n^2p} : 0]'$ ja $a = 0$, asetetaan π :n n viimeistä komponenttia nolaksi eli oletetaan, että y_{t-p} ei selitä y_t :n viimeistä komponenttia y_{nt} . Tällöin selvästikin $R = [0 : I_n]$ ja $b = 0$.

Rajoitetun mallin parametrien δ ja Ω log-uskottavuusfunktio saadaan yksinkertaisesti suorittamalla rajoittamattoman mallin log-uskottavuusfunktiossa (4.7) sijoitus $\pi \rightarrow H\delta + a$. Merkitsemällä

$$z_t = y_t - X_t' a \quad \text{ja} \quad W_t = H' X_t$$

saadaan log-uskottavuusfunktioksi

$$l(\delta, \Omega) = -\frac{T}{2} \log \det(\Omega) - \frac{1}{2} \sum_{t=1}^T (z_t - W_t' \delta)' \Omega^{-1} (z_t - W_t' \delta).$$

Uskottavuusyhtälöiden tutkimiseksi derivoidaan parametrien suhteen. Kuten derivaatan (4.8) tapauksessa saadaan

$$\frac{\partial}{\partial \delta} l(\delta, \Omega) = \sum_{t=1}^T W_t \Omega^{-1} (z_t - W_t' \delta). \quad (4.12)$$

Toisin kuin parametrin π tapauksessa (ks. yhtälöt (4.8) ja (4.9)) ei matriisia Ω^{-1} voida nyt kuitenkaan "siirtää" summan eteen. Tämä merkitsee, ettei yhtälöä $\partial l(\delta, \Omega) / \partial \delta = 0$ voida ratkaista tuntematta matriisia Ω , vaan ratkaisu jää (yleensä) muotoon

$$\delta = \left(\sum_{t=1}^T W_t \Omega^{-1} W_t' \right)^{-1} \sum_{t=1}^T W_t \Omega^{-1} z_t. \quad (4.13)$$

Merkitsemällä jälleen $\Omega^{-1} = \Phi$ voidaan edellisen jakson laskelmat toistaa ja todeta, että Φ :n uskottavuusyhtälö $\partial l(\delta, \Phi) / \partial \Phi = 0$ johtaa yhtälöön

$$\Omega = \frac{1}{T} \sum_{t=1}^T (z_t - W_t' \delta) (z_t - W_t' \delta)'. \quad (4.14)$$

Uskottavuusyhtälöiden ratkaiseminen on yhtäpitävää yhtälöiden (4.13) ja (4.14) ratkaisemisen kanssa. Koska ratkaisu δ :n suhteen vaatii Ω :n arvon tuntemisen ja ratkaisu Ω :n suhteen vaatii δ :n arvon tuntemisen (erikoistapauksia lukuun ottamatta), vaatii ratkaisu (yleensä) numeeristen menetelmien käyttöä. Havainnollinen tapa ratkaista uskottavuusyhtälöt etenee seuraavasti. Korvataan tuntematon Ω yhtälön (4.13) oikealla puolella edellisen jakson rajoittamattomalla estimaatilla $\hat{\Omega}$, jolloin saadaan

$$\tilde{\delta}^{(1)} = \left(\sum_{t=1}^T W_t \hat{\Omega}^{-1} W_t' \right)^{-1} \sum_{t=1}^T W_t \hat{\Omega}^{-1} z_t.$$

Tämä voidaan seuraavaksi sijoittaa yhtälön (4.14) oikealle puolelle, jolloin saadaan uusi Ω :n estimaatti $\tilde{\Omega}^{(1)}$, joka voidaan sijoittaa estimaatin $\tilde{\delta}^{(1)}$:n yhtälöön $\hat{\Omega}$:n paikalle. Näin saadaan uusi δ :n estimaatti $\tilde{\delta}^{(2)}$. Tätä menettelyä voidaan jatkaa, kunnes estimaatit eivät enää muutu. Tämä menettely esitetään lähinnä sen havainnollisuuden vuoksi. Käytännössä tietokoneohjelmat suorittavat ratkaisun jostain muuta algoritmia käyttäen. Mainittakoon kuitenkin, että jo yksinkertaisella yhden askeleen estimaattorilla $\tilde{\delta}^{(1)}$ on hyvät asymptoottiset ominaisuudet, sillä sen voidaan osoittaa olevan asymptoottisesti yhtäpitävä SU-estimaattorin kanssa (selitys piilee parametrien π ja Ω ortogonaalisuudessa).

Merkitään parametrien δ ja Ω SU-estimaattoreita symboleilla $\tilde{\delta}$ ja $\tilde{\Omega}$ ja johdetaan edellisen asymptoottinen jakauma vedoten SU-estimoinnin yleiseen teoriaan. Derivoimalla log-uskottavuusfunktiota toisen kerran δ :n suhteen saadaan (ks. (4.12) ja Liite A.7)

$$\frac{\partial^2 \delta}{\partial \delta \partial \delta'} l(\delta, \Omega) = - \sum_{t=1}^T W_t \Omega^{-1} W_t',$$

joten parametrin δ Fisherin informaatiomatriisi on

$$\mathcal{I}_{\delta\delta}(\delta, \Omega) = \sum_{t=1}^T \mathbf{E}(W_t \Omega^{-1} W_t'),$$

joka stationarisessa tapauksessa on yhtä kuin $T \mathbf{E}(W_t \Omega^{-1} W_t')$. Samanlainen laskelma kuin edellisessä jaksossa osoittaa, että $\mathbf{E}(\partial^2 l(\delta, \Omega) / \partial \omega_{ij}) = 0$ eli parametrit δ ja Ω ovat ortogonaaliset. Estimaattorin $\tilde{\delta}$ asymptoottiseksi jakaumaksi saadaan siten

$$\tilde{\delta}_{as} \sim \mathbf{N}(\delta, \mathcal{I}_{\delta\delta}(\delta, \Omega)^{-1}).$$

Fisherin informaatiomatriisissa voidaan ottaa huomioon määritelmä $W_t = H' X_t$, jossa H on tunnettu ja kiinteä, mutta kovin havainnollista lauseketta ei silti saada.

Kovarianssimatriisiin Ω rajoitetun SU-estimaattorin $\tilde{\Omega}$ asymptoottisista ominaisuuksista mainitaan vain tarkentuvuus, jota käytetään seuraavassa jaksossa. Edellä sanotusta seuraa, että estimaattori $\tilde{\Omega}$ voidaan kirjottaa

$$\tilde{\Omega} = \frac{1}{T} \sum_{t=1}^T (z_t - W_t' \tilde{\delta}) (z_t - W_t' \tilde{\delta})'$$

4.3 Hypoteesien testaus

Johdetaan Waldin testi lineaariselle hypoteesille (ks. (4.11))

$$H_0 : R\pi = b,$$

jonka voimassaolo seuraavassa oletetaan. Lisäksi oletetaan, että SU-estimaattoreilla $\hat{\pi}$ ja $\hat{\Omega}$ on jaksossa 4.1 esitetyt asymptoottiset ominaisuudet.

SU-estimaattorin $\hat{\pi}$ asymptoottinen normaalisuus on formaalisti esitettyä (oletetaan merkintöjen yksinkertaistamiseksi stationaarisuus)

$$\sqrt{T}(\hat{\pi} - \pi) \xrightarrow{d} \mathbf{N}(0, \Omega \otimes \Gamma_x^{-1}).$$

Koska $R\pi = b$, seuraa tästä (ks. Liite B, Lause B.2)

$$\sqrt{T}(R\hat{\pi} - b) = R\sqrt{T}(\hat{\pi} - \pi) \xrightarrow{d} Z \sim \mathbf{N}(0, R(\Omega \otimes \Gamma_x^{-1})R')$$

ja edelleen (ks. Liite B, Lause B.3)

$$T(R\hat{\pi} - b)' [R(\Omega \otimes \Gamma_x^{-1})R']^{-1} (R\hat{\pi} - b) \xrightarrow{d} Z' [R(\Omega \otimes \Gamma_x^{-1})R']^{-1} Z \sim \chi_q^2,$$

jossa viimeksi mainittu tulos seuraa tunnetusta multinormaalijakaumaa ja χ^2 -jakaumaa koskevasta tuloksesta.¹⁰ Vasemmalla oleva neliömuoto mittaa selväsikin hypoteesin $R\pi = b$ realistisuutta, mutta sisältää tuntemattoman matriisin $\Omega \otimes \Gamma_x^{-1}$, joka on testisuureessa korvattava (tarkentuvalla) estimaattorilla. Koska $\Gamma_x = \mathbf{E}(x_t x_t')$, jossa $x_t = [1 \ \mathbf{y}'_t]'$ ja $\mathbf{y}'_t = [y'_{t-1} \ \cdots \ y'_{t-p+1}]'$, ja koska jaksossa 3 todetun perusteella (stationaarisen) VAR(p)-prosessin tapauksessa ensimmäiset ja toiset otosmomentit ovat vastaavien teoreettisten momenttien tarkentuvia estimaattoreita, saadaan tämän ja estimaattorin $\hat{\Omega}$ tarkentuvuuden nojalla

$$[R(\hat{\Omega} \otimes \hat{\Gamma}_x^{-1})R']^{-1} \xrightarrow{p} [R(\Omega \otimes \Gamma_x^{-1})R']^{-1},$$

jossa

$$\hat{\Gamma}_x = T^{-1} \sum_{t=1}^T x_t x_t'.$$

Yhteenvedonä edellä esitetystä saadaan nyt *Waldin testisuure* (ks. Liite B, Lause B.3)

$$W = T(R\hat{\pi} - b)' [R(\hat{\Omega} \otimes \hat{\Gamma}_x^{-1})R']^{-1} (R\hat{\pi} - b) \xrightarrow{d} \chi_q^2.$$

Suuret testisuureen arvot ovat tietenkin kriittisiä. Käytännössä testiä sovelletaan laskemalla approksimatiivinen P-arvo

$$P = \mathbf{P}_{H_0} \{W \geq W(\mathbf{y})\} \approx \mathbf{P} \{\chi_q^2 \geq W(\mathbf{y})\},$$

jossa χ_q^2 on χ_q^2 -jakaumaa noudattava satunnaismuuttuja ja $W(\mathbf{y})$ on (satunnaisen) testisuureen W aineistosta laskettu arvo.

¹⁰Tämän tuloksen mukaan $Z \sim \mathbf{N}_k(\mu, \Sigma) \Rightarrow (Z - \mu)' \Sigma^{-1} (Z - \mu) \sim \chi_k^2$ (olettaen, että Σ on positiivisesti definiitti). Tässä kuten tarvittaessa myöhemminkin ilmaistaan multinormaalijakauman dimensio alaindeksillä.

Waldin testin asemasta voidaan käyttää *uskottavuusosamäärätestiä*. Testisuureen yleinen lauseke on

$$\text{LR} = 2 \left[l(\hat{\pi}, \hat{\Omega}) - l(\tilde{\pi}, \tilde{\Omega}) \right],$$

jossa $\tilde{\pi}$ ja $\tilde{\Omega}$ ovat parametrien π ja Ω jaksossa 4.2 johdetut rajoitetut SU-estimaattorit. Käyttäen hyväksi estimaattoreiden $\hat{\Omega}$ ja $\tilde{\Omega}$ sekä log-uskottavuusfunktion lausekkeita voidaan suoralla laskulla todeta, että (perustelu jätetään tehtäväksi)

$$l(\hat{\pi}, \hat{\Omega}) = -\frac{T}{2} \log \det(\hat{\Omega}) - \frac{Tn}{2} \quad \text{ja} \quad l(\tilde{\pi}, \tilde{\Omega}) = -\frac{T}{2} \log \det(\tilde{\Omega}) - \frac{Tn}{2}.$$

Uskottavuusosamäärätesti voidaan siten perustaa testisuureeseen

$$\text{LR} = T \left[\log \det(\tilde{\Omega}) - \log \det(\hat{\Omega}) \right] \xrightarrow{d} \chi_q^2,$$

jossa asymptoottinen jakauma voidaan perustella samoilla ehdoilla kuin Waldin testin tapauksessa. Suuret testisuureen arvot ovat jälleen kriittisiä, joten approksimatiiviset P-arvot lasketaan kuten Waldin testissä.

Edellä esitetyt testit voidaan yleistää koskemaan myös jaksossa 4.2 tarkastellun rajoitetun mallin parametrivektoria δ . Rajoitetun SU-estimaattorin asymptoottinen normaalisuus on formaalisti esitettyä (ks. jakson 4.2 loppu)

$$\sqrt{T}(\tilde{\delta} - \delta) \underset{as}{\sim} \text{N} \left(0, \left(\sum_{t=1}^T \text{E} (W_t \Omega^{-1} W_t') \right)^{-1} \right),$$

joten Waldin testiä varten tarvitaan asymptoottisesti pätevä empiirinen vastine matriisille $\sum_{t=1}^T \text{E} (W_t \Omega^{-1} W_t')$. Ilmeinen vaihtoehto on $T^{-1} \sum_{t=1}^T W_t \tilde{\Omega}^{-1} W_t'$, jota käyttäen Waldin testi voidaan johtaa samaan tapaan kuin edellä rajoittamattoman SU-estimaattorin tapauksessa. Matriisin $\left(\sum_{t=1}^T W_t \tilde{\Omega}^{-1} W_t' \right)^{-1}$ diagonaalialkioiden neliöjuurista saadaan approksimatiiviset keskivirheet estimaattorin $\tilde{\delta}$ komponenteille. Jos i :n komponentin keskivirheestä käytetään merkintää $\text{s.e.}(\tilde{\delta}_i)$, pätee $\tilde{\delta}_i / \text{s.e.}(\tilde{\delta}_i) \underset{as}{\sim} \text{N}(0, 1)$, joten parametrin δ_i approksimatiivinen 95%:n luottamusväli on $\tilde{\delta}_i \pm 1.96 \text{s.e.}(\tilde{\delta}_i)$. Vastaavalla tavalla voidaan muodostaa approksimatiivisia luottamusvälejä rajoittamattoman mallin parametrivektorin π komponenteille.

4.4 Mallinvalinta ja valitun mallin sopivuuden tutkiminen

Testien käyttö mallinvallinnassa. Usein VAR-mallin rakentaminen aloitetaan rajoittamattomasta mallista, koska tarpeeksi vahvaa etukäteistietoa parametrirajoitteista ei ole tai mielessä olevia rajoitteita halutaan testata käyttäen Waldin testiä tai uskottavuusosamäärätestiä. Tällöin ensimmäinen tehtävä on valita mallin (yleensä)

tuntematon aste. Jos asteeksi on valittu tavalla tai toisella p , voidaan valinnan sopivuutta tutkia testaamalla valittua VAR(p)–mallia jotain korkeampaa astetta olevaa vaihtoehtoa vastaan. Jos vaihtoehtona on VAR($p+s$)–malli

$$y_t = \nu + \sum_{j=1}^{p+s} A_j y_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim \text{nid}(0, \Omega),$$

testataan (lineaarista) hypoteesia

$$H_0 : A_{p+1} = \dots = A_{p+s} = 0.$$

Jos $\hat{\Omega}(k) = T^{-1} \sum_{t=1}^T (y_t - X_t' \hat{\pi})(y_t - X_t' \hat{\pi})'$ on rajoittamattoman VAR(k)–mallin virhekovarianssimatriisin Ω SU–estimaattori, voidaan käyttää uskottavuusosamäärätestiä ja testisuuretta

$$\text{LR} = T \left[\log \det(\hat{\Omega}(p)) - \log \det(\hat{\Omega}(p+s)) \right] \underset{as}{\sim} \chi_{n^2 s}^2.$$

Huomaa, että tässä (samoin kuin vastaavan tyyppisissä testisuureissa alempana) estimaattorit $\hat{\Omega}(p)$ ja $\hat{\Omega}(p+s)$ perustuvat samaan määrään havaintoja, joten (ehdolisessa) SU–estimoinnissa y_{-p-s+1}, \dots, y_0 toimivat alkuarvoina.

Usein testaus suoritetaan peräkkäin aloittaen jostain ”tarpeeksi suuresta” asteesta P . Jos hypoteesi $A_P = 0$ jää voimaan, testataan hypoteesia $A_{P-1} = 0$. Jos tämäkin hypoteesi jää voimaan, jatketaan testausta, kunnes saadaan ensimmäinen hylkäävä testitulos tai päädytään asteeseen 0. Tämä menetelmä vastaa yksiulotteisesta tapauksesta tutun osittaisautokorrelaatiofunktion käyttämistä AR–mallin asteen valinnassa. Jos ensimmäinen hylkäys tapahtuu suurelta tuntuvan asteen kohdalla, voidaan testausta silti jatkaa samaan tapaan kuin osittaisautokorrelaatiofunktioita käytettäessä välttämättä tiukkaa formaalista testausta. Jonon i . testisuure on

$$\text{LR}(i) = T \left[\log \det(\hat{\Omega}(P-i)) - \log \det(\hat{\Omega}(P-i+1)) \right] \quad (1 \leq i \leq P),$$

jonka saamia arvoja verrataan $\chi_{n^2}^2$ –jakauman prosenttipisteisiin. On syytä huomata, että yksittäisistä testeistä muodostetun yhteistestin merkitsevyytaso (vastaavasti P –arvo) riippuu kaikkien suoritettujen testien merkitsevyytasoista (vastaavasti P –arvoista). Voidaan osoittaa, että jos k ensimmäistä hypoteesia on voimassa, ovat testisuuret $\text{LR}(i)$ ja $\text{LR}(j)$, $i, j \leq k$ ($i \neq j$) asympotoottisesti riippumattomia ja (approssimatiivinen) todennäköisyys hylätä jokin niistä saadaan kaavasta

$$\alpha_k = 1 - (1 - \gamma_1) \cdots (1 - \gamma_k) \quad (1 \leq k \leq P),$$

jossa γ_i on i :nnessä testissä käytetty merkitsevyytaso. Jos $\gamma_i = 0.05$ kaikilla i , on $\alpha_1 = 0.05$, $\alpha_2 = 0.0975$, $\alpha_3 = 0.143$ ja $\alpha_4 = 0.185$.

Kun rajoittamattoman mallin aste on valittu, voidaan Waldin testiä tai uskottavuusosamäärätestiä soveltaa rajoitetun mallin yksityiskohtaisemmassa täsmennyksessä. Tämä on luontevaa, kun esimerkiksi taustateoriaa käyttäen voidaan muotoilla kiinnostavia hypoteeseja. Jos näin ei ole ja verrattavia mallivaihtoehtoja on useita, voivat alempana tarkasteltavat mallinvalintakriteerit olla luontevampia vaihtoehtoja.

Esimerkki 3.1 (jatkoa). Sovelletaan edellä kuvattua asteen valintamenettelyä Kuvion 1.1 valuuttakurssi/korkoaineistoon. Valittaessa maksimiasteeksi $P = 6$ saadaan seuraavat tulokset.

Testisuure :	LR (2)	LR (3)	LR (4)	LR (5)	LR (6)
p-arvo :	0.003	0.678	0.359	0.434	0.866

Tulokset viittaavat (olettaen käytetyn maksimiasteen riittävyys) selvästi asteeseen $p = 2$, jota käytettäessä saadaan estimoiduksi malliksi

$$\begin{aligned} \begin{bmatrix} dexch_t \\ rdiff_t \end{bmatrix} &= \begin{bmatrix} -0.002 \\ (0.002) \\ 0.004 \\ (0.014) \end{bmatrix} + \begin{bmatrix} 0.418 & 0.023 \\ (0.089) & (0.012) \\ -0.187 & 1.144 \\ (0.672) & (0.090) \end{bmatrix} \begin{bmatrix} dexch_{t-1} \\ rdiff_{t-1} \end{bmatrix} \\ &+ \begin{bmatrix} -0.178 & -0.026 \\ (0.087) & (0.012) \\ 0.603 & -0.179 \\ (0.659) & (0.089) \end{bmatrix} \begin{bmatrix} dexch_{t-2} \\ rdiff_{t-2} \end{bmatrix} + \begin{bmatrix} \hat{\varepsilon}_{1t} \\ \hat{\varepsilon}_{2t} \end{bmatrix}. \end{aligned} \quad (4.15)$$

Estimaattien alla suluissa on (approksimatiiviset) keskivirheet. Residuaalien välinen korrelaatiokerroin on -0.259 , joten mallin virhetermit ovat lievästi korreloituneita. Estimoitu malli toteuttaa stationaarisuusehdon (2.9), sillä matriisin \mathbf{A} SU-estimaatin suurin ominaisarvo on itseisarvoltaan 0.953 . Jotkut estimaatit ovat pieniä keskivirheisiinsä verrattuna, joten mallin rajoittaminen edelleen näyttää mahdolliselta. Tähän kysymykseen palataan myöhemmin.

Mallivalintakriteerit. Rajoittamattoman mallin asteen valinnassa voidaan käyttää myös *mallinvalintakriteerejä*, joita on johdettu useita periaatteita käyttäen. Olkoon jälleen P ”tarpeeksi suuri” ennalta valittu maksimiaste ja $\hat{\Omega}(k)$ on kuten edellä. Tällöin eräs yleinen kriteerityyppi on

$$C(k) = \log \det(\hat{\Omega}(k)) + \frac{f(T)}{T}d, \quad k = 1, \dots, P, \quad (4.16)$$

jossa k on mallin aste, d vastaava vapaasti estimoitujen autoregressiivisten parametrien ja vakiotermien lukumäärä (rajoittamattoman mallin tapauksessa $d = n(nk + 1)$) ja ns. *sakkofunktio* f on positiivinen ja toteuttaa $f(T)/T \rightarrow 0$, kun $T \rightarrow \infty$. Sakko-funktion idea on rankaista tarpeettoman laajan mallin käyttämisestä. Kriteerifunktion ensimmäinen termi mittaa sovituksen hyvyttä ja, jos asteen k kasvattaminen ei

pienennä sitä tarpeeksi, ei laajempaa mallia suosita. Tämä tulee ilmeiseksi, kun huomataan, että kriteerifunktion arvoja verrattaessa ensimmäisen termin paikalla voisi olla $-n - (2/T) \times \log$ -uskottavuusfunktion maksimi-arvo (ks. jakso 4.3). ”Hyvillä” malleilla $C(k)$ saa siten pienen arvon. Tunnettuja sakkofunktioita ovat

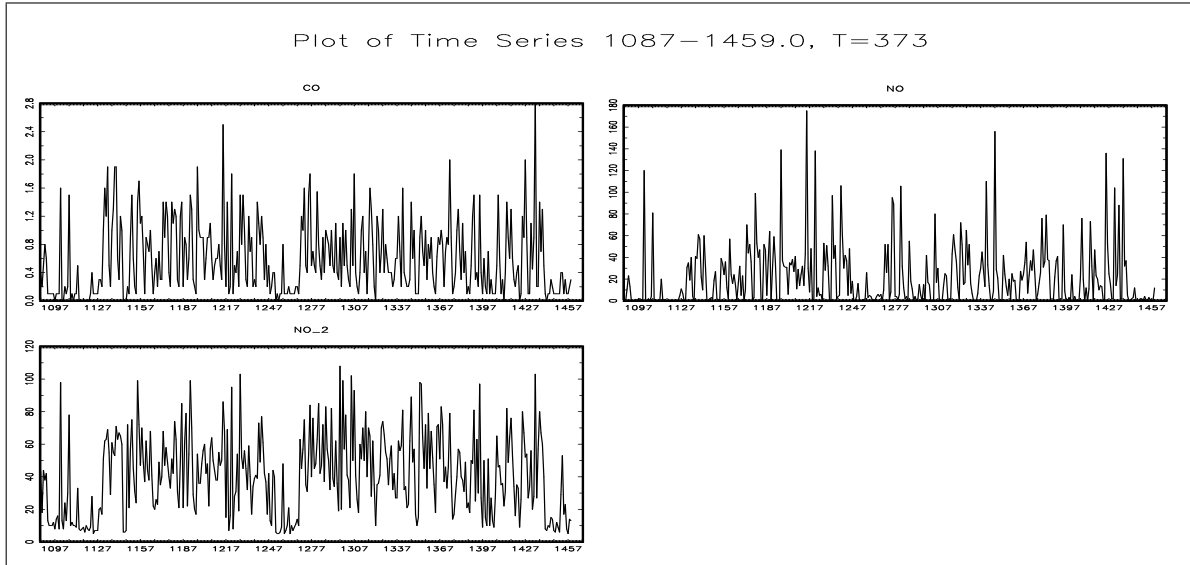
$$\begin{aligned} \text{AIC: } f(T) &= 2 && (\text{Akaike, 1974}) \\ \text{HQ: } f(T) &= 2 \log(\log T) && (\text{Hannan ja Quinn, 1979}) \\ \text{BIC: } f(T) &= \log T && (\text{Schwarz, 1978, Rissanen, 1978}). \end{aligned}$$

Näistä ensimmäinen sakottaa vähiten (suosii laajempia malleja) ja viimeinen eniten (suosii suppeampia malleja). Kriteerifunktiosta HQ on myös versioita, joissa vakion 2 paikalla on joku muu vakio. Näistä kolmesta kriteerifunktiosta kaksi viimeksi mainittua tuottaa asteen tarkentuvan estimaattorin, kun valinta perustetaan kriteerifunktion minimointiin. Käytännössä mallinvalintakriteerejä on kuitenkin suositeltavaa käyttää vain yhtenä mallinvalinnan apuvälineenä eikä valita mallia minimoimalla kriteerifunktiota mekaanisesti. Kun Esimerkin 3.1 tapauksessa maksimiasteeksi valitaan $P = 6$, saadaan BIC:tä käyttäen ”parhaaksi” asteeksi $p = 1$, kun taas AIC ja HQ suosittelevat valintaa $p = 2$.

Kuten edellä mainittiin, voidaan mallinvalintakriteerejä käyttää myös erilaisten rajoitettujen mallien vertailuun. Tällöin yhtälön (4.16) oikealla puolella SU-estimaattori $\hat{\Omega}(k)$ muodostetaan rajoitetusta VAR(k)-mallista ja d on rajoitetun mallin vapaasti estimoitujen AR-parametrien ja vakiotermien lukumäärä (jakson 4.2 VAR(p)-mallissa $d = \dim(\delta)$). Tätä menettelyä käytetään erityisesti, kun hypoteesien muotoilu ei ole selkeää ja vertailtavia mallivaihtoehtoja on useita. Yksi vaihtoehto on estimoida kaikki mahdolliset mallikombinaatiot ja perustaa valinta saatuihin kriteerifunktion arvoihin kuten rajoittamattoman mallin tapauksessa. Koska tämä voi johtaa sangen raskaisiin laskelmiin, on esitetty erilaisia askeltavia vaihtoehtoja. Nollarajoitteiden tapauksessa yksi mahdollisuus on poistaa lähtökohdaksi olevasta mallista vuorotellen yksi selittävä muuttuja (eli viipymän y_{t-i} , $i = 1, \dots, p$ komponentti) ja laskea näin saaduille malleille käytettävän kriteerifunktion arvot. Seuraavassa vaiheessa valitaan lähtökohdaksi suppeammista malleista se, jota vastaava kriteerifunktion arvo on pienentynyt eniten lähtökohdaksi olevaan malliin verrattuna. Näin jatketaan poistaen aina yksi selittävästä muuttujista kerrallaan, kunnes kriteerifunktion arvo ei pienene edellisen vaiheen malliin verrattuna. Seuraavassa esimerkki tästä menettelystä.

Esimerkki 4.1. Kuviossa 4.1 esitetyt kolme aikasarjaa kuvaavat Heidelbergissä heinä- ja elokuussa 1991 tehtyjä päivittäisiä mittauksia hiilimonoksidin (CO), typioksidin (NO) ja typpidioksidin (NO₂) määristä (mg/m³). Kultakin päivältä on kuusi tasavälistä mittausta ja havaintoja on kaikkiaan 373. Hiilimonoksidin ja typioksidin määrään vaikuttavat liikenteen, teollisuuden ja energiatuotannon päästöt, jotka osaltaan kasvattavat typpidioksidin määrää.

On odotettavaa, että peräkkäisten päivien välillä on riippuvuutta, joten aivan



Kuvio 4.1. Hiilimonoksidin (CO), typpioksidin (NO) ja typpidioksidin (NO₂) määrät (mg/m³) Heidelbergissä heinä- ja elokuussa 1991. Kuusi tasavälistä mittausta päivittäin.

pieniasteinen VAR-malli ei välttämättä toimi. Rajoittamattoman mallin astetta valittaessa käytettiin mallinvalintakriteerejä maksimiasteena 15 (kaksi ja puoli päivää). AIC suositteli astetta 8, HQ astetta 6 ja BIC astetta 2. Valitaan asteeksi AIC:n suosituksen mukaan $p = 8$ ja käytetään edellä kuvattua askeltavaa menettelyä sovellettaessa ankarampaa HQ-kriteeriä. Tulokseksi saadaan parametrimatriiseille A_1, \dots, A_8 seuraava rakenne. Estimaattien asemesta ilmoitetaan ”t-suhteet” eli estimaatit jaetuna approksimatiivisilla keskivirheillään. Lihavoidut nollat osoittavat parametrit, jotka menettely ehdottaa rajoitettavaksi nollassi.

$$\tilde{A}_1 = \begin{bmatrix} 5.39 & \mathbf{0} & \mathbf{0} \\ -2.58 & 6.10 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 5.36 \end{bmatrix} \quad \tilde{A}_2 = \begin{bmatrix} -2.67 & 2.76 & 3.64 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 5.07 \end{bmatrix}$$

$$\tilde{A}_3 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 2.26 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \tilde{A}_4 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 3.17 & \mathbf{0} & -2.29 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$\tilde{A}_5 = \begin{bmatrix} 4.02 & -3.95 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -1.99 & \mathbf{0} \end{bmatrix} \quad \tilde{A}_6 = \begin{bmatrix} \mathbf{0} & 5.14 & 2.49 \\ \mathbf{0} & 9.45 & \mathbf{0} \\ -1.90 & 4.00 & 3.86 \end{bmatrix}$$

$$\tilde{A}_7 = \begin{bmatrix} 2.34 & -2.30 & -2.00 \\ \mathbf{0} & -3.10 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \tilde{A}_8 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 2.18 \\ 4.14 & -4.03 & \mathbf{0} \end{bmatrix}$$

Estimoinnissa käytettiin myös vakiotermiä, jonka kaikkien komponenttien estimaatit olivat ”selvästi merkitseviä”. Malli toteuttaa stationaarisuusehdon (2.9), sillä matriisin \mathbf{A} SU-estimaatin suurin ominaisarvo on itseisarvoltaan 0.893. Virheiden välillä ilmenee selvää korreloituneisuutta, sillä kovarianssimatriisiestimaatista $\tilde{\Omega}$ muodostetut korrelaatiokertoimet vaihtelevat 0.6:n ja 0.8:n välillä. Estimaateista $\tilde{A}_1, \dots, \tilde{A}_8$ on hankala tehdä yksityiskohtaisia päätelmiä. Nollat niiden viimeisen rivin kahdella ensimmäisellä sarakkeella osoittavat kuitenkin, että hiilimonoksidin ja typpioksidin vaikutus typpidioksiidiin ilmenee vasta viipymiltä 5 ja 6 alkaen. Muut vaikutussuhteet ilmenevät sen sijaan jo viipymilla 1 ja 2.

Valitun mallin sopivuuden tutkiminen. Edellä esitettyjen menetelmien lisäksi mallin sopivuutta voidaan tutkia myös käyttäen residuaaleja. Rajoittamattoman mallin tapauksessa residuaalien määritelmä on

$$\hat{\varepsilon}_t = y_t - \hat{\nu} - \hat{A}_1 y_{t-1} - \dots - \hat{A}_p y_{t-p}, \quad t = 1, \dots, T.$$

Residuaalien tulisi muistuttaa ominaisuuksiltaan teoreettisia virheitä ε_t . Tämän tarkistamiseksi kannattaa ensiksi piirtää residuaalisarjojen kuvat ja tutkia löytyykö poikkeavia residuaaleja tai residuaaliryhmiä, varianssin vaihtelua tai muita selviä systemaattisia piirteitä. Myös residuaalien normaalisuutta on hyvä tutkia, vaikka mallin asymptoottinen estimointi- ja testiteoria ei normaalisuutta vaadikaan (iid-oletus riittää).

Residuaalien auto- ja ristikorreloituneisuutta voidaan tutkia laskemalla residuaalisarjoista jaksossa 3 esitetyt auto- ja ristikorrelaatiofunktiot. Jos teoreettisille virheille pätee $\varepsilon_t \sim \text{iid}(0, \Omega)$, ovat auto- ja ristikorrelaatioestimaattorit asymptoottisesti normaalisti jakautuneita odotusarvona nolla. Toisin kuin teoreettisten virheiden tapauksessa ne eivät kuitenkaan ole asymptoottisesti riippumattomia eikä jaksossa 3 esitetty $N(0, 1/T)$ -jakauma-approksimaatio ole pätevä. Oikea asymptoottinen jakauma tunnetaan, mutta sitä ei tarkastella tällä kurssilla. Jotkut tietokoneohjelmat tulostavat residuaalien auto- ja ristikorrelaatioestimaattorien oikeat asymptoottiset keskivirheet tai niihin perustuvat kriittiset rajat, joiden avulla residuaalien mahdollista korreloituneisuutta voidaan arvioida.

Residuaalien mahdollisen auto- ja ristikorreloituneisuuden tutkimisessa voidaan käyttää apuna myös tilastollisia testejä. Ns. *portmanteau-testi* perustuu suoraan residuaalien auto- ja ristikorrelaatiofunktioihin. Eräs versio testisuureesta on

$$Q_K^* = T^2 \sum_{k=1}^K \frac{1}{T-k} \text{tr}(\mathcal{S}'_k \mathcal{S}_0^{-1} \mathcal{S}_k \mathcal{S}_0^{-1}), \quad \mathcal{S}_k = \sum_{t=1}^{T-k} \hat{\varepsilon}_t \hat{\varepsilon}'_{t+k}.$$

Jos valitun rajoittamattoman VAR(p)-mallin virheille pätee $\varepsilon_t \sim \text{iid}(0, \Omega)$, noudattaa testisuure Q_K^* asympotoottisesti χ^2 -jakaumaa vapausastein $n^2(K-p)$. Jos kysymyksessä on rajoitettu malli, tulee χ^2 -jakauman vapausasteluvuksi n^2K miinus vapaasti estimoitujen autoregressiivisten parametrien lukumäärä. Asympotoottinen tulos edellyttää lisäksi, että K on ”suuri”.

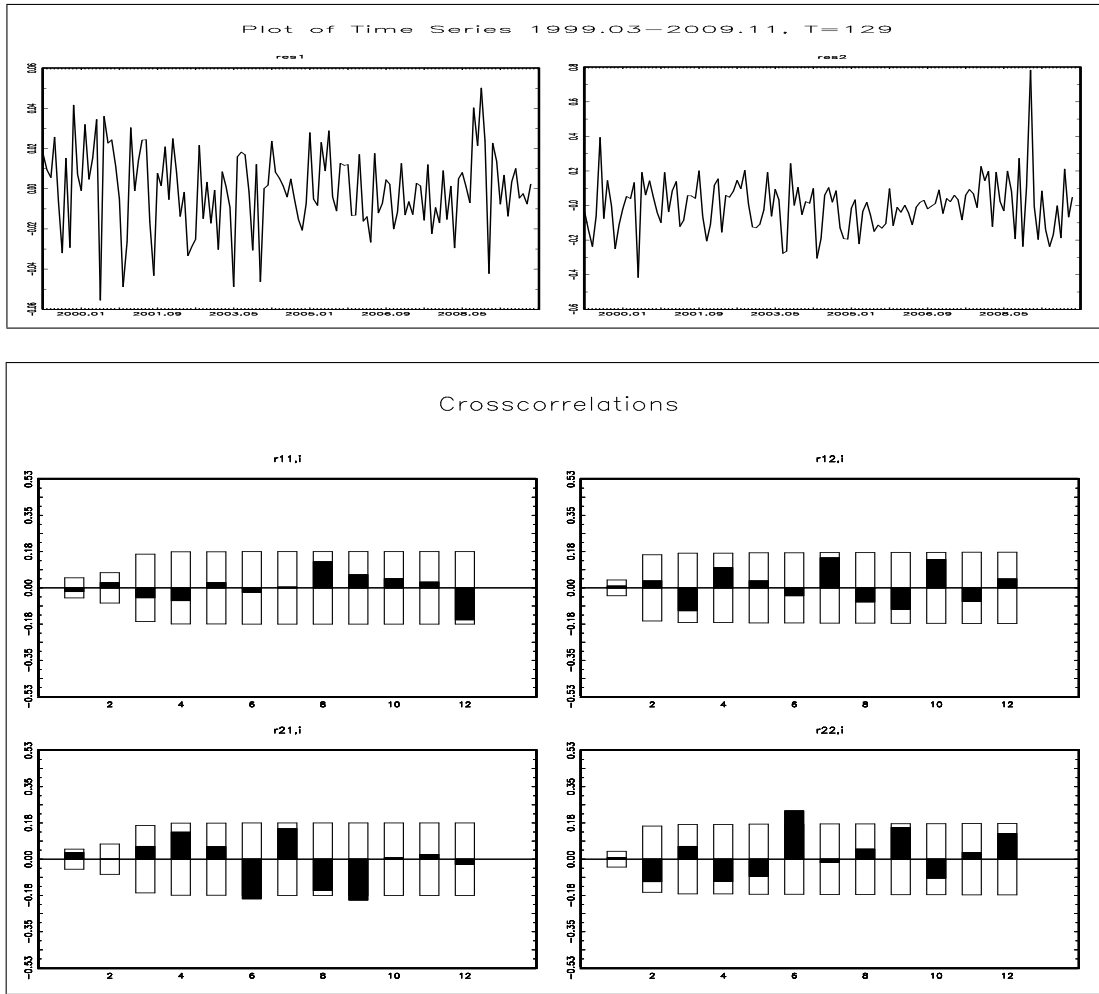
Vaikka residuaalien korreloitunaisuutta ei olisikaan syytä epäillä, eivät residuaalit ole silti välttämättä riippumattomia, sillä korreloimattomuus merkitsee vain lineaarisen riippuvuuden puuttumista. Mahdollisen epälineaarisen riippuvuuden tutkimiseksi voidaan (rajoitetussa mielessä) käyttää neliöidyistä residuaaleista laskettuja auto- ja ristikorrelaatiofunktioita.

Esimerkki 3.1 (jatkoa). Kuviossa 4.2 nähdään estimoidun mallin (4.15) residuaalit ja niistä lasketut auto- ja ristikorrelaatiofunktiot. Ensimmäisen residuaalin vaihtelu näyttää olevan sarjan alkupuolella hieman voimakkaampaa kuin loppupuolella ja toisessa residuaalisarjassa silmään pistävät muutamat suurehkot arvot. Näistä piirteistä huolimatta residuaalit ovat kuitenkin ”kohtuullisen epämääräisen” näköisiä. Residuaalien auto- ja ristikorrelaatiot ovat pieniä muutamaa suurehkoa ja melko kaukaisilla viipymillä esiintyvää ristikorrelaatiota lukuun ottamatta. Portmanteautestisuureen arvo ($Q_{12}^* = 51.38$) ei kuitenkaan hälytä, sillä likimääräinen p-arvo on 0.11. Piirtämällä (standardoitujen) residuaalien histogrammit saadaan viitteitä normaalijakaumaa paksuhäntäisemmästä jakaumasta, mihin residuaalien graafinen tarkastelu myös viittaa. Ei-normaalisuuteen (ja ns. ehdolliseen heteroskedastisuuteen) viittaa myös neliöityjen residuaalien lievä korreloituneisuus muutamalla viipymällä. Kaiken kaikkiaan mallia voidaan kuitenkin pitää kohtuullisen onnistuneena.

4.5 Grangerin kausaalisuus

Yleinen määritelmä. Tarkastellaan aluksi kahta reaaliarvoista muuttujaa x ja z ja pohditaan miten voitaisiin määritellä kausaalisuus edellisestä jälkimmäiseen. Menemättä filosofisiin kysymyksiin voidaan ajatella, että kausaalisuudelle on välttämätöntä, ettei syytapahduma tapahdu seurauksen jälkeen. Jos muuttuja x on muuttujan z syy, täytyisi muuttujasta x olla siten hyötyä ennustettaessa muuttujaa z .

Olkoon nyt x_t ja z_t *reaaliarvoisia* stokastisia prosesseja ja tarkastellaan prosessin z_t tulevien arvojen ennustamista ajankohtana t . Merkitään symbolilla \mathcal{F}_t hypoteettista muuttujajoukkoa, joka x_t :n lisäksi sisältää kaikki relevantit muuttujat, jotka ajankohdана t voivat tulla kysymykseen z_t :n tulevia arvoja z_{t+h} , $h \geq 1$, ennustettaessa. Tätä muuttujajoukkoa sanotaan usein informaatiojoukoksi. Siihen perustuvaa z_{t+h} :n (keskineliövirheen mielessä) optimaalista ennustetta merkitään symbolilla $z_t(h|\mathcal{F}_t)$, kun taas $z_t(h|\mathcal{F}_t \setminus \{x_s, s \leq t\})$ on z_{t+h} :n optimaalinen ennuste, kun $\{x_s, s \leq t\}$ on poistettu joukosta \mathcal{F}_t . Olkoon $\Sigma_z(h|\mathcal{F}_t)$ ja $\Sigma_z(h|\mathcal{F}_t \setminus \{x_s, s \leq t\})$ vastaavat keskineliövirheet, jotka ovat tässä siis skalaareja (ks. jakso 2.3 tässä mainituista käsitteistä). Prosessien



Kuvio 4.2. Mallin (4.15) residuaalit (ylhäällä vas. $\hat{\varepsilon}_{1t}$ ja ylhäällä oik. $\hat{\varepsilon}_{2t}$), ja niistä lasketut auto- ja ristikorrelaatiofunktiot. Valkoiset pylväät osoittavat kriittiset rajat ($\pm 2 \times$ likimääräinen keskivirhe, kun $\varepsilon_t \sim \text{iid}(0, \Omega)$).

x_t ja z_t välillä sanotaan olevan *Grangerin kausaalisuus* edellisestä jälkimmäiseen eli $x \rightarrow z$, jos

$$\Sigma_z(h|\mathcal{F}_t) < \Sigma_z(h|\mathcal{F}_t \setminus \{x_s, s \leq t\})$$

ainakin yhdellä t ja $h = 1, 2, \dots$. Toisin sanoen, prosessia z_t voidaan ennustaa tarkemmin käyttämällä prosessia x_t kuin ilman sitä.

Okoon nyt x_t ja z_t vektoriarvoisia prosesseja ja $\Sigma_z(h|\mathcal{F}_t)$ ja $\Sigma_z(h|\mathcal{F}_t \setminus \{x_s, s \leq t\})$ edellä määriteltyjen reaalisten keskineliövirheiden ilmeisiä matriisivastineita. Tällöin pätee välttämättä (vrt. jaksossa 2.3 ennusteosion alku)

$$\Sigma_z(h|\mathcal{F}_t \setminus \{x_s, s \leq t\}) - \Sigma_z(h|\mathcal{F}_t) \geq 0.$$

Koska ei ole kohtuullista vaatia, että vasemmalla oleva matriisi olisi positiivisesti definiitti (miksi?), sanotaan *Grangerin kausaalisuuden* pätevän suuntaan $x \rightarrow z$, jos

$$\Sigma_z(h|\mathcal{F}_t) \neq \Sigma_z(h|\mathcal{F}_t \setminus \{x_s, s \leq t\})$$

ainakin yhdellä t ja $h = 1, 2, \dots$. Tulkinta on kuitenkin olennaisesti sama kuin reaalisessa tapauksessa. Huomaa, että esitetty määritelmä ei ota kantaa saman aikaiseen kausaalisuuteen eli siihen voidaanko z_t :n ennustetta parantaa käyttämällä hyväksi myös muuttujaa x_t . Tällaista saman aikaista kausaalisuutta on tarkasteltu kirjallisuudessa, mutta se sivuutetaan tällä kurssilla.

Grangerin kausaalisuus VAR-mallissa. Grangerin kausaalisuuden määritelmässä käytetty informaatiojoukko \mathcal{F}_t on käytännön kannalta tietenkin ongelmallinen. Käytännössä se rajataan käytettävissä olevien tai valittujen muuttujien joukoksi, joka VAR(p)-mallin tapauksessa muodostuu vektorin y_t komponenteista ja niiden viipymistä eli $\mathcal{F}_t = \{y_s, s \leq t\}$.

Tarkastellaan nyt stationaarista VAR(p)-prosessia

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \Omega).$$

Ositetaan $y_t = (z_t, x_t)$ ja kirjoitetaan tämä yhtälö vastaavasti ositettuna

$$\begin{bmatrix} z_t \\ x_t \end{bmatrix} = \sum_{j=1}^p \begin{bmatrix} A_{11,j} & A_{12,j} \\ A_{21,j} & A_{22,j} \end{bmatrix} \begin{bmatrix} z_{t-j} \\ x_{t-j} \end{bmatrix} + \begin{bmatrix} \zeta_t \\ \xi_t \end{bmatrix}. \quad (4.17)$$

Tuntuu varsin selvältä, että Grangerin kausaalisuus suuntaan $x \rightarrow z$ liittyy siihen onko $A_{12,j} \neq 0$ jollain $1 \leq j \leq p$. Seuraavassa tämä perustellaan formaalisti.

Lause 4.1. Jos $y_t = (z_t, x_t)$ on kuten edellä, niin Grangerin kausaalisuus suuntaan $x \rightarrow z$ ei päde jos ja vain jos $A_{12,j} = 0$ kaikilla $j = 1, \dots, p$.

Todistus: Oletetaan yksinkertaisuuden vuoksi, että $p = 1$ ja merkitään $A_{ij,1} = A_{ij}$ (yleisen tapauksen todistus sujuu samalla tavalla).

Oletetaan $A_{12} = 0$. Tällöin prosessille z_t saadaan yhtälöstä (4.17) stationaarinen VAR(1)-esitys (miksi?)

$$z_t = A_{11} z_{t-1} + \zeta_t.$$

Tästä seuraa (ks. jakso 2.3)

$$\mathbf{E}(z_{t+h} | z_s, s \leq t) = A_{11}^h \mathbf{E}(z_{t+h-1} | z_s, s \leq t), \quad h \geq 1.$$

Toisaalta yhtälöstä (4.17) seuraa tapauksessa $A_{12} = 0$ myös

$$\mathbf{E}(z_{t+h} | y_s, s \leq t) = A_{11}^h \mathbf{E}(z_{t+h-1} | y_s, s \leq t), \quad h \geq 1.$$

Näistä yhdessä seuraa

$$\begin{aligned}
\mathbf{E}(z_{t+1}|z_s, s \leq t) &= A_{11}z_t = \mathbf{E}(z_{t+1}|y_s, s \leq t) \\
\mathbf{E}(z_{t+2}|z_s, s \leq t) &= A_{11}\mathbf{E}(z_{t+1}|z_s, s \leq t) \\
&= A_{11}(A_{11}z_t) \\
&= A_{11}\mathbf{E}(z_{t+1}|y_s, s \leq t) \\
&= \mathbf{E}(z_{t+2}|y_s, s \leq t) \\
&\vdots
\end{aligned}$$

ja jatkaen induktiivisesti nähdään, että ennusteet $\mathbf{E}(z_{t+h}|z_s, s \leq t)$ ja $\mathbf{E}(z_{t+h}|y_s, s \leq t)$ ovat samat kaikilla h ja t . Tästä seuraa, ettei Grangerin kausaalisuutta ole.

Oletetaan nyt, että $A_{12} \neq 0$. Koska $\varepsilon_{t+1} \perp\!\!\!\perp \{y_s, s \leq t\}$, on $\mathbf{E}(\varepsilon_{t+1}|y_s, s \leq t) = \mathbf{E}(\varepsilon_{t+1}) = 0$ ja yhtälöstä (4.17) saadaan

$$z_{t+1} - \mathbf{E}(z_{t+1}|y_s, s \leq t) = z_{t+1} - A_{11}z_t - A_{12}x_t = \zeta_{t+1}$$

ja

$$\begin{aligned}
z_{t+1} - \mathbf{E}(z_{t+1}|z_s, s \leq t) &= z_{t+1} - A_{11}z_t - A_{12}\mathbf{E}(x_t|z_s, s \leq t) \\
&= \zeta_{t+1} + A_{12}[x_t - \mathbf{E}(x_t|z_s, s \leq t)].
\end{aligned}$$

Viimeisellä rivillä olevat kaksi sv:ia ovat riippumattomia (miksi?) ja jälkimmäinen on nolasta poikkeava, koska muutoin x_t olisi sv:ien $z_s, s \leq t$ funktio. Ennusteen $\mathbf{E}(z_{t+1}|z_s, s \leq t)$ keskineliövirhe poikkeaa näin ollen optimaalisen ennusteen $\mathbf{E}(z_{t+1}|y_s, s \leq t)$ keskineliövirheestä eli ζ_{t+1} :n kovarianssimatriisista, joten Grangerin kausaalisuus pätee suuntaan $x \rightarrow z$. \square

On selvää, että Lauseen 4.1 tulos pätee myös silloin, kun malliin lisätään vakiotermi (tehdään muunnos $y_t \rightarrow y_t - \mathbf{E}(y_t)$). Grangerin kausaalisuutta VAR(p)-mallissa voidaan edellä kuvatussa tilanteessa siten testata testaamalla lineaarista hypoteesia

$$H_0 : A_{12,1} = \dots = A_{12,p} = 0.$$

Näiden rajoitteiden ilmaiseminen yhtälöä (4.10) tai (4.11) käyttäen (ks. jakso 4.2) on kuitenkin hankalaa yleisillä $n:n$ ja $p:n$ arvoilla. Periaatteessa ongelmia ei kuitenkaan ole. Jos z_t on $n_1 \times 1$ vektori ja x_t on $n_2 \times 1$ vektori ($n_1 + n_2 = n$) on $A_{12,j}$ $n_1 \times n_2$ matriisi ja rajoitteiden lukumäärä on $n_1 n_2 p$. Waldin testiä tai uskottavuusosamäärätestiä voidaan siten käyttää yhdessä (asymptoottisen) $\chi^2_{n_1 n_2 p}$ -jakauman kanssa.

Esimerkki 3.1 (jatkoa). Mallin (4.15) estimointitulosten perusteella on selvää, ettei Grangerin kausaalisuus suuntaan $dexh_t \rightarrow rdif_t$ saa tukea aineistolta (vertaa relevantteja estimaatteja niiden keskivirheisiin). Käyttäen formaalia uskottavuusosamäärätestiä saadaan (likimääräiseksi) p-arvoksi 0.65. Grangerin kausaalisuus

suuntaan $rdif_t \rightarrow dexch_t$ näyttää estimointitulosten perusteella sen sijaan realistiselta. Formaalisissa testauksessa P-arvoksi tulee 0.06. Näiden tarkastelujen perusteella saadun rajoitetun mallin estimointi tuottaa tulokseksi

$$\begin{bmatrix} dexch_t \\ rdif_t \end{bmatrix} = \begin{bmatrix} -0.002 \\ (0.002) \\ 0.003 \\ (0.014) \end{bmatrix} + \begin{bmatrix} 0.411 & 0.023 \\ (0.084) & (0.012) \\ 0 & 1.147 \\ & (0.086) \end{bmatrix} \begin{bmatrix} dexch_{t-1} \\ rdif_{t-1} \end{bmatrix} \quad (4.18)$$

$$+ \begin{bmatrix} -0.158 & -0.026 \\ (0.082) & (0.012) \\ 0 & -0.185 \\ & (0.086) \end{bmatrix} \begin{bmatrix} dexch_{t-2} \\ rdif_{t-2} \end{bmatrix} + \begin{bmatrix} \tilde{\varepsilon}_{1t} \\ \tilde{\varepsilon}_{2t} \end{bmatrix}.$$

Verrattuna mallin (4.15) estimointiin on estimointitarkkuus (odotetusti) parantunut, kuten keskivirheitä vertaamalla voidaan todeta.

4.6 Impulssivasteanalyysi

Impulssivasteiden määrittely. Tarkastellaan samaa VAR(p)-mallia kuin edellisessä jaksossa, jossa esitetyn kausaalisuustarkastelun lisäksi mallia voidaan yrittää tulkita ns. impulssivasteanalyysin avulla. Kuten jaksossa 2.3 mainittiin, impulssivasteanalyysissä mallin virheet tai innovaatiot ε_t ajatellaan ”impulseiksi” tai ”satunnaissokeiksi”, joiden vaikutusta havaittaviin muuttujiin (eli y_t :n komponentteihin) voidaan tutkia ratkaisemalla y_t :n VMA(∞)-esitys (ks. jakso 2.3):

$$y_t = A(\mathbf{B})^{-1} \varepsilon_t = \Psi(\mathbf{B}) \varepsilon_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}, \quad (4.19)$$

jossa $\Psi(\mathbf{B}) = \sum_{j=0}^{\infty} \Psi_j \mathbf{B}^j$, $\Psi_0 = I_n$.

Oletetaan nyt, että innovaation ε_t yhdessä komponentissa ε_{kt} ($1 \leq k \leq n$) tapahtuu yhden yksikön muutos (esimerkiksi) ajankohtana $t = 0$. Jotta tämän muutoksen tai impulssin vaikutukset voitaisiin eristää ajatellaan (keinotekoisesti), että ennen ajankohtaa $t = 0$ prosessi on ollut odotusarvossaan eli tässä tapauksessa $y_t = 0$, $t < 0$. Olkoon edelleen $\varepsilon_{k0} = 1$ ja, jotta voitaisiin nähdä miten impulssin vaikutus leviää prosessin y_t tulevissa arvoissa, oletetaan (edelleen keinotekoisesti), että $\varepsilon_{i0} = 0$, $i \neq k$, ja $\varepsilon_1 = \varepsilon_2 = \dots = 0$. Jos merkitään $e_k = (0, \dots, 0, 1, 0, \dots, 0)$ (ykköinen komponentissa k), saadaan tehtyjen oletusten voimassa ollessa yhtälöstä (4.19)

$$\begin{aligned} y_0 &= \varepsilon_0 = e_k \\ y_1 &= \Psi_1 \varepsilon_0 = \Psi_1 e_k \\ &\vdots \\ y_h &= \Psi_h \varepsilon_0 = \Psi_h e_k. \end{aligned}$$

Tästä nähdään, että matriisin Ψ_h k . sarake kuvaa impulssin $\varepsilon_{k0} = 1$ vaikutusta prosessiin y_t ajankohtana $t = h$ eli h aikayksikköä impulssin tapahtuman jälkeen. Vastaavasti matriisin Ψ_h alkio eli *impulssivaste* $\psi_{jk,h}$ kuvaa tämän impulssin vaikutusta prosessin y_t komponenttiin y_{jt} ajankohtana $t = h$. Koska $\Psi_h \rightarrow 0$ geometrisesti, kun $h \rightarrow \infty$, häipyvät nämä vaikutukset ajan kuluessa. Havainnollisen kuvan asiasta saa piirtämällä impulssivasteista $\psi_{jk,h}$, $h = 1, \dots, H$, (tai niiden estimaateista) muodostuva (empiirinen) *impulssivastefunktio* ”tarpeeksi suurella” H :n arvolla (eli $\psi_{jk,h} \approx 0$, kun $h > H$). Näin määritellyjä impulssivasteita nimitetään myös pidemmin *ennustevirheimpulssivasteiksi* erotuksena alempana määriteltävistä vaihtoehtoisista impulssivasteista.

Matriisin Ψ_h avulla voidaan siis tutkia miten impulssin ε_t vaikutus ilmenee prosessissa y_t ajankohtana $t + h$. Tällaisten yksittäisten vaikutusten lisäksi voidaan tutkia myös kumulatiivisia vaikutuksia ja kokonaisvaikutusta käyttäen matriiseja $\sum_{j=0}^h \Psi_j$ ja $\sum_{j=0}^{\infty} \Psi_j = A(1)^{-1}$. Edellistä käytetään, kun innovaatioissa tapahtunut yhden yksikön muutoksesta on kulunut h aikayksikköä ja jälkimmäistä, kun tutkitaan kokonaisvaikutusta.

Edellä kuvatulla impulssivasteanalyysillä on yhteys myös edellisen jakson kausaalisuustarkasteluihin. Lauseen 4.1 perusteella kausaalisuutta suuntaan $x \rightarrow z$ ei ole jos ja vain jos $A_{12}(\mathbf{B}) = 0$ ($A(\mathbf{B}) = [A_{ij}(\mathbf{B})]_{i,j=1,2}$). Jos $\Psi(\mathbf{B}) = [\Psi_{ij}(\mathbf{B})]_{i,j=1,2}$ ositetaan $A(\mathbf{B})$:n ositusta vastaavalla tavalla, nähdään käänteismatriisin määritelmän avulla helposti, että $A_{12}(\mathbf{B}) = 0$ jos ja vain jos $\Psi_{12}(\mathbf{B}) = 0$. Toisin sanoen, kausaalisuutta voidaan tarkastella myös prosessin VMA(∞)-esityksen avulla.

Edellä oletettiin, että innovaatioissa (eli ε_t :n komponenteissa) tapahtuvat muutokset ovat yhden yksikön suuruisia. Jos innovaatioiden varianssit ovat kovin erisuuruisia, voi olla luontevampaa tarkastella yhden hajontayksikön suuruisia muutoksia. Edellä tarkasteltiin vain yksittäisessä innovaatioissa tapahtuvan muutoksen vaikutuksia, mitä voidaan pitää epäloogisena, jos innovaatiot ovat korreloituneita. Innovaatioiden välinen korrelaatio on kuitenkin varsin hankalasti huomioon otettava kysymys. Eräs ehdotettu ratkaisu on ns. ortogonaalisiin eli korreloimattomiin innovaatioihin siirtyminen (normaalissa tapauksessa innovaatiot ovat tällöin jopa riippumaattomia). Ideana on kirjoittaa VMA(∞)-esitys muotoon

$$y_t = \sum_{j=0}^{\infty} (\Psi_j C) C^{-1} \varepsilon_{t-j},$$

jossa alakolmiomatriisi tai vaihtoehtoisesti yläkolmiomatriisi C ($n \times n$) on valittu siten, että $\text{Cov}(C^{-1}\varepsilon_t) = I_n$.¹¹ Muunnetut innovaatiot $C^{-1}\varepsilon_t$ ovat siten korreloimattomia ja varianssiltaan saman suuruisia, joten niihin liittyvät ns. *ortogonaaliset impulssivasteet* (eli matriisin $\Psi_j C$ alkio) välttävät edellä kuvatut ongelmat. Uudeksi

¹¹Koska $\text{Cov}(\varepsilon_t) = \Omega$ on oletettu positiivisesti definiitiksi, tiedetään matriisilaskennasta, että se voidaan esittää tulona $\Omega = CC'$ (ns. Cholesky-hajotelma), jossa C on epäsingulaarinen alakolmio- tai yläkolmiomatriisi.

hankaluudeksi muodostuu kuitenkin se, että muunnosmatriisi C ei ole yksikäsitteinen. Matriisin C kolmiorakenteesta johtuen riippuvat tulokset (mahdollisesti paljonkin) mm. siitä missä järjestyksessä tarkasteltavat aikasarjat esiintyvät vektorissa y_t . Sopivan muunnosmatriisin valinta vaatii (ei-tilastollista) lisäinformaatiota. Tällaisiin rakenteellisia impulssivasteita ei kuitenkaan tarkastella tällä kurssilla.

Impulssivastefunktion estimointi. Tarkastellaan nyt impulssivasteiden estimointia. Olkoon $\hat{A}_1, \dots, \hat{A}_p$ valitun VAR(p)-mallin kerroinmatriisien (rajoittamattomat tai rajoitetut) SU-estimaattorit. Jaksossa 2.3 esitetyn perusteella on matriisit Ψ_j tällöin luonteva estimoida käyttäen yhtälöitä

$$\begin{aligned}\hat{\Psi}_1 &= \hat{A}_1 \\ \hat{\Psi}_2 &= \hat{\Psi}_1 \hat{A}_1 + \hat{A}_2 \\ &\vdots\end{aligned}$$

tai yleisesti

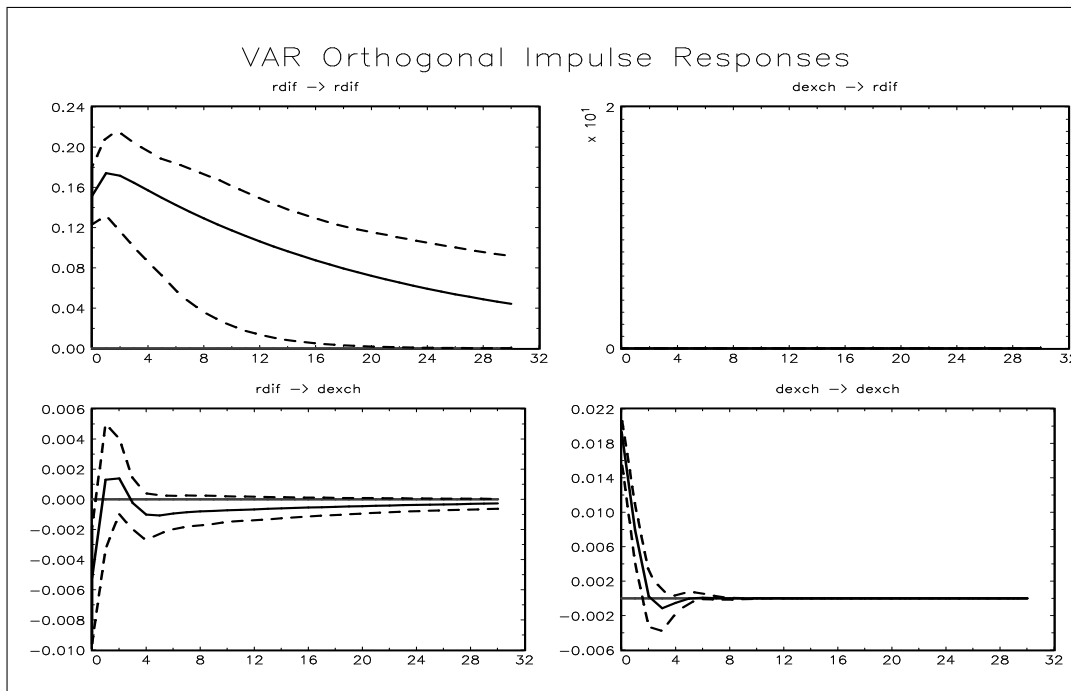
$$\hat{\Psi}_j = \sum_{i=1}^j \hat{\Psi}_{j-i} \hat{A}_i, \quad j = 1, 2, \dots, \quad \hat{\Psi}_0 = I_n, \quad \hat{A}_i = 0, \quad i > p.$$

Vaihtoehtoisesti voidaan kirjoittaa $\hat{\Psi}_j = J \hat{\mathbf{A}}^j J'$, jossa $\hat{\mathbf{A}}$ on yhtälössä (2.8) määritellyn matriisin \mathbf{A} ilmeinen estimaattori ja $J = [I_n : 0 : \dots : 0]$ ($n \times np$) (ks. jakso 2.3).

Koska estimaattorit $\hat{\Psi}_j$ saadaan asympotoottisesti normaaleista estimaattoreista jatkuvasti derivoituvan funktion avulla, voidaan ns. deltamenetelmää käyttäen osoittaa, että estimaattorit $\hat{\Psi}_1, \dots, \hat{\Psi}_H$ (H kiinteä) ovat asympotoottisesti normaalisti jakautuneita (ks. Liite B). Rajajakauman kovarianssimatriisi on monimutkainen, joten sen tarkastelu sivuutetaan. Todetaan kuitenkin, että tapauksessa $H > p$ rajajakauma on välttämättä singulaarinen ja että näin voi käydä myös, kun $H < p$.¹² Tästä syystä impulssivasteille $\psi_{ij,k}$ muodostetut luottamusvälit saattavat epätarkkoja ja hankalia tulkita. Vaihtoehtoinen tapa muodostaa luottamusvälejä perustuu ns. bootstrap-menetelmien käyttöön.

Esimerkki 3.1 (jatkoa). Havainnollistetaan impulsivasteiden käyttöä käyttäen Kuvion 1.1 valuuttakurssi/korkoaineistoon sovitettua rajoitettua VAR(2)-mallia (4.18), jossa Granger-kausalisuus ilmenee korkoerosta valuuttakurssiin, mutta ei päinvastoin. *Oletetaan*, että samanaikainen vaikutus ilmenee myös korkoerosta valuuttakurssiin ja käytetään ortogonaalisia impulssivasteita valiten muunnosmatriisi C yläkolmionmatriisiksi. Kuvioista 4.3 nähdään, että (mallin oletetusta rakenteesta johtuen) valuuttakurssiin tuleva impulssi ei vaikuta tällöin korkoeron tulevaan kehitykseen (ks.

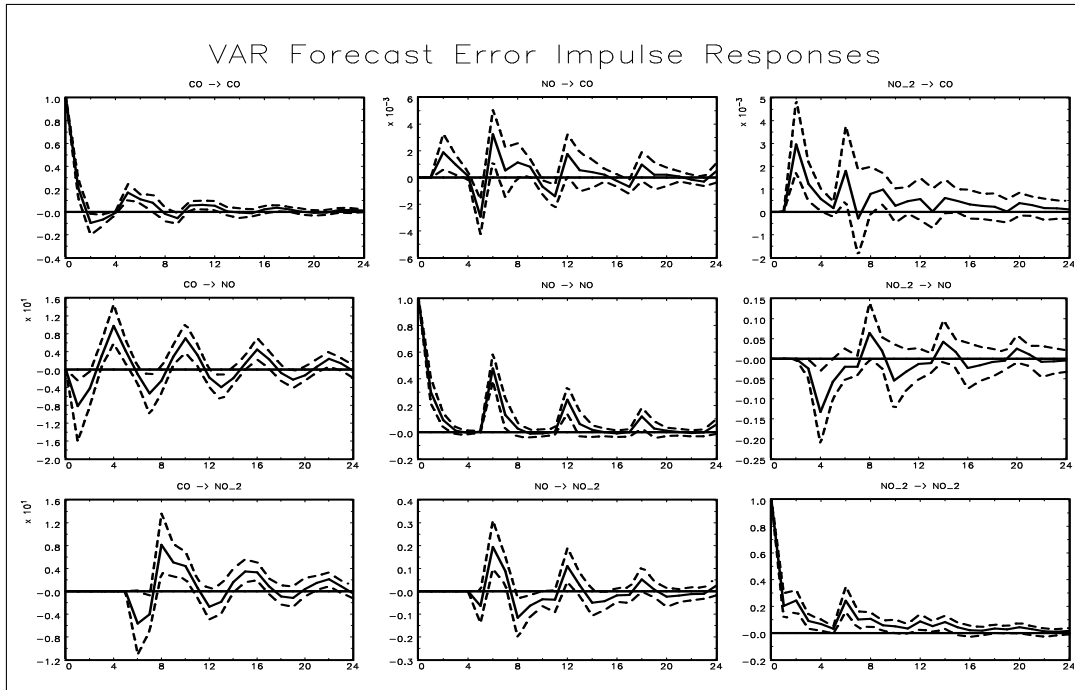
¹²Tapauksessa $H > p$ tämä seuraa siitä, että estimaattoreita $\hat{\Psi}_1, \dots, \hat{\Psi}_H$ on enemmän kuin niiden muodostamisessa käytettäviä estimaattoreita $\hat{A}_1, \dots, \hat{A}_p$.



Kuvio 4.3. Mallista (4.18) lasketut ortogonaaliset impulssivasteet (oletettu samanaikainen vaikutus $rdif_t \rightarrow dexch_t$). Katkoviivat osoittavat likimääräiset 95%:n luottamusvälit, jotka on muodostettu eräällä bootstrap-menetelmän versiolla.

oikealla ylhäällä olevaa kuvion osaa). Se vaikuttaa kuitenkin valuuttakurssin tulevaan kehitykseen n . puolen vuoden ajan (ks. oikealla alhaalla olevaa kuvion osaa). Korkoeroon tuleva impulssi sen sijaan vaikuttaa sekä korkoeron että valuuttakurssin tulevaan kehitykseen. Edellinen vaikutus häviää varsin hitaasti (ks. vasemmalla ylhäällä olevaa kuvion osaa), kun taas jälkimmäinen ilmenee pääasiallisesti muutamien ensimmäisten kuukausien aikana (ks. vasemmalla alhaalla olevaa kuvion osaa). Tuloksia voidaan tulkita niin, että systeemin sopeutuminen muutoksiin tapahtuu valuuttakurssin kautta.

Esimerkki 4.1 (jatkoa). Tarkastellaan toisena esimerkkinä Kuvion 4.1 saasteaineistolle valitun rajoitetun VAR(8)-mallin tavanomaisia ennustevirheimpulssivastefunktioita, jotka on esitetty Kuviossa 4.4. Impulssivasteista nähdään jaksossa 4.4 todetut viivästetyt vaikutukset. Hiilimonoksidin- ja typpioksidin yhtälöissä ilmenevän impulssin vaikutus typpidioksidiin ilmenee vasta viipymiltä 5 ja 6 alkaen, kun taas muissa tapauksissa impulssien vaikutukset ilmenevät jo viipymilla 1 ja 2.



Kuvio 4.4. Kuvion 4.1 saasteaineistolle jaksossa 4.4 valitun rajoitetun VAR(8)-mallin tavanomaiset impulssivastefunktiot. Katkoviivat osoittavat likimääräiset 95%:n luottamusvälit, jotka on muodostettu eräällä bootstrap-menetelmän versiolla.

5 Integroituneet ja yhteisintegroituneet prosessit

5.1 Integroituneet prosessit

Joillakin sovellusalueilla trendejä sisältävät aikasarjat ovat tavallisia, joten niitä ei voida kuvata stationaarisilla prosesseilla (tässä sanaa ”trendi” käytetään hyvin yleisluonteisesti). Tällaisten aikasarjojen muutokset eli differenssit $y_t - y_{t-1} = \Delta y_t$ ($\Delta = 1 - B$) saattavat kuitenkin näyttää stationaarisilta ja siten kohtuullisilta mallintaa stationaarisilla prosesseilla. Jos myös differensseissä on trendejä, voidaan vastaavasti ajatella käytettävän differenssien muutoksia eli toisia differenssejä $\Delta(\Delta y_t) = y_t - 2y_{t-1} + y_{t-2} = \Delta^2 y_t$. Tätä ajatusta voidaan ajatella jatkettavan edelleen korkeampiin differensseihin, mikä on käytännössä kuitenkin harvinaista.

Edellä sanottuun liittyen esitetään seuraava määritelmä.

Määritelmä 5.1. Stokastista prosessia $\{y_t, t = 1, 2, \dots\}$ (mahdollisesti vektori) sanotaan *astetta d olevaksi integroituneeksi prosessiksi eli $I(d)$ -prosessiksi*, jos d kertaa differensoitu prosessi $\{\Delta^d(y_t - E(y_t)), t = 1, 2, \dots\}$ on (asymptoottisesti) stationaarinen, mutta $d-1$ kertaa differensoitu prosessi $\{\Delta^{d-1}(y_t - E(y_t)), t = 1, 2, \dots\}$ ei ole (asymptoottisesti) stationaarinen. Jos y_t on $I(d)$ -prosessi, merkitään $y_t \sim I(d)$ ($d \geq 1$).

Stationaarisen tai asympotoottisesti stationaarisen prosessin yhteydessä käytetään merkintää $y_t \sim I(0)$. $I(d)$ -prosessia ei voida määritellä kaikilla $t \in \mathbb{Z}$, vaan alkuaikajankohdalla on kiinnitettävä (edellä $t = 1$). Kuten jaksossa 2.3, asympotoottisesti stationaarinen prosessi on sellainen, jolla epästationaarisuus aiheutuu alkuarvoista ja häviää, kun $t \rightarrow \infty$ (ks. yhtälön (2.10) yhteydessä esitetyt tarkastelut). Tällä kurssilla rajoitutaan tapaukseen $d = 1$ (ja $d = 0$).

Satunnaiskulku. Yksinkertainen esimerkki $I(1)$ -prosessista on *satunnaiskulku* (random walk), joka määritellään yhtälöllä

$$\Delta y_t = \nu + \varepsilon_t, \quad t = 1, 2, \dots,$$

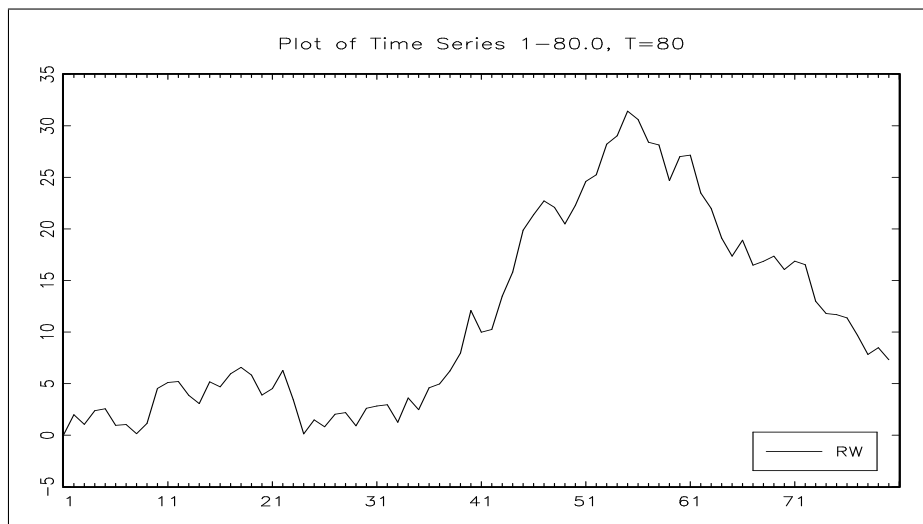
jossa $\varepsilon_t \sim \text{iid}(0, \Omega)$. Tarkastellaan yksiulotteista tapautta ja merkitään $\text{Var}(\varepsilon_t) = \omega^2$. Peräkkäisillä sijoituksilla saadaan

$$y_t = y_0 + \nu t + \sum_{j=1}^t \varepsilon_j, \quad t = 1, 2, \dots .$$

Olettaen alkuarvo y_0 kiinteäksi seuraa tästä (yksityiskohdat jätetään tehtäväksi)

$$E(y_t) = y_0 + \nu t,$$

$$\text{Var}(y_t) = \text{Var}\left(\sum_{j=1}^t \varepsilon_j\right) = \omega^2 t$$



Kuvio 5.1. Simuloitu realisaatio satunnaiskulusta $\Delta y_t = \varepsilon_t$, $y_0 = 0$, $\varepsilon_t \sim \text{nid}(0, 1)$ ($t = 1, \dots, 80$).

ja edelleen

$$\text{Cor}(y_t, y_{t+k}) = \frac{1}{\sqrt{1+k/t}}, \quad k \geq 0.$$

Tästä nähdään, että satunnaiskulku on epästationaarinen eikä tilanne selvästikään muutu, vaikka alkuarvo-oletusta muutettaisiin. Koska $\Delta y_t \sim \text{iid}(\nu, \omega^2)$ on stationaarinen, on satunnaiskulku I(1)–prosessi.

Edellä esitetyistä laskelmista havaitaan, että satunnaiskululle pätee $\text{Var}(y_t) \rightarrow \infty$ ja $\text{Cor}(y_t, y_{t+k}) \rightarrow 1$ kaikilla (myös suurilla) $k > 0$, kun $t \rightarrow \infty$. Satunnaiskulkua noudattavat aikasarjat ovat siten erittäin voimakkaasti autokorreloituneita ja niissä on vaeltelevasta ilmeestä aiheutuvia trendinomaisia piirteitä. Oheinen kuvio havainnollistaa tätä. Kuvion satunnaiskulussa on $y_0 = \nu = 0$, joten edellä oletettu deterministinen trenditermi $y_0 + \nu t$ häviää prosessista. Koska $y_t = \sum_{j=1}^t \varepsilon_j$ on satunnainen, sanotaan trendiä usein stokastiseksi. Jos $\nu \neq 0$, ilmenee tällainen stokastinen trendi deterministisen lineaarisen trendin $y_0 + \nu t$ ympärillä. Edellä esitetyt tulokset ja ominaisuudet pätevät pääpiirteissään myös yleisemmille I(1)–prosesseille, joista seuraavassa esimerkki.

Autoregressiivinen yksikköjuuriprosessi. Tarkastellaan yhtälössä (2.6) määritellyä (yksiulotteista) AR(p)–prosessia, kun $t = 1, 2, \dots$. Tämä prosessi voidaan kirjoittaa $a(\mathbf{B})x_t = \varepsilon_t$, jossa $a(\mathbf{B}) = 1 - a_1\mathbf{B} - \dots - a_p\mathbf{B}^p$.¹³ Liitteen A.4 yhtälöstä (A.2) nähdään, että polynomi $a(\mathbf{B})$ voidaan kirjoittaa $a(\mathbf{B}) = \Delta - \phi\mathbf{B} - \phi_1\Delta\mathbf{B} - \dots - \phi_{p-1}\Delta\mathbf{B}^{p-1}$ ($\phi = -a(1)$), joten prosessille x_t saadaan esitys

$$\Delta x_t = \phi x_{t-1} + \phi_1 \Delta x_{t-1} + \dots + \phi_{p-1} \Delta x_{t-p+1} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (5.1)$$

jossa $\varepsilon_t \sim \text{iid}(0, \omega^2)$.

Oletetaan nyt, että polynomi $a(z)$ toteuttaa seuraavan ehdon:

$$\text{Jos } a(z) = 0, \text{ niin } |z| > 1 \text{ tai } z = 1. \quad (5.2)$$

Toisin sanoen, polynomien $a(z)$ nollakohdat eli juuret sijaitsevat kompleksitasossa yksikköympyrän kehän ulkopuolella tai pisteessä $z = 1$. Jos kaikki juuret sijaitsevat yksikköympyrän kehän ulkopuolella, on x_t (ainakin) asymptoottisesti stationaarinen. Oletetaan nyt, että yksikköjuuria on täsmälleen yksi. Polynomien tunnetuista ominaisuuksista (ns. algebran peruslauseesta) seuraa, että tällöin $a(z) = (1-z)b(z)$, jossa $b(z)$ on (korkeintaan) astetta $p-1$ oleva polynomi ja sen juuret ovat välttämättä yksikköympyrän kehän ulkopuolella (ks. Liite A.4). On selvää, että tämä on yhtäpitävää sen kanssa, että yhtälössä (5.1) $\phi = 0$ ja, kun $\phi = 0$, on edellä mainittu polynomi $b(z) = 1 - \phi_1 z - \dots - \phi_{p-1} z^{p-1} := \phi(z)$. Näin ollen, $\Delta x_t \sim \text{I}(0)$ ja, koska polynomien $\phi(z)$ juuret sijaitsevat yksikköympyrän kehän ulkopuolella, voidaan löytää

¹³Tässä kuten usein myöhemminkin käytetään merkintää x_t , kun prosessissa ei oleteta olevan vakiota tai muita deterministisiä termejä.

alkuarvot $\Delta x_{t-1}, \dots, \Delta x_{t-p+1}$ siten, että Δx_t on stationaarinen (ks. jakso 2.3). Tällöin prosessi Δx_t voidaan kirjoittaa myös muotoon $\Delta x_t = \phi(\mathbf{B})^{-1} \varepsilon_t$, joka on edellä tarkastellun satunnaiskulun yleistys tapaukseen, jossa innovaation ε_t paikalla on stationaarinen AR(p-1)-prosessi $\phi(\mathbf{B})^{-1} \varepsilon_t$.

5.2 Yhteisintegroituneet prosessit

Tarkastellaan nyt moniulotteista I(1)-prosessia $y_t = (y_{1t}, \dots, y_{nt})$ ($t \geq 1$), jonka differenssoituun versioon $\Delta y_t = (\Delta y_{1t}, \dots, \Delta y_{nt})$ voidaan periaatteessa soveltaa (asymptoottisesti) stationaaristen prosessien teoriaa. Tällaisten differenssien käyttäminen voi kuitenkin aiheuttaa hankaluuksia, joiden havainnollistamiseksi tarkastellaan kaksiulotteista tapausta, jossa $y_{1t} \sim I(1)$ ja $y_{2t} \sim I(0)$. Tällöin $y_t \sim I(1)$, mutta toinen y_t :n komponentti on I(0) ilman differenssointia eli differenssointi on sen osalta turhaa. Oletetaan edelleen, että vektori $(\Delta y_{1t}, \Delta y_{2t})$ noudattaa (stabiilia) VAR(1)-prosessia

$$\begin{bmatrix} 1 - a_{11}\mathbf{B} & 0 \\ 0 & 1 - a_{22}\mathbf{B} \end{bmatrix} \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

jossa $|a_{ii}| < 1$ ($i = 1, 2$) ja virhetermi on tavanomainen iid $(0, \Omega)$. Jos analyysi perustetaan differenssoituun prosessiin $(\Delta y_{1t}, \Delta y_{2t})$ tulee mallin jälkimmäisestä yhtälöstä $(1 - a_{22}\mathbf{B}) \Delta y_{2t} = (1 - \mathbf{B}) \varepsilon_{2t}$ eli päädytään VARMA(1,1)-malliin. VAR(p)-malli on siten väärin spesifioitu eikä toimi approksimaationakaan valittiinpa aste p kuinka suureksi tahansa. Tämä johtuu siitä, että Δy_{2t} :n ARMA(1,1)-malli ei ole käännettävä ja sama pätee prosessin $(\Delta y_{1t}, \Delta y_{2t})$ VARMA(1,1)-malliin (ks. jakson 2.4 lopussa mainittu käännettävyysehto). Samanlainen tilanne syntyy monimutkaisemmissa malleissa, jos joku komponentti differensoidaan turhaan. Ongelmalliseksi tämän tekee se, että ei-käännettävien VARMA-mallien soveltaminen on hankalaa ja niiden teoria on kehittymätön.

Edellä kuvattu tilanne voidaan ilmaista formaalisti myös niin, että $y_t \sim I(1)$, mutta on olemassa lineaarikombinaatio, jolle pätee $\beta' y_t \sim I(0)$. Edellä vektori β oli tyyppiä $\beta = (0, \dots, 0, 1, 0, \dots, 0)$, mutta paljon tärkeämpi ja mielenkiintoisempi tilanne syntyy, kun ainakin kaksi β :n komponenttia poikkeaa nolasta. Tähän liittyen esitetään seuraava määritelmä.

Määritelmä 5.2. I(1)-prosessia $\{y_t, t = 1, 2, \dots\}$ ($n \times 1$) sanotaan *yhteisintegroitu-neeksi*, jos on olemassa nolasta poikkeava parametrivektori β ($n \times 1$) siten, että $\beta' y_t \sim I(0)$.¹⁴

Huomaa, että määritelmä sallii mahdollisuuden $\beta = (0, \dots, 0, 1, 0, \dots, 0)$, joka joissakin määritelmässä suljetaan pois. Parametrivektoria β sanotaan *yhteisintegroitu-vuusvektoriksi* (yi-vektoriksi; cointegrating vector). Yi-vektoreita voi olla enemmän kuin yksi. Jos on olemassa r lineaarisesti riippumatonta (vapaata) yi-vektoria

¹⁴Yleisemmin I(d)-prosessin y_t sanotaan olevan yhteisintegroitunut astetta $d - c$, jos on olemassa nolasta poikkeava parametrivektori β ($n \times 1$) siten, että $\beta' y_t \sim I(d - c)$.

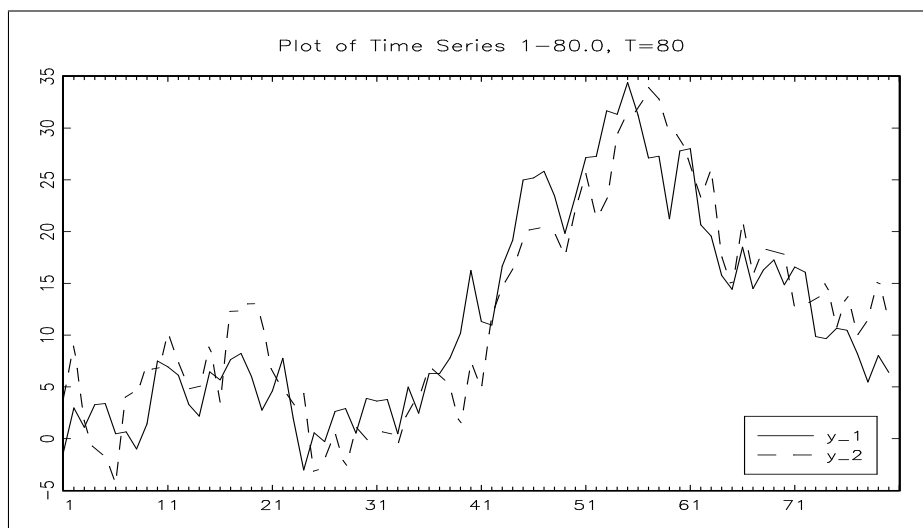
β_1, \dots, β_r , niistä voidaan muodostaa matriisi $[\beta_1 \ \dots \ \beta_r]$, jonka astetta r sanotaan *yhteisintegroituvuusasteeksi* (yi-asteeksi; cointegrating rank). Edellä esitettyssä määritelmässä ajatellaan, että $r \leq n - 1$. Joissakin yhteyksissä voi kuitenkin olla kätevää sallia raja-tapaus $r = n$ eli $y_t \sim I(0)$. Jatkossa merkinnällä β tarkoitetaan (myös) matriisia $[\beta_1 \ \dots \ \beta_r]$, jota nimitetään *yhteisintegroituvuusmatriisiksi* (yi-matriisiksi).

Kaksiulotteinen esimerkki. Yhteisintegroituvuuden käsitteen havainnollistamiseksi tarkastellaan kaksiulotteista prosessia

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} b \\ 1 \end{bmatrix} \tau_t + \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \end{bmatrix}, \quad t = 1, 2, \dots, \quad (5.3)$$

jossa $(\xi_{1t}, \xi_{2t}) = \xi_t \sim \text{iid}(0, \Omega)$, $\tau_t = \sum_{j=1}^t \eta_j$ ja $\eta_t \sim \text{iid}(0, \sigma^2)$. Oletetaan lisäksi, että $\sigma^2 > 0$ ja että kovarianssmatriisi Ω on positiivisesti definiitti. Tällöin y_{1t} ja y_{2t} ovat selvästikin $I(1)$ -prosesseja, joten niiden tuottamat aikasarjat vaeltelevat ajan kuluessa kasvavalla varianssilla kuten Kuvion 5.1 esimerkissä. Prosessien y_{1t} ja y_{2t} epästationaarisuuden selittää yksiulotteinen satunnaiskulku τ_t , joka häviää lineaarikombinaatiosta $y_{1t} - by_{2t} = \mu_1 - b\mu_2 + \xi_{1t} - b\xi_{2t}$. Koska tämä lineaarikombinaatio on stationaarinen (jopa iid), se ei käyttäydy samalla tavalla kuin y_{1t} ja y_{2t} , vaan vaihtelee vakiovarianssilla kiinteän odotusarvonsa $\mu_1 - b\mu_2$ ympärillä. Yhtälö (5.3) määrittelee näin ollen yhteisintegroituneen prosessin, jonka yi-vektori on $(1, -b)$.

Edellä todettu merkitsee, että lineaarikombinaatio $y_{1t} - by_{2t}$ määrittelee prosessien



Kuvio 5.2. Simuloidut realisaatiot prosessista (5.3), kun $b = 1$, $\mu_1 = \mu_2 = 0$, $\xi_{1t} = 0.8\xi_{1,t-1} + \varepsilon_{1t}$, $\xi_{2t} = 0.4\xi_{2,t-1} + \varepsilon_{2t}$ ja $\varepsilon_{1t} \sim \text{nid}(0, 1)$ ja $\varepsilon_{2t} \sim \text{nid}(0, 2)$ ovat riippumattomia ($t = 1, \dots, 80$). Satunnaiskulku τ_t on sama kuin Kuviossa 5.1 ja riippumaton prosessista $(\varepsilon_{1t}, \varepsilon_{2t})$.

y_{1t} ja y_{2t} välille erittäin voimakkaan lineaarisen riippuvuuden, joka pakottaa ne vaeltelemaan ”käsi kädessä”. Ehkä konkreettisimmin tätä ilmenee, kun $b = 1$ ja $\mu_1 = \mu_2$, jolloin lineaarikombinaatio on $y_{1t} - y_{2t}$ on stationaarinen odotusarvona nolla. Vaikka prosessit y_{1t} ja y_{2t} vaeltelevatkin villisti ajan kuluessa, estää yhteisintegroituvuus niitä loitontumasta kauas toisistaan. Kuvio 5.2 havainnollistaa tätä.

Konkreettisena esimerkkinä edellä todetusta voidaan tarkastella koko kansantalouden tuloja (Y) ja kulutusta (C), joiden aikasarjat ovat tyypillisesti epästationaarisia. Tuntuu selvältä, että ”pitkällä aikavälillä” kulutuksen ja tulojen suhteen täytyy olla tasapainossa, mitä kuvaa ehto $Y/C = a$ (= vakio). Tästä saadaan logaritmoimalla yhtälö $\log(Y) - \log(C) = a$, jota voidaan pitää malliin (5.3) liittyvän yi-relaation ”idealisoituna” vastineena, kun parametri $b = 1$. Tämä havainnollistaa, miksi yhteisintegroituvuutta pidetään taloudellisissa sovelluksissa taloustieteen pitkän aikavälin relaatioiden tai tasapainoehtojen tilastollisena vastineena. Se havainnollistaa myös sitä, että mallissa (5.3) lineaarikombinaatio $y_{1t} - by_{2t} - (\mu_1 - b\mu_2)$ voidaan tulkita poikkeamaksi tällaisesta pitkän aikavälin relaatiosta tai tasapainosta. Yleisessä tapauksessa, jossa yi-vektoreita voi olla useita, nämä poikkeamat ovat $\beta'y_t$ tai $\beta'y_t - E(\beta'y_t)$.

Yhtälössä (5.3) esiintyvä satunnaiskulku on esimerkki tarkasteltavan kaksiulotteisen I(1)-prosessin *yhteisestä trendikomponentista*. Sen innovaation η_t suhteesta innovaatioon ξ_t ei oletettu mitään. Äärimmäiset vaihtoehdot ovat, että $\{\eta_t\}$ ja $\{\xi_t\}$ ovat riippumattomia tai että η_t on ξ_t :n lineaarikombinaatio. Jälkimmäiseen palataan myöhemmin yleisemmän mallin yhteydessä.

Yi-vektoreiden monikäsiteisyys. Todetaan vielä, että ilman lisäehtoja yi-vektorit eivät ole yksikäsitteisiä. Esimerkiksi mallissa (5.3) $(c, -cb)$ on yi-vektori kaikilla $c \neq 0$. Yleisesti, jos yi-aste on r ja $\beta = [\beta_1 \ \cdots \ \beta_r]$ ($n \times r$) on lineaarisesti riippumattomien yi-vektoreiden matriisi, niin matriisin $\beta\zeta$ sarakkeet ovat yi-vektoreita olipa ζ mikä tahansa epäsingulaarinen $r \times r$ matriisi. (Perustelu: $\beta'y_t \sim I(0) \Rightarrow (\beta\zeta)'y_t = \zeta'\beta'y_t \sim I(0)$ ja $r(\beta\zeta) = r(\beta) = r$.) Yi-matriisin β sarakkeiden virittämä \mathbb{R}^n :n aliavaruus eli ns. *yhteisintegroituvuusavaruus* (yi-avaruus; cointegrating space) on sen sijaan yksikäsitteinen. Mallissa (5.3) yi-vektorin edellisen komponentin rajoittaminen ykköseksi takaa yksikäsiteisyyden. Vastaaviin yleisiin ehtoihin palataan myöhemmin.

5.3 Yhteisintegroitunut VAR(p)-prosessi ja virheenkorjausmalli

Tarkastellaan n -ulotteista VAR(p)-prosessia

$$x_t = A_1x_{t-1} + \cdots + A_px_{t-p} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (5.4)$$

jossa $\varepsilon_t \sim \text{iid}(0, \Omega)$. Merkitään $A(\mathbf{B}) = I_n - A_1\mathbf{B} - \dots - A_p\mathbf{B}^p$. Liitteen A.4 tulosta (A.4) käyttäen saadaan esitys

$$A(\mathbf{B}) = A(1)\mathbf{B} + \left(I_n - \sum_{j=1}^{p-1} \Lambda_j \mathbf{B}^j \right) \Delta,$$

jossa $\Lambda_j = -A_{j+1} - \dots - A_p$. Prosessin x_t VAR-yhtälö (5.4) voidaan siten kirjoittaa

$$\Delta x_t = \Theta x_{t-1} + \sum_{j=1}^{p-1} \Lambda_j \Delta x_{t-j} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (5.5)$$

jossa $\Theta = -A(1)$ (vrt. (5.1)).

Oletetaan nyt, että stationaarisuus (tai stabiilisuusehto) $\det(A(z)) \neq 0$, $|z| \leq 1$, ei päde, vaan polynomilla $\det(A(z))$ on ainakin yksi yksikköjuuri eli juuri pisteessä $z = 1$ (muuta juuria yksikköympyrän kehällä ei sallita). Tällöin $\det(A(1)) = 0$, joten matriisi Θ on vajaa-asteinen. Olkoon $r < n$ Θ :n aste eli $r(\Theta) = r$. Tunnetun matriisilaskennan tuloksen mukaan matriisi Θ voidaan esittää tulona¹⁵

$$\Theta = \alpha\beta', \quad (5.6)$$

jossa α ($n \times r$) ja β ($n \times r$) ovat astetta r eli $r(\alpha) = r(\beta) = r$. Sijoittamalla tämä yhtälöön (5.5) saadaan

$$\Delta x_t = \alpha\beta' x_{t-1} + \sum_{j=1}^{p-1} \Lambda_j \Delta x_{t-j} + \varepsilon_t, \quad t = 1, 2, \dots \quad (5.7)$$

Yhtälön (5.7) määrittelemää mallia sanotaan *virheenkorjausmalliksi* tai prosessin x_t *virheenkorjausesitykseksi*. Nimitys perustuu seuraavaan ajatteluun. Oletetaan, että $x_t \sim I(1)$ ja myös, että $E(x_t) = 0$ kaikilla t . Tällöin $\Delta x_t \sim I(0)$ ja, koska myös $\varepsilon_t \sim I(0)$, seuraa yhtälöstä (5.7), että $\beta' x_{t-1} \sim I(0)$. Prosessi x_t on siis yhteisintegroitunut. Matriisin β sarakkeet ovat yi-vektoreita ja niiden määrittelemien lineaarikombinaatioiden $\beta' x_{t-1}$ poikkeamat odotusarvostaan nollassa voidaan tulkita poikkeamiksi prosessin x_t komponenttien välisestä pitkän aikavälin relaatioista tai tasapainoehdoista (vrt. edellisen jakson esimerkki ja huomaa, että myöhemmin malli yleistetään tapaukseen, jossa $E(\beta' x_{t-1}) \neq 0$ sallitaan).

Edellä mainitut poikkeamat tasapainosta ajatellaan usein ”virheinä”. Näiden virheiden odotusarvo on nolla, mikä vastaa prosessin x_t olemista (hypoteettisessa) tasapainotilassa. Virheiden stationaarisuudesta seuraa, että ne vaihtelevat odotusarvonsa ympärillä vakiovarianssilla ja poiketessaan odotusarvostaan palaavat sinne takaisin (suhteellisen) lyhyen ajan kuluessa. Toisin sanoen, syntyvät virheet korjautuvat

¹⁵Mainittu tulos sanoo, että mikä tahansa astetta k oleva $n \times n$ matriisi A voidaan kirjoittaa tulona $A = BC'$, jossa B ($n \times k$) ja C ($n \times k$) ovat astetta k . Jos A on symmetrinen, voidaan valita $B = C$.

ajan mittaan ja prosessin x_t vaeltelevat I(1)–komponentit palaavat tasapainoon (tai lähelle sitä).

Edellä kuvattu virheiden korjausmekanismi voidaan nähdä mallissa (5.7), jossa virheet $\beta'x_{t-1}$ vaikuttavat prosessin x_t muutoksiin Δx_t ja prosessin dynamiikan mukaisesti edelleen tuleviin muutoksiin. Kuvattu virheiden korjaus tapahtuu lopulta näiden muutosten seurauksena. Virheiden korjautumisen dynamiikkaa mallissa (5.7) kuvaavat parametrit α ja $\Lambda_1, \dots, \Lambda_{p-1}$. Näistä α on tässä suhteessa keskeinen, sillä se liittyy suoraan virheisiin $\beta'x_{t-1}$ ja kuvaa siten suoraan virheen korjaamiseen liittyvää sopeutumisprosessia. Tästä syystä parametria α sanotaan usein *sopeutumismatriisiksi* (adjustment matrix), joskin nimityksiä *latausmatriisi* (loadings matrix) ja *painomatriisi* (weight matrix) käytetään myös.

Grangerin esityslause. Edellä esitetyissä tarkasteluissa oletettiin, että $x_t \sim I(1)$. Tämän oletuksen yhteys yksikköjuurioletukseen $\det A(1) = 0$ ja yhtälöön (5.7) kaippaa kuitenkin täsmenämistä. Tämä on seuraavassa esitettävän ns. Grangerin esityslauseen sisältö. Esitetään ensin tarvittavat oletukset ja merkinnät.

Jos A on astetta m oleva $n \times m$ matriisi ($m < n$), niin merkinnällä A_\perp tarkoitetaan astetta $n - m$ olevaa $n \times (n - m)$ matriisia, jolla on ominaisuus $A'_\perp A = 0$ eli matriisien A_\perp ja A sarakkeet ovat ortogonaaliset (kohtisuorassa toisiaan vastaan). Tällaista (ei-yksikäsitteistä) matriisia sanotaan usein matriisiin A ortogonaaliseksi komplementiksi ja sen olemassaolo seuraa tunnetuista lineaarialgebran tuloksista. Asetetaan nyt seuraava oletus, jossa merkinnät Θ ja $A(z)$ ovat kuten yhtälössä (5.5) ja sitä edeltävässä keskustelussa (vrt. oletus (5.2) ja sitä seuraava keskustelu).

Oletus 5.1. (i) Jos $\det A(z) = 0$, niin $|z| > 1$ tai $z = 1$.

(ii) $r(\Theta) = r$, $0 \leq r < n$, joten yhtälö (5.6) pätee.

(iii) Kohdassa (i) yksikköjuurien ($z = 1$) lukumäärä on täsmälleen $n - r$.

Ensimmäinen oletus takaa, että prosessi x_t voidaan muuntaa (asymptoottisesti) stationaariseksi differensoimalla tarpeeksi monta kertaa. Se sulkee pois muut yksikköjuuret kuin $z = 1$ (eli kompleksiset ja -1). Yhdessä kahdesta ensimmäisestä oletuksesta seuraa, että yksikköjuurien lukumäärä on vähintään $n - r$. Kolmas oletus täsmenää yksikköjuurien lukumääräksi $n - r$. Voidaan osoittaa, että kahden ensimmäisen oletuksen voimassa ollessa kolmas on yhtäpitävä sen kanssa, että

$$r \left(\alpha'_\perp \sum_{j=1}^{p-1} \Lambda_j \beta_\perp \right) = n - r, \quad (5.8)$$

jossa matriisi $\Lambda_1, \dots, \Lambda_{p-1}$ ovat kuten yhtälössä (5.5). Toisin sanoen, $(n - r) \times (n - r)$ matriisi $\alpha'_\perp \sum_{j=1}^{p-1} \Lambda_j \beta_\perp$ on epäsingulaarinen.

Nyt voidaan esittää seuraava tulos, jonka todistus sivuutetaan.

Lause 5.1. (Grangerin esityslause) (i) Tarkastellaan VAR(p)–prosessia (5.4). Jos Oletus 5.1 on voimassa, niin $\Delta x_t \sim I(0)$ ja $\beta'x_t \sim I(0)$, joten sopivilla alkuarvovalinnoilla Δx_t ja $\beta'x_t$ ovat stationaarisia (ks. yhtälö (5.7)).

(ii) Prosessilla x_t on esitys

$$x_t = C \sum_{j=1}^t \varepsilon_j + \sum_{j=0}^{\infty} C_j \varepsilon_{t-j} + P_{\beta_{\perp}} x_0, \quad t = 1, 2, \dots, \quad (5.9)$$

jossa $C = \beta_{\perp} \left(\alpha'_{\perp} \sum_{j=1}^{p-1} \Lambda_j \beta_{\perp} \right)^{-1} \alpha'_{\perp}$ (ks. (5.8)), $P_{\beta_{\perp}} = \beta_{\perp} (\beta'_{\perp} \beta_{\perp})^{-1} \beta'_{\perp}$ ja $\|C_j\| \rightarrow 0$ geometrisesti, kun $j \rightarrow \infty$.

Lauseen tulos osoittaa, että Oletuksen 5.1 voimassa ollessa virheenkorjausmallin (5.7) yhteydessä tehty oletus $x_t \sim I(1)$ on voimassa ja että $\beta'x_t \sim I(0)$ kuten aiemmin pääteltiin. Huomaa, että nämä tulokset eivät päde ilman Oletusta 5.1(iii) (tai (5.8)). Esimerkiksi, kun $p = 2$ ja $A(z) = I_n - 2I_n z + I_n z^2 = (1 - z)^2 I_n$, ovat Oletukset 5.1 (i) ja (ii) voimassa ja, koska $A(1) = 0$, on jälkimmäisessä $r = 0$. Yksikköjuuria on kuitenkin enemmän kuin n , joten Oletus 5.1(iii) ei ole voimassa ja Lauseen 5.1 tulos ei päde, sillä $x_t \sim I(2)$.

Tulokset $\Delta x_t \sim I(0)$ ja $\beta'x_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j} \sim I(0)$ nähdään suoraan yhtälöstä (5.9), joka on tulkinnallisesti samanlainen kuin sen yksinkertainen vastine (5.3). Yhtälön (5.9) oikean puolen ensimmäinen termi on satunnaiskulku, joka määrittelee prosessin x_t yhteiset trendikomponentit. Matriisin C (vajaa-asteisen) rakenteen vuoksi nämä yhteiset trendikomponentit on luonteva määrittellä käyttäen satunnaiskulkua $\alpha'_{\perp} \sum_{j=1}^t \varepsilon_j$ $((n - r) \times 1)$. Yhteisiä trendikomponentteja on siten $n - r$ ja stationaarisia komponentteja r (eli $\beta'x_t$). Huomaa kuitenkin, että yhteiset trendikomponentit voidaan määrittellä myös muilla tavoin (esimerkiksi $\beta'_{\perp} x_t$).

Impulssivasteet. Jaksoissa 2.3 ja 4.6 tarkasteltiin stationaarisen VAR(p)–prosessin impulssivasteita, joiden määrittely perustui prosessin lineaarisen VMA(∞)–esityksen kerroinmatriiseihin Ψ_j . Nyt tarkasteltavassa I(1)–tapauksessa samanlaista lineaarista VMA(∞)–esitystä ei ole, mutta impulssivasteet voidaan kuitenkin määrittellä samaan tapaan kuin stationaarisessa tapauksessakin. Jaksossa 2.3 todettiin, että kerroinmatriisit Ψ_j voidaan laskea rekursiivisesti yhtälöistä

$$\Psi_j = \sum_{i=1}^j \Psi_{j-i} A_i, \quad \Psi_0 = I_n, \quad j = 1, 2, \dots,$$

jossa $A_i = 0$, kun $i > p$. Tämä laskukaava on hyvin määritelty, vaikka lähtökohtana olisikin yhtälön (5.4) mukainen yhteisintegroitunut VAR(p)–prosessi x_t . Kerroinmatriisien Ψ_j yhteys innovaatioihin ε_t nähdään lausumalla x_t peräkkäisiä sijoituksia

käyttäen innovaatioiden $\varepsilon_1, \dots, \varepsilon_t$ ja alkuarvojen x_0, \dots, x_{-p+1} lineaarisena funktiona (yksityiskohdat, jotka näkyvät helpoiten VAR(1)–yhtälöä (2.8) käyttäen, jätetään tehtäväksi). Toisin kuin stationaarisessa tapauksessa impulssivasteet eli kerroinmatriisien Ψ_j alkioit eivät konvergoi nyt nollaan j :n kasvaessa rajatta, vaan (yleensä) kohti nollasta poikkeavaa vakiota. Tämä tulee ymmärrettäväksi Grangerin esityslauseen yhtälöstä (5.9), jonka ensimmäisessä termissä oikealla innovaation ε_j edessä on kerroinmatriisi C . Edellä sanotusta seuraa myös, ettei jaksossa 4.6 määriteltyä impulssien kokonaisvaikutusta voida määrittellä I(1)–prosessien tapauksessa. Muuten impulssivasteita voidaan kuitenkin käyttää kuten jaksossa 4.6 esitettiin.

Determinististen termien lisääminen malliin. Kuten yhtälöstä (5.9) nähdään, ovat prosessin x_t stationaaristen komponenttien $\beta'x_t$ odotusarvot nolliä, mikä on usein epärealistista. Nollasta poikkeava odotusarvo voidaan ottaa huomioon tarkastelemalla prosessia

$$y_t = \mu + x_t, \quad t = 1, 2, \dots, \quad (5.10)$$

jossa x_t on kuten yhtälössä (5.4). Oletetaan lisäksi, että x_t toteuttaa Grangerin esityslauseen ehdot, jolloin prosessille y_t pätee $y_t \sim I(1)$, mutta $\Delta y_t \sim I(0)$ ja $\beta'y_t \sim I(0)$. Kertomalla yhtälö (5.10) vasemmalta polynomimatriisilla $A(B)$ ja menettelemällä kuten yhtälöitä (5.5) ja (5.7) johdettaessa saadaan prosessille y_t virheenkorjausesitys

$$\Delta y_t = \alpha(\beta'y_{t-1} - \nu) + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (5.11)$$

jossa $\nu = \beta'\mu$ ja muut merkinnät ovat kuten aikaisemminkin. Tämän virheenkorjausesityksen tulkinta poikkeaa prosessin x_t virheenkorjausesityksen tulkinnasta vain siinä, että stationaaristen lineaarikombinaatioiden $\beta'y_{t-1}$ odotusarvo ν voi poiketa nollassa. Sijoittamalla yhtälöön (5.9) $x_t = y_t - \mu$ saadaan prosessille y_t ilmeinen vastaavanlainen esitys.

Edellä esitettyä yleisempi malli saadaan, lisäämällä yhtälöön (5.10) lineaarinen aikatrendi eli olettamalla

$$y_t = \mu_0 + \mu_1 t + x_t, \quad t = 1, 2, \dots, \quad (5.12)$$

jossa x_t on kuten edellä. Tällöin yhtälön (5.11) vastineeksi tulee (yksityiskohdat jätetään tehtäväksi)

$$\Delta y_t = \nu_0 + \alpha(\beta'y_{t-1} - \nu_1(t-1)) + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (5.13)$$

jossa $\nu_1 = \beta'\mu_1$ ja vakion ν_0 riippuvuus muista parametreista jätetään täsmentämättä.

Virheenkorjausmallin (5.13) tulkinta on samanlainen kuin sen aikaisempien vastineiden. Nyt lineaarikombinaatiot $\beta' y_{t-1}$ eivät kuitenkaan ole stationaarisia, sillä niiden odotusarvossa on yleensä aikatrendi (tämän näkee sijoittamalla yhtälöön (5.9) $x_t = y_t - \mu_0 - \mu_1 t$). Poikkeuksen muodostaa tapaus, jossa parametri μ_1 toteuttaa ehdon $\mu_1 = \beta_{\perp} \mu_*$ ($\mu_* \neq 0$). Tällöin yhtälössä (5.13) $\nu_1 = 0$, joten malliksi saadaan (ν_0 ei ole sama kuin mallissa (5.13))

$$\Delta y_t = \nu_0 + \alpha \beta' y_{t-1} + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (5.14)$$

Tässäkin tapauksessa prosessissa y_t on kuitenkin aikatrendi (ks. (5.12)). Tämä voidaan todeta myös seuraavasti (yksityiskohdat sivuutetaan). Jos yhtälön (5.12) asemesta lähtökohdaksi otetaan yhtälö (5.14), jossa parametria ν_0 ei voida esittää muodossa $\nu_0 = -\alpha \nu$ (eli yhtälö (5.11) ei päde), päädytään Grangerin esityslauseessa tilanteeseen, jossa prosessissa y_t on aikatrendi. Perustelu on periaatteessa samanlainen kuin jakson 5.1 alkupuolen satunnaiskulkuesimerkissä, mutta algebrallisesti monimutkaisempi.¹⁶

YI-vektoreiden yksikäsitteisyys. Aiemmin todettiin, että yi-vektorit eivät ole yksikäsitteisiä. Tämä näkyy virheenkorjausmallissa (5.7) siten, että $\alpha \beta' = (\alpha \zeta^{-1}) (\beta \zeta)'$ kaikilla epäsingulaarisilla matriiseilla ζ ($r \times r$). Virheenkorjausmalleissa (5.11) ja (5.13) pätee sama, kun suoritetaan myös muunnokset $\nu \rightarrow \zeta \nu$ ja $\nu_1 \rightarrow \zeta \nu_1$. Yksikäsitteisyys, joka on tulkinallisesti miellyttävä, voidaan saavuttaa rajoittamalla yi-vektoreita. Seuraavassa yksinkertainen tapa tehdä tämä.

Ositetaan $y_t = (z_t, q_t)$, jossa z_t on $r \times 1$ ja q_t on $(n - r) \times 1$ vektori, ja tarkastellaan millä ehdoilla (yksikäsitteinen) valinta $\beta' = [I_r : A]$ jollain $r \times (n - r)$ matriisilla A on mahdollinen. Koska $r(\beta) = r$, on matriisissa β ($n \times r$) aina r lineaarisesti riippumatonta riviä ja koska matriisin $\beta \zeta$ (ζ epäsingulaarinen), sarakkeet ovat yi-vektoreita, voidaan valinta $\beta' = [I_r : A]$ tehdä jos ja vain jos matriisin β r ensimmäistä riviä ovat lineaarisesti riippumattomia. Tämä on yhtäpitävää sen kanssa, että $\beta' = [\beta'_{(1)} : \beta'_{(2)}]$, jossa $\beta_{(1)}$ ($r \times r$) on epäsingulaarinen. Rajoittumalla tapaukseen $\beta_{(1)} = I_r$ pakotetaan edellä mainituksi muunnosmatriisiksi $\zeta = I_r$, mistä seuraa esityksen $\beta' = [I_r : A]$ yksikäsitteisyys.

On helppo nähdä, että valinta $\beta_{(1)} = I_r$ on mahdollinen jos ja vain jos prosessi q_t ei ole yhteisintegroitu. Tämä voidaan todeta huomaamalla, että yi-matriisi voidaan kirjoittaa $\beta' = [I_r : A]$ jos ja vain jos vektori $b = [0 \quad a']'$, a $(n - r) \times 1$, ei kuulu yi-avaruuteen eli b ei ole yi-vektori. Tämä on puolestaan yhtäpitävää sen kanssa, että q_t ei voi olla yhteisintegroitu. Käytännössä valinta $\beta' = [I_r : A]$ edellyttää siis tietoa

¹⁶Tässä suhteessa kurssilla käytetty esitystapa poikkeaa joistakin oppikirjoista, joissa lähtökohdانا on virheenkorjausmalli (5.13) (tai joku sen erikoistapaus) eikä yhtälö (5.12) (tai joku sen erikoistapaus).

yhteisintegroituvuuden rakenteesta, jotta ositus $y_t = (z_t, q_t)$ osataan tehdä (mahdollisesti y_t :n komponenttien järjestyksen vaihtamisen jälkeen). Tieto voi perustua taustateoriaan ja/tai siihen, että prosessi q_t määrää prosessin y_t epästationaarisuuden eli toimii yhteisenä trendikomponenttina.

Edellä kuvattu tapa saavuttaa yi -vektoreiden yksikäsitteisyys yleistyy seuraavasti. Jos oletetaan, että $r(c'\beta) = r$ jollain valitulla $n \times r$ matriisilla c , on $\beta_c = \beta(c'\beta)^{-1}$ yksikäsitteinen. Valitsemalla $c' = [I_r : 0]$ saadaan edellä tarkasteltu tapa tästä erikoistapauksena. Tällaisten rajoitteiden yhteydessä puhutaan *yi-vektoreiden normalisoinnista*. Niissä käytetään rajoitteita vain minimimäärä yksikäsitteisyyden saavuttamiseksi. Jos mielekkäitä rajoitteita on tiedossa enemmän (esim. taustateoriaan perustuen), voidaan yi -vektoreita rajoittaa minimiä enemmän. Esimerkki tästä on, kun yksi yi -vektoreista voidaan olettaa tunnetuksi. Yleisemmin voidaan käyttää jaksossa 4.2 esitettyjä lineaarisia rajoitteita (4.10) tai (4.11). Esimerkiksi edellisen tapauksessa yi -vektoreiden oletetaan toteuttavan

$$\beta_i = H_i \delta_i + a_i, \quad i = 1, \dots, r,$$

jossa H_i ($n \times s_i$) ja a_i ($n \times 1$) ovat tunnettuja ja $r(H_i) = s_i$. Näin spesifioitujen yi -vektoreiden yksikäsitteisyys vaatii lisäehtoja, joiden esittäminen on kuitenkin melko mutkikasta ja sivuutetaan.

6 I(1)–prosessien asymptotiikkaa

I(1)–prosesseja sisältävien aikasarjamallien asymptoottinen estimointi- ja testiteoria on tavanomaisesta täysin poikkeava, sillä jaksossa 3 tarkastellun tyyppiset suurten lukujen lait ja keskeiset raja-arvolauseet eivät päde. Tämän havainnollistamiseksi tarkastellaan normaalista satunnaiskulkua (vrt. jakson 5 alku)

$$x_t = x_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots, \quad x_0 = 0,$$

jossa $\varepsilon_t \sim \text{nid}(0, \omega^2)$. Peräkkäisiä sijoituksia käyttäen nähdään, että $x_t = \sum_{j=1}^t \varepsilon_j$. Koska $\varepsilon_t = x_t - x_{t-1}$, saadaan

$$\begin{aligned} \sum_{t=1}^T \varepsilon_t^2 &= \sum_{t=1}^T x_t^2 + \sum_{t=1}^T x_{t-1}^2 - 2 \sum_{t=1}^T x_{t-1} x_t \\ &= x_T^2 + 2 \sum_{t=1}^T x_{t-1}^2 - 2 \sum_{t=1}^T x_{t-1} x_t, \end{aligned}$$

josta voidaan edelleen johtaa identiteetti

$$T^{-1} \sum_{t=1}^T x_{t-1} \varepsilon_t = \frac{1}{2} \left(T^{-1} x_T^2 - T^{-1} \sum_{t=1}^T \varepsilon_t^2 \right).$$

Oikealla

$$T^{-1}x_T^2 = \omega^2 \left(\frac{1}{\omega\sqrt{T}} \sum_{t=1}^T \varepsilon_t \right)^2 \sim \omega^2 \chi_1^2,$$

jossa jakaumatulos seuraa siitä, että suluisissa oleva satunnaismuuttuja noudattaa $\mathbf{N}(0, 1)$ -jakaumaa. Suurten lukujen lain nojalla

$$T^{-1} \sum_{t=1}^T \varepsilon_t^2 \xrightarrow{p} \mathbf{E}(\varepsilon_t^2) = \omega^2,$$

joten edellä todetusta voidaan päätellä tulos (ks. Liite B, Lause B.3)

$$T^{-1} \sum_{t=1}^T x_{t-1} \varepsilon_t \xrightarrow{d} \frac{\omega^2}{2} (\chi_1^2 - 1).$$

Tämä osoittaa, että mikään tavanomainen todennäköisyyslaskennan raja-arvolause ei päde sm:iien x_{t-1} ja ε_t väliselle toiselle otosmomentille. Syynä on satunnaiskulun erittäin voimakas autokorreloituisuus ja varianssin kasvaminen ajan funktiona (ks. jakson 5 alku). Sama tulos saadaan, kun oletetaan $\varepsilon_t \sim \text{iid}(0, \omega^2)$, joten tehty normaalisoletus ei ole olennainen.

Brownin liike. I(1)-prosessien asymptotiikassa käytetään teknisenä apuvälineenä aikajatkuvaa prosessia, jota kutsutaan Brownin liikkeeksi (vaihtoehtoinen nimitys on Wienerin prosessi). Määritelmä on seuraava.

Määritelmä 6.1. Aikajatkuva stokastinen prosessi $B(t) = (B_1(t), \dots, B_n(t))$ on (n -ulotteinen) *Brownin liike*, jos

- (i) $B(t)$:n realisaatiot ovat jatkuvia (todennäköisyydellä yksi)
- (ii) $B(0) = 0$ ja millä tahansa $0 \leq t_1 \leq \dots \leq t_m < \infty$ ja m $B(t)$:n ns. lisäykset $B(t_2) - B(t_1), \dots, B(t_m) - B(t_{m-1})$ ovat riippumattomia ja

$$B(t_i) - B(t_{i-1}) \sim \mathbf{N}(0, (t_i - t_{i-1})\Sigma).$$

Jos $B(t)$ on kuten määritelmässä, merkitään $B(t) \sim \mathbf{BM}(\Sigma)$. Kun $\Sigma = I_n$, puhutaan standardista Brownin liikkeestä. Asettamalla $n = 1$ saadaan vastaava yksiulotteinen tapaus.

Määritelmästä nähdään, että

$$B(t) \sim \mathbf{N}(0, t\Sigma) \quad \text{ja} \quad \text{Cov}(B(t), B(t+u)) = t\Sigma, \quad u > 0.$$

Tästä voidaan päätellä, että Brownin liikkeen jakaumaominaisuudet ovat samanlaiset kuin diskreettiaikaisen satunnaiskulun $x_t = \sum_{j=1}^t \varepsilon_j$, $\varepsilon_t \sim \text{nid}(0, \Sigma)$.

Jos $B(t) \sim \mathbf{BM}(\Sigma)$ ($n \times 1$), niin $B(t)$:n realisaatiot ovat jatkuvia ja siten integroituvia, kun $0 \leq t \leq a < \infty$. Näin ollen voidaan määritellä ns. *integroitu Brownin liike*

$$\int_0^t B(u) du = \left(\int_0^t B_1(u) du, \dots, \int_0^t B_n(u) du \right) := B^{(1)}(t).$$

Myöhemmin tarkastellaan tilannetta, jossa $0 \leq t \leq 1$, johon liittyen voidaan edelleen määritellä ns. *keskistetty Brownin liike*

$$B(t) - \int_0^1 B(u) du := \bar{B}(t) \quad (n \times 1).$$

Stokastinen integraali. Olkoon $B(t)$ reaaliarvoinen Brownin liike ja $X(t)$ jokin aikajatkuva reaalinen prosessi. Jos satunnaismuuttuja

$$\sum_{i=1}^N X(u_i) [B(u_i) - B(u_{i-1})], \quad 0 \leq u_0 < \dots < u_N = t,$$

konvergoi kvadraattisesti, kun $N \rightarrow \infty$ ja $\max_{1 \leq i \leq N} (u_i - u_{i-1}) \rightarrow 0$, merkitään (satunnaista) raja-arvoa symbolilla

$$\int_0^t X(u) dB(u) \quad (t < \infty)$$

ja sitä kutsutaan *stokastiseksi integraaliksi* (nimitystä Ito-integraali käytetään myös). Tällaisen stokastisen integraalin olemassaolo voidaan osoittaa mm. tapauksissa

$$X(t) = B(t), \quad B^{(1)}(t), \quad \bar{B}(t)$$

ja myös silloin, kun $X(t)$ on ei-satunnainen ja (esimerkiksi) jatkuva.

Jos $B(t) = (B_1(t), \dots, B_n(t))$ ja $X(t) = (X_1(t), \dots, X_m(t))$ ja jos stokastiset integraalit $\int_0^t X_i(u) dB_j(u)$ ovat määriteltyjä, voidaan määritellä matriisiarvoinen stokastinen integraali

$$\int_0^t X(u) dB(u)' = \left[\int_0^t X_i(u) dB_j(u) \right], \quad (m \times n).$$

Tämän olemassaolo voidaan osoittaa mm. edellä mainituissa erikoistapauksissa.

Asymptoottisia tuloksia: Yksikköjuuren testaus. Tarkastellaan jakson 5.1 lopussa esitetyn autoregressiivisen yksikköjuuriprosessin (5.1) erikoistapausta

$$\Delta x_t = \phi x_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (6.1)$$

jossa (yksinkertaisuuden vuoksi) $x_0 = 0$ ja $\varepsilon_t \sim \text{nid}(0, \omega^2)$. Kiinnostuksen kohteena on yksikköjuurihypoteesin $H_0 : \phi = 0$ testaaminen stationaarista (tai stabiilia) vaihtoehtoa $\phi < 0$ vastaan. Tuntuu luontevalta perustaa testaus parametrin ϕ PNS-estimaattoriin (tai SU-estimaattoriin)

$$\hat{\phi} = \left(\sum_{t=1}^T x_{t-1}^2 \right)^{-1} \sum_{t=1}^T x_{t-1} \Delta x_t \stackrel{H_0}{=} \left(\sum_{t=1}^T x_{t-1}^2 \right)^{-1} \sum_{t=1}^T x_{t-1} \varepsilon_t,$$

jossa T on tavanomaiseen tapaan havaintojen lukumäärä.

Testiä varten tarkastellaan estimaattorin $\hat{\phi}$ asymptoottista jakaumaa nollahypoteesin voimassa ollessa. Tällöin $x_t = \sum_{j=1}^t \varepsilon_j$ on (normaalinen) satunnaiskulku, joten tavanomaiset raja-arvolauseet eivät päde (ks. jakson alku). Kuten edellä todettiin, ovat x_t :n jakaumaominaisuudet samanlaiset kuin Brownin liikkeen $B(u) \sim \text{BM}(\omega^2)$. Tähän liittyen voidaan osoittaa, että

$$T^{-2} \sum_{t=1}^T x_{t-1}^2 \xrightarrow{d} \int_0^1 B(u)^2 du, \quad B(u) \sim \text{BM}(\omega^2). \quad (6.2)$$

Intuitio tälle on, että vasemmalla oleva summa on ”integraali”, jossa ”integroidaan” Brownin liikkeen diskreetin vastineen neliötä, joten asymptoottisesti saadaan sama jakauma kuin integroimalla neliöityä Brownin liikettä. Tämä intuitio ei välttämättä tee selväksi miten integrointiväliin $[0, 1]$ päädytään. Taustalla on seuraava tempu, joka havainnollistaa myös miksi summa $\sum_{t=1}^T x_{t-1}^2$ täytyy jakaa T :n neliöllä.

Merkitään symbolilla $[a]$ reaaliarvon a kokonaisosaa (esim. $[5.73] = 5$) ja asetetaan ($t = 1, \dots, T$)

$$X_T(u) = T^{-1/2} x_{[Tu]} = \begin{cases} T^{-1/2} x_t, & \text{kun } (t-1)/T \leq u < t/T \\ T^{-1/2} x_T, & \text{kun } u = 1. \end{cases}$$

Ideana on siirtyä diskreettiaikaisesta havaintosarjasta x_1, \dots, x_T jatkuva-aikaiseen funktioon $X_T(u)$, joka on jokaisella T :n arvolla aikajatkuva stokastinen prosessi. $X_T(u)$:n realisaatiot eivät kuitenkaan ole jatkuvia u :n funktioita, vaan niissä on hyppäykset pisteissä $u = (t-1)/T$ ja $u = 1$ (piirrä kuva!)¹⁷ Jokaisella kiinteällä u pätee $\mathbf{E}[X_T(u)] = 0$ ja, koska $[Tu]/T \approx u$,

$$\text{Var}[X_T(u)] = T^{-1} \text{Var} \left(\sum_{j=1}^{[Tu]} \varepsilon_j \right) = \frac{[Tu]}{T} \omega^2 \approx u \omega^2$$

ja

$$X_T(u) = \left(\frac{[Tu]}{T} \right)^{1/2} [Tu]^{-1/2} \sum_{j=1}^{[Tu]} \varepsilon_j \underset{as}{\sim} \mathbf{N}(0, u \omega^2).$$

¹⁷Yhdistämällä epäjatkuvuuspisteet interpoloimalla voitaisiin $X_T(u)$ korvata prosessilla, jonka realisaatiot ovat jatkuvia ja saavat arvoja joukossa $C[0, 1]$.

Lisäksi,

$$T^{-2} \sum_{t=1}^T x_{t-1}^2 = T^{-1} \sum_{t=1}^T (T^{-1/2} x_{t-1})^2 = \int_0^1 X_T(u)^2 du$$

ja, koska prosessin $X_T(u)$ jakaumaominaisuudet ovat samanlaiset kuin Brownin liikkeen $B(u) \sim \mathbf{BM}(\omega^2)$, saadaan tulokselle (6.2) intuitiivinen selitys (matemaattinen selitys vaatii kehittyneitä heikon konvergenssin teoriaa).

Edellä kuvatun kaltaiset ideat ovat myös taustalla tulokselle

$$T^{-1} \sum_{t=1}^T x_{t-1} \varepsilon_t \xrightarrow{d} \int_0^1 B(u) dB(u), \quad B(u) \sim \mathbf{BM}(\omega^2). \quad (6.3)$$

Vasemmalla olevan muuttujan asymptoottinen jakauma johdettiin yksinkertaisemmin jakson alussa, jossa sen todettiin olevan $\omega^2(\chi_1^2 - 1)/2$. Ns. Iton lemman avulla voidaan osoittaa, että tuloksessa (6.3) oleva stokastinen integraali jakautuu kuten satunnaismuuttuja $\omega^2(\chi_1^2 - 1)/2$.

Määrittelemällä standardi Brownin liike $W(u) = \omega^{-1}B(u)$ ja yhdistämällä tulokset (6.2) ja (6.3) pätee H_0 :n voimassa ollessa (ks. Liite B, Lause B.2)

$$\begin{aligned} T\hat{\phi} &= \left(\frac{1}{\omega^2 T^2} \sum_{t=1}^T x_{t-1}^2 \right)^{-1} \frac{1}{\omega^2 T} \sum_{t=1}^T x_{t-1} \varepsilon_t \\ &\xrightarrow{d} \left(\int_0^1 W(u)^2 du \right)^{-1} \int_0^1 W(u) dW(u). \end{aligned}$$

Huomaa, että tässä estimaattori $\hat{\phi}$ täytyy kertoa havaintojen lukumäärällä eikä sen neliöjuurella kuten vastaavissa tavanomaisissa tuloksissa. Koska rajajakauma ei riipu tuntemattomista parametreista, se voidaan taulukoida, joten yksikköjuuren testaus voidaan perustaa suoraan estimaattoriin $\hat{\phi}$. Tavallisempaa on kuitenkin käyttää testauksessa regressio-ohjelmistojen tulostamaa ”t-suhdetta”

$$\tau := \frac{\sqrt{\sum_{t=1}^T x_{t-1}^2}}{\hat{\omega}} \hat{\phi} \xrightarrow{d} \left(\int_0^1 W(u)^2 du \right)^{-1/2} \int_0^1 W(u) dW(u),$$

jossa $\hat{\omega}^2 = (T-1)^{-1} \sum_{t=1}^T (\Delta x_t - \hat{\phi} x_{t-1})^2$ on virhevarianssin ω^2 tavanomainen regressioestimaattori (jakajana voi olla myös T). Rajajakauma on jälleen riippumaton tuntemattomista parametreista ja voidaan taulukoida (taulukoida löytyy useista kirjoista ja tietokoneohjelmistoista). Rajajakauman prosenttipisteiden (tai p-arvojen) laskeminen vaatii numeerisia menetelmiä tai simulointia. Sekä estimaattorin $\hat{\phi}$ että ”t-suhteen” τ asymptoottinen jakauma on ei-normaalinen ja vasemmalle vino. Esimerkiksi jälkimmäisellä 5%:n prosenttipiste on n. -1.94 , kun se $N(0, 1)$ -jakaumalla on -1.645 . Käytettäessä virheellisesti normaalijakaumaa hylätään yksikköjuurihypooteesi siten liian helposti.

Edellä esitetty yksikköjuuritestit voidaan yleistää tapaukseen, jossa on deterministisiä termejä. Yleinen malli tälle on

$$y_t = d_t + x_t, \quad t = 1, 2, \dots,$$

jossa x_t on kuten yhtälössä (6.1) ja deterministinen termi $d_t = \mu$ tai $d_t = \mu_1 + \mu_2 t$. Myös indikaattorimuuttujia kausivaihtelun huomioon ottamiseksi voidaan käyttää. Kertomalla edellä esitetty yhtälö puolittain polynomilla $\Delta - \phi B$ ja järjestämällä termit saadaan tapauksessa $d_t = \mu$ yhtälö (yksityiskohdat jätetään tehtäväksi)

$$\Delta y_t = \nu + \phi y_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots,$$

jossa $\varepsilon_t \sim \text{nid}(0, \omega^2)$ ja $\nu = \phi\mu$. Parametriin ϕ liittyvälle "t-suhteelle" τ_μ pätee yksikköjuurihypoteesin $\phi = 0$ voimassa ollessa

$$\tau_\mu \xrightarrow{d} \left(\int_0^1 \bar{W}(u)^2 du \right)^{-1/2} \int_0^1 \bar{W}(u) dW(u),$$

jossa $\bar{W}(u)$ on standardin Brownin liikkeen keskistetty versio. Rajajakauma, joka voidaan jälleen taulukoida, poikkeaa vakiottomassa tapauksessa saadusta ja on aikaisempaa vinompi vasemmalle (5%:n prosenttipiste on n. -2.86). Asymptoottisen jakauman vinous lisääntyy edelleen, kun "t-suhteeseen" perustuva testi yleistetään lineaarisen trendin $d_t = \mu_1 + \mu_2 t$ tapaukseen (yksityiskohdat sivuutetaan).

Todetaan vielä lopuksi, että edellä esitetyt yksikköjuuritestit voidaan yleistää AR(p)-tapaukseen. Toisin kuin determinististen termien lisääminen malliin, ei AR-asteen kasvattaminen vaikuta testisuureiden asymptoottisiin jakaumiin (olettaen, että prosessi on I(1) nollahypoteesin voimassa ollessa). Esimerkiksi tapauksessa $y_t = \mu + x_t$, jossa $x_t \sim \text{AR}(p)$, käytetään mallia

$$\Delta y_t = \nu + \phi y_{t-1} + \sum_{j=1}^{p-1} \phi_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1, 2, \dots,$$

jossa $\varepsilon_t \sim \text{nid}(0, \omega^2)$, $\nu = -\phi\mu$ (ks. yhtälöä (5.1) edeltävä keskustelu) ja nollahypoteesina on $\phi = 0$. Testisuureena voidaan käyttää parametriin ϕ liittyvää "t-suhdetta" τ_μ , jonka asymptoottinen jakauma on sama kuin edellä. Kaikki asymptoottiset tulokset pätevät lisäksi ilman normaalisuusoletusta (eli $\varepsilon_t \sim \text{iid}(0, \omega^2)$ riittää). Vastavaanlaiset otosmomenttien jakaumakonvergenssit voidaan yleistää myös seuraavassa tarkasteltavaan VAR(p)-tapaukseen.

7 Yhteisintegroituneiden VAR(p)-mallien teoriaa

7.1 Uskottavuusfunktio ja SU-estimointi

Uskottavuusfunktio. Oletetaan, että n :n aikasarjan aineisto voidaan kuvata jaksossa 5.3 tarkasteltua yhteisintegroitunutta VAR(p)-mallia käyttäen. Yksityiskohdat esitetään mallille, jossa deterministisenä terminä on vain vakio, minkä jälkeen

kuvataan lyhyesti miten muiden determinististen termien tapauksissa menetellään. Lähtökohtana on siis yhtälöissä (5.10) ja (5.11) esitetyt mallit. Edellisessä esiintyvän prosessin x_t oletetaan olevan kuten yhtälössä (5.4), jossa polynomimatriisi $A(\mathbf{B})$ toteuttaa myöhemmin jaksossa 5.3 esitetyn Grangerin esityslauseen vaatiman Oletuksen 5.1 ehdot.

Edellä sanotun perusteella saadaan havainnoille malli

$$\Delta y_t = \alpha(\beta' y_{t-1} - \nu) + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1, \dots, T, \quad (7.1)$$

jossa alkuarvot y_{-p}, \dots, y_0 oletetaan havaituiksi ja $\varepsilon_t \sim \text{nid}(0, \Omega)$ (Ω positiivisesti definiitti). Asymptoottiset tulokset pätevät jälleen oletuksella $\varepsilon_t \sim \text{iid}(0, \Omega)$. Oteetaan käyttöön merkinnät $Z_{0t} = \Delta y_t$, $Z_{1t} = [y'_{t-1} \ 1]'$ ja $Z_{2t} = [\Delta y'_{t-1} \ \dots \ \Delta y'_{t-p+1}]'$. Tällöin yhtälö (7.1) voidaan kirjoittaa

$$Z_{0t} = \alpha \beta'_* Z_{1t} + \Lambda Z_{2t} + \varepsilon_t, \quad t = 1, \dots, T, \quad (7.2)$$

jossa $\beta_* = [\beta' : -\nu]'$ ja $\Lambda = [\Lambda_1 : \dots : \Lambda_{p-1}]$.

Menettelemällä aivan kuten jaksossa 4.1 saadaan ehdolliseksi log-uskottavuusfunktiksi¹⁸

$$l(\alpha, \beta_*, \Lambda, \Omega) = -\frac{T}{2} \log \det(\Omega) - \frac{1}{2} \sum_{t=1}^T (Z_{0t} - \alpha \beta'_* Z_{1t} - \Lambda Z_{2t})' \Omega^{-1} (Z_{0t} - \alpha \beta'_* Z_{1t} - \Lambda Z_{2t}). \quad (7.3)$$

Tämän funktion maksimointi vaatii numeerisia menetelmiä (johtuen tulon $\alpha \beta'_*$ epälineaarisuudesta). Periaatteessa voidaan käyttää mitä tahansa optimointialgoritmia, kunhan α ja β vain ensin rajoitetaan yksikäsitteiksi (esimerkiksi asettamalla $\beta'_* = [I_r : A : -\nu]$; ks. jakson 5.3 loppu). Seuraavassa esitetetään vaiheittainen profiiliuskottavuusfunktion perustuva menettely, josta on tullut tässä yhteydessä suosituksi. Tilastollisen päättelyn kurssilla esitetyn profiiliuskottavuusfunktion idea on kerrattu harjoitustehtävissä 6, joiden tuloksia käytetään myös seuraavassa. Ensin esitetään kaksi aputulosta.

Kaksi aputulosta. Log-uskottavuusfunktion maksimoinnissa käytetään profiiliuskottavuusfunktion lisäksi myös seuraavia matriisilaskennan tuloksia.

Apulause 7.1. Ositetun positiivisesti definiitin matriisin $A = [A_{ij}]_{i,j=1,2}$ determinantille pätee (A_{11} ja A_{22} neliömatriiseja)

$$\det(A) = \det(A_{11}) \det(A_{22} - A_{21} A_{11}^{-1} A_{12}) = \det(A_{22}) \det(A_{11} - A_{12} A_{22}^{-1} A_{21})$$

¹⁸Huomaa, että log-uskottavuusfunktioita (4.7) edeltävien laskelmien mukaan $X'_t \pi = \Pi x_t$ ja että stationaarisuusoletusta ei käytetty uskottavuusfunktioita johdettaessa missään kohtaa.

Todistus: Merkitään $A_{i,j} = A_{ii} - A_{ij}A_{jj}^{-1}A_{ji}$ (käänteismatriisit ovat olemassa positiivisdefiniittisyyden nojalla). Tällöin (suora lasku ja $A_{12} = A'_{21}$)

$$A = \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & A_{2.1} \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} := BDB',$$

josta ensimmäinen tulos seuraa, koska $\det(A) = \det(B) \det(D) \det(B') = \det(D)$. Toinen tulos saadaan vastaavasti.

Apulause 7.2. Olkoon M ($n \times n$) positiivisesti semidefiniitti ja N ($n \times n$) positiivisesti definiitti matriisi. Tällöin,

$$\min \frac{\det(X'MX)}{\det(X'NX)} = \prod_{i=1}^r \rho_{n+1-i},$$

jossa minimi on yli kaikkien astetta r olevien $n \times r$ matriisien X ja $\rho_1 \geq \dots \geq \rho_n \geq 0$ saadaan ratkaisemalla determinanttiyhtälö

$$\det(\rho N - M) = 0.$$

Minimi saavutetaan, kun $X = [w_1 \ \dots \ w_r]$, jossa vektori w_i ($n \times 1$) toteuttaa yhtälön

$$\rho_i N w_i = M w_i \quad (i = 1, \dots, n).$$

Tämän tuloksen todistus sivuutetaan. Huomaa, että asettamalla $N = GG'$ nähdään, että ρ_i on positiivisesti semidefiniitin matriisin $G^{-1}M(G^{-1})'$ i . ominaisarvo, mistä seuraa tunnetusti $\rho_i \geq 0$ ja $\rho_i > 0$, jos M on positiivisesti definiitti.

Uskottavuusfunktion maksimointi. Kuten edellä mainittiin, maksimoidaan log-uskottavuusfunktio (7.3) vaiheittain käyttäen profiiliuskottavuusfunktiota.

Vaihe 1. Maksimoidaan (7.3) parametrin Λ suhteen pitäen parametrien α , β_* ja Ω arvot kiinnitettyinä. Tämä maksimointi voidaan suorittaa kuten jaksossa 4.1 käyttäen malliyhtälöä $Z_{0t} - \alpha\beta'_*Z_{1t} = \Lambda Z_{2t} + \varepsilon_t$, jossa vasen puoli tulkitaan selitettävien muuttujien vektoriksi. Suoralla laskulla nähdään, että PNS-menetelmällä saatava ratkaisu (ks. jakso 4.1) ei riipu parametrissa Ω ja voidaan kirjoittaa matriisimuodossa (ks. HT 4.3)

$$\hat{\Lambda}(\alpha, \beta_*) = \sum_{t=1}^T (Z_{0t} - \alpha\beta'_*Z_{1t}) Z'_{2t} \left(\sum_{t=1}^T Z_{2t} Z'_{2t} \right)^{-1} \quad (7.4)$$

Tätä (parametreista α ja β^* riippuvaa) estimaattoria ja log-uskottavuusfunktion (7.3) lauseketta käyttäen määritellään logaritmoitu profiiliuskottavuusfunktio

$$l_1^{(p)}(\alpha, \beta_*, \Omega) = l(\alpha, \beta_*, \hat{\Lambda}(\alpha, \beta_*), \Omega), \quad (7.5)$$

jonka maksimointi tuottaa parametrien α , β_* ja Ω SU-estimaattorit $\hat{\alpha}$, $\hat{\beta}_*$ ja $\hat{\Omega}$. Sijoittamalla estimaattorit $\hat{\alpha}$ ja $\hat{\beta}_*$ yhtälöön (7.4) saadaan edelleen parametrin Λ SU-estimaattori kaavalla $\hat{\Lambda} = \hat{\Lambda}(\hat{\alpha}, \hat{\beta}_*)$.

Vaihe 2. Edellä sanotun perusteella riittää tarkastella logaritmoidun profiiliuskottavuusfunktion $l_1^p(\alpha, \beta_*, \Omega)$ maksimointia. Merkitään

$$R_{it} = Z_{it} - \hat{\Phi}_i Z_{2t}, \quad \hat{\Phi}_i = \sum_{t=1}^T Z_{it} Z'_{2t} \left(\sum_{t=1}^T Z_{2t} Z'_{2t} \right)^{-1}, \quad i = 0, 1.$$

Käyttäen näitä merkintöjä ja matriisin $\hat{\Lambda}(\alpha, \beta_*)$ määritelmää nähdään suoralla laskulla, että $\hat{\Lambda}(\alpha, \beta_*) Z_{2t} = \hat{\Phi}_0 Z_{2t} - \alpha \beta'_* \hat{\Phi}_1 Z_{2t}$ ja edelleen että

$$Z_{0t} - \alpha \beta'_* Z_{1t} - \hat{\Lambda}(\alpha, \beta_*) Z_{2t} = R_{0t} - \alpha \beta'_* R_{1t}.$$

Sijoittamalla tämä funktion $l(\alpha, \beta_*, \hat{\Lambda}(\alpha, \beta_*), \Omega)$ lausekkeeseen (ks. yhtälö (7.3)) saadaan yhtälön (7.5) logaritmoidulle profiiliuskottavuusfunktiolle lauseke

$$l_1^{(p)}(\alpha, \beta_*, \Omega) = -\frac{T}{2} \log \det(\Omega) - \frac{1}{2} \sum_{t=1}^T (R_{0t} - \alpha \beta'_* R_{1t})' \Omega^{-1} (R_{0t} - \alpha \beta'_* R_{1t}). \quad (7.6)$$

Maksimoidaan tämä nyt kiinnittämällä parametrien β_* ja Ω arvot, jolloin $\beta'_* R_{1t}$ tulkitaan selittävien muuttujien vektoriksi regressiomallissa $R_{0t} = \alpha(\beta'_* R_{1t}) + \varepsilon_t$. Kuten vaiheessa 1, nähdään jälleen jakson 4.1 laskelmista, että PNS-menetelmällä saatava α :n ratkaisu ei riipu parametrissa Ω ja voidaan kirjoittaa (ks. HT 4.3)

$$\begin{aligned} \hat{\alpha}(\beta_*) &= \sum_{t=1}^T R_{0t} R'_{1t} \beta_* \left(\sum_{t=1}^T \beta'_* R_{1t} R'_{1t} \beta_* \right)^{-1} \\ &= S_{01} \beta_* (\beta'_* S_{11} \beta_*)^{-1}, \end{aligned} \quad (7.7)$$

jossa

$$S_{ij} = \frac{1}{T} \sum_{t=1}^T R_{it} R'_{jt}, \quad i = 0, 1.$$

Kuten vaiheessa 1 seuraa, että parametrin α SU-estimaattori $\hat{\alpha}$ saadaan sijoittamalla yhtälössä (7.7) β_* :n paikalle sen SU-estimaattori $\hat{\beta}_*$ eli $\hat{\alpha} = \alpha(\hat{\beta}_*)$.

Vaihe 3. Edellä todetun perusteella riittää johtaa parametrien β_* ja Ω SU-estimaattorit. Tämä voidaan suorittaa soveltamalla logaritmoitua profiiliuskottavuusfunktioon (7.6) edelleen profiiliuskottavuusfunktion ideaa. Toisin sanoen, maksimoidaan funktio $l_2^{(p)}(\beta_*, \Omega) = l_1^{(p)}(\hat{\alpha}(\beta_*), \beta_*, \Omega)$. Kiinnitettyä β_* :n arvolla tämä maksimointiongelmaksi voidaan tulkita kovarianssimatriisin Ω SU-estimointiongelmaksi,

kun malliyhtälönä on $R_{0t} - \hat{\alpha}(\beta_*) \beta_*' R_{1t} = \varepsilon_t$. Kuten jaksossa 4.1 saadaan tulokseksi siten

$$\hat{\Omega}(\beta_*) = \frac{1}{T} \sum_{t=1}^T (R_{0t} - \hat{\alpha}(\beta_*) \beta_*' R_{1t}) (R_{0t} - \hat{\alpha}(\beta_*) \beta_*' R_{1t})'. \quad (7.8)$$

Suora lasku ja matriisien $\hat{\alpha}(\beta_*)$ ja S_{ij} ($i, j = 0, 1$) määritelmät johtavat vaihtoehtoisiin esityksiin

$$\begin{aligned} \hat{\Omega}(\beta_*) &= S_{00} - \hat{\alpha}(\beta_*) (\beta_*' S_{11} \beta_*) \hat{\alpha}(\beta_*)' \\ &= S_{00} - S_{01} \beta_* (\beta_*' S_{11} \beta_*)^{-1} \beta_*' S_{10}. \end{aligned} \quad (7.9)$$

Käyttäen mitä tahansa näistä vaihtoehtoisista esitysmuodoista saadaan parametrin Ω SU-estimaattori $\hat{\Omega}$ sijoittamalla β_* :n paikalle sen SU-estimaattori $\hat{\beta}_*$ eli $\hat{\Omega} = \hat{\Omega}(\hat{\beta}_*)$.

Vaihe 4. Edellä esitetystä seuraa, että parametrien Λ , α ja Ω SU-estimaattorit voidaan muodostaa helposti, kunhan parametrin β_* SU-estimaattori $\hat{\beta}_*$ tunnetaan. Parametrin β_* SU-estimaattori maksimoi logaritmoidun profiiliuskottavuusfunktion $l_3^{(p)}(\beta_*) = l_2^{(p)}(\beta_*, \hat{\Omega}(\beta_*)) = l_1^{(p)}(\hat{\alpha}(\beta_*), \beta_*, \hat{\Omega}(\beta_*))$. Yhtälöistä (7.6) ja (7.8) seuraa (vrt. HT 5.3)

$$l_3^{(p)}(\beta_*) = -\frac{T}{2} \log \det(\hat{\Omega}(\beta_*)) - \frac{Tn}{2}.$$

Näin ollen, $\hat{\beta}_*$ löydetään minimoimalla

$$\det(\hat{\Omega}(\beta_*)) = \det \left(S_{00} - S_{01} \beta_* (\beta_*' S_{11} \beta_*)^{-1} \beta_*' S_{10} \right).$$

Apulauseen 7.1 avulla nähdään, että

$$\begin{aligned} \det \left(\begin{bmatrix} S_{00} & S_{01} \beta_* \\ \beta_*' S_{01} & \beta_*' S_{11} \beta_* \end{bmatrix} \right) &= \det(S_{00}) \det(\beta_*' (S_{11} - S_{10} S_{00}^{-1} S_{01}) \beta_*) \\ &= \det(\beta_*' S_{11} \beta_*) \det \left(S_{00} - S_{01} \beta_* (\beta_*' S_{11} \beta_*)^{-1} \beta_*' S_{10} \right). \end{aligned}$$

Näin ollen,

$$\det(\hat{\Omega}(\beta_*)) = \det(S_{00}) \frac{\det(\beta_*' (S_{11} - S_{10} S_{00}^{-1} S_{01}) \beta_*)}{\det(\beta_*' S_{11} \beta_*)},$$

joten ongelma voidaan kutistaa oikean puolen osamäärän minimointiin. Apulauseesta 7.2 nähdään, että ratkaisu johtaa determinanttiyhtälöön

$$\det(\rho S_{11} - (S_{11} - S_{10} S_{00}^{-1} S_{01})) = 0$$

tai yhtäpitävästi determinanttiyhtälöön

$$\det(\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}) = 0, \quad (7.10)$$

jossa $\lambda = 1 - \rho$. Jos $\hat{\rho}_1 \geq \dots \geq \hat{\rho}_n$ ovat edellisen yhtälön ratkaisut, niin $\hat{\lambda}_i = 1 - \hat{\rho}_{n+1-i}$ ($i = 1, \dots, n$) ovat jälkimmäisen yhtälön ratkaisut ja $1 > \hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_n > 0$ ($1 > \hat{\lambda}_1$,

koska $\hat{\rho}_n > 0$ ja $\hat{\lambda}_n > 0$ samoin perustein kuin $\hat{\rho}_n > 0$; ks. Apulauseen 7.2 jälkeinen keskustelu). Ratkaisemalla vektorit \hat{w}_i ($n \times 1$) yhtälöstä

$$\hat{\lambda}_i S_{11} \hat{w}_i = S_{10} S_{00}^{-1} S_{01} \hat{w}_i \quad (i = 1, \dots, n) \quad (7.11)$$

saadaan parametrin β_* SU-estimaattoriksi $\hat{\beta}_* = [\hat{w}_1 \cdots \hat{w}_r]$.

Yhtälön (7.10) yhteydessä puhutaan usein yleistetystä ominaisarvotehtävästä.¹⁹ Ominaisarvojen $\hat{\lambda}_i$ ja ominaisvektoreiden \hat{w}_i laskemiseen on olemassa valmiita algoritmeja. Tavallisesti ominaisvektorit normalisoidaan vaatimalla $\hat{w}' S_{11} \hat{w} = I_n$, jossa $\hat{w} = [\hat{w}_1 \cdots \hat{w}_n]$. Tällöin SU-estimaattori $\hat{\beta}_*$ toteuttaa rajoitteen $\hat{\beta}_*' S_{11} \hat{\beta}_* = I_r$ ja siten $\hat{\alpha} = S_{01} \hat{\beta}_*$ (ks. yhtälö (7.7) ja sitä seuraava keskustelu). Nämä rajoitteet eivät ole tulkinnallisesti mielekkäitä, mutta sikäli käteviä, että ne toimivat ilman mitään etukäteistietoa yhteisintegroituvuuden rakenteesta (vrt. jakson 5.3 viimeinen kohta). Sen jälkeen kun $\hat{\beta}_*$ on käytettävissä voidaan muodostaa vaihtoehtoisia SU-estimaattoreita käyttäen jakson 5.3 viimeisessä kohdassa esitettyä menettelyä eli muodostamalla $\hat{\beta}_{*c} = \hat{\beta}_* (c' \hat{\beta}_*)^{-1}$ halutulla matriisilla c (esimerkiksi $c' = [I_r : 0]$). Tällöin parametrin α SU-estimaattoriksi tulee $\hat{\alpha}_c = \hat{\alpha} (\hat{\beta}_*' c)$.

Yhteenvedo SU-estimoinnista. Kootaan vielä edellä esitetty yhteen.

- (i) Parametrin β_* SU-estimaattori on $\hat{\beta}_* = [\hat{w}_1 \cdots \hat{w}_r]$, jossa vektorit $\hat{w}_1, \dots, \hat{w}_r$ saadaan yhtälöstä (7.11) valitsemalla $\hat{\lambda}_i$:n paikalle $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_r$ eli determinanttiyhtälön (7.10) r suurinta ratkaisua. Estimaattori $\hat{\beta}_*$ normeerataan vaatimalla $\hat{\beta}_*' S_{11} \hat{\beta}_* = I_r$.
- (ii) Kun $\hat{\beta}_*$ on muodostettu, saadaan parametrien α ja Ω SU-estimaattorit $\hat{\alpha} = \alpha(\hat{\beta}_*)$ ja $\hat{\Omega} = \hat{\Omega}(\hat{\beta}_*)$ sijoittamalla yhtälöihin (7.7) ja (7.8) tai (7.9) $\beta_* = \hat{\beta}_*$.
- (iii) Lopuksi, parametrin Λ SU-estimaattori $\hat{\Lambda} = \hat{\Lambda}(\hat{\alpha}, \hat{\beta}_*)$ saadaan sijoittamalla yhtälöön (7.4) $\alpha = \hat{\alpha}$ ja $\beta_* = \hat{\beta}_*$.

Huomataan vielä, että jos $r = 0$, voidaan mallissa (7.2) tulkita $\alpha = 0$ ja $\hat{\Lambda}$ saadaan yhtälöstä (7.4) asettamalla vastaavasti $\alpha = 0$. Tällöin $\hat{\Omega}$ voidaan muodostaa yhtälöstä (7.8) asettamalla $\hat{\alpha}(\beta_*) = 0$ eli $\hat{\Omega} = S_{00}$ (ks. (7.9)). Tapaus $r = n$ voidaan myös käsitellä edellä esitettyssä kehikossa, mutta suora PNS:ään perustuva estimointi on tietysti suoraviivaisempi tapa (ks. jakso 4.1).

¹⁹Kanonista analyysia tuntevat tunnistavat ominaisarvot $\hat{\lambda}_i$ muuttujien R_{0t} ja R_{1t} välisten empiiristen kanonisten korrelaatiokertoimien neliöiksi.

Deterministisen aikatrendin tapaus. Tarkastellaan nyt deterministisen trendin sisältävää virheenkorjausmallia (ks. (5.13))

$$\Delta y_t = \nu_0 + \alpha(\beta' y_{t-1} - \nu_1(t-1)) + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1, \dots, T, \quad (7.12)$$

jossa aikaisempaan tapaan alkuarvot y_{-p}, \dots, y_0 oletetaan havaituiksi ja $\varepsilon_t \sim \text{nid}(0, \Omega)$ (Ω positiivisesti definiitti). Tämä malli voidaan esittää yhtälöä (7.2) käyttäen, kun merkinnät määritellään uudelleen eli nyt muuttujat Z_{1t} ja Z_{2t} ovat

$$Z_{1t} = [y'_{t-1} \ (t-1)]' \quad \text{ja} \quad Z_{2t} = [1 \ \Delta y'_{t-1} \ \cdots \ \Delta y'_{t-p+1}]'$$

ja niitä vastaavat parametrit

$$\beta_* = [\beta' : -\nu_1]' \quad \text{ja} \quad \Lambda = [\nu_0 : \Lambda_1 : \cdots : \Lambda_{p-1}].$$

Edellä esitetty uskottavuusfunktion vaiheittainen maksimointimenettely voidaan nyt toistaa käyttäen näitä määritelmiä. Tällöin parametrin β_* SU-estimaattori muodostetaan ensin yhtälöistä (7.10) ja (7.11), kun vektorit R_{it} ja matriisit S_{ij} ($i, j = 0, 1$) määritellään uudelleen käyttäen tähän tilanteeseen sopivia vektorien Z_{1t} ja Z_{2t} määritelmiä. Parametrien α , Ω ja Λ estimaattorit voidaan tämän jälkeen perustaa yhtälöihin (7.7) ja (7.8) tai (7.9) aivan kuten edellä yksityiskohtaisesti käsitellyssä tapauksessa.

Jos mallissa (7.12) oletetaan $\nu_1 = 0$ ja ettei parametria ν_0 voida esittää muodossa $\nu_0 = -\alpha\nu$, saadaan malliksi

$$\Delta y_t = \nu_0 + \alpha\beta' y_{t-1} + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1, \dots, T. \quad (7.13)$$

Kuten jaksossa 5.3 todettiin, on prosessissa y_t tällöinkin aikatrendi. Tämä malli voidaan esittää yhtälöä (7.2) käyttäen, kun muuttujat Z_{1t} ja Z_{2t} määritellään

$$Z_{1t} = y_{t-1} \quad \text{ja} \quad Z_{2t} = [1 \ \Delta y'_{t-1} \ \cdots \ \Delta y'_{t-p+1}]'$$

ja niitä vastaavat parametrit

$$\beta_* = \beta \quad \text{ja} \quad \Lambda = [\nu_0 : \Lambda_1 : \cdots : \Lambda_{p-1}].$$

Käyttäen näitä määritelmiä ja muuttaen vektorien R_{it} ja matriisien S_{ij} ($i, j = 0, 1$) määritelmät vastaavasti voidaan esitetty uskottavuusfunktion vaiheittainen maksimointimenettely jälleen toistaa ja muodostaa SU-estimaattorit kuten edelläkin.

Edellä sanotun perusteella on ilmeistä, että poistamalla malliyhtälöstä (7.13) vakio-termi ν_0 ja määrittelemällä Z_{2t} ja Λ tätä vastaavalla tavalla nähdään miten uskottavuusfunktion maksimointi hoituu (käytännössä melko harvoin soveltuvassa) mallissa, jossa ei ole deterministisiä termejä lainkaan.

7.2 Yhteisintegroituvuusasteen testaus

Uskottavuusosamäärätesti: Testisuure. Edellisessä jaksossa yi-aste oletettiin tunnetuksi, mikä on käytännössä usein enemmän tai vähemmän epärealistista. Seuraavassa johdettava uskottavuusosamäärätestin testisuure esitetään konkreettisuuden vuoksi vakiolisessa mallissa, mutta kuten edellisen jakson estimointitarkasteluissa voidaan testisuureen johtoa soveltaa merkintöjä muuttaen myös muille tarkastelluille virheenkorjausmalleille.

Tarkastellaan virheenkorjausesitykseen (5.11) perustuvaa rajoittamatonta mallia

$$\Delta y_t = \Theta(y_{t-1} - \mu) + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1, \dots, T.$$

Testattava hypoteesi väittää, että matriisin Θ aste on r , mikä on yhtäpitävää sen kanssa, että Θ voidaan esittää tulona $\Theta = \alpha\beta'$, jossa α ja β ovat kuten aikaisemminkin. Yhtälön (7.2) merkintöjä käyttäen edellä esitetty malli voidaan kirjoittaa

$$Z_{0t} = \Theta_* Z_{1t} + \Lambda Z_{2t} + \varepsilon_t, \quad t = 1, \dots, T,$$

jossa $\Theta_* = [\Theta : -\nu_0]$ ja $\nu_0 = \Theta\mu$. Lineaarialgebraa käyttäen nähdään helposti, että matriisien Θ ja Θ_* asteet ovat samat, joten testattavaksi hypoteesiksi saadaan

$$H(r) : r(\Theta_*) = r.$$

Seuraavassa tälle hypoteesille johdetaan uskottavuusosamäärätesti, kun vaihtoehtona on, että Θ_* on täysin rajoittamaton eli $H(n)$. Ellei toisin mainita, oletetaan nollahypoteesin $H(r)$ olevan voimassa.

Koska uskottavuusfunktion maksimiarvo on sama kuin profiiliuskottavuusfunktion maksimiarvo (ks. HT 6 ja edellisen jakson vaiheen 4 alku), pätee edellisen jakson merkinnöin

$$l(H(r)) := l(\hat{\alpha}, \hat{\beta}_*, \hat{\Lambda}, \hat{\Omega}) = l_1^{(p)}(\hat{\alpha}, \hat{\beta}_*, \hat{\Omega}) = l_3^{(p)}(\hat{\beta}_*)$$

ja käyttäen aiemmin esitettyä $l_3^{(p)}(\hat{\beta}_*)$:n lauseketta (ks. edellisen jakson vaihe 4)

$$l(H(r)) = -\frac{T}{2} \log \det(\hat{\Omega}(\hat{\beta}_*)) - \frac{Tn}{2}.$$

Ottaen huomioon normalisointi $\hat{\beta}'_* S_{11} \hat{\beta}_* = I_r$ saadaan (ks. edellisen jakson vaihe 4)

$$\begin{aligned} \det(\hat{\Omega}(\hat{\beta}_*)) &= \det(S_{00}) \frac{\det(\hat{\beta}'_* (S_{11} - S_{10} S_{00}^{-1} S_{01}) \hat{\beta}_*)}{\det(\hat{\beta}'_* S_{11} \hat{\beta}_*)} \\ &= \det(S_{00}) \det(\hat{\beta}'_* (S_{11} - S_{10} S_{00}^{-1} S_{01}) \hat{\beta}_*). \end{aligned}$$

Kertomalla yhtälö (7.11) vasemmalta vektoreilla $\hat{w}'_1, \dots, \hat{w}'_r$ ja käyttäen määritelmää $\hat{\beta}_* = [\hat{w}_1 \ \dots \ \hat{w}_r]$ sekä normalisointia $\hat{\beta}'_* S_{11} \hat{\beta}_* = I_r$ nähdään, että

$$\hat{\beta}'_* S_{10} S_{00}^{-1} S_{01} \hat{\beta}_* = \text{diag}[\hat{\lambda}_1 \ \dots \ \hat{\lambda}_r] \quad (7.14)$$

ja edelleen, että

$$\hat{\beta}'_* (S_{11} - S_{10} S_{00}^{-1} S_{01}) \hat{\beta}_* = \text{diag}[(1 - \hat{\lambda}_1) \ \dots \ (1 - \hat{\lambda}_r)].$$

Kaiken kaikkiaan voidaan siis todeta, että

$$l(H(r)) = -\frac{T}{2} \log \det(S_{00}) - \frac{T}{2} \sum_{i=1}^r \log(1 - \hat{\lambda}_i) - \frac{Tn}{2},$$

jossa keskimäinen termi oikealla jätetään pois, kun $r = 0$ (ks. edellisen jakson toiseksi viimeisen kohdan loppu)

Koska rajoittamattoman (eli $H(n)$:n toteuttavan) mallin tapauksessa voidaan edellisen jakson tuloksia soveltaa asettamalla $r = n$ (ks. jälleen toiseksi viimeisen kohdan loppu), saadaan

$$l(H(n)) = -\frac{T}{2} \log \det(S_{00}) - \frac{T}{2} \sum_{i=1}^n \log(1 - \hat{\lambda}_i) - \frac{Tn}{2}.$$

Uskottavuusosamäärätestin testisuureeksi saadaan siten²⁰

$$\text{LR}(r) = 2[l(H(n)) - l(H(r))] = -T \sum_{i=r+1}^n \log(1 - \hat{\lambda}_i).$$

Käyttäen normalisointia $\hat{\beta}'_* S_{11} \hat{\beta}_* = I_r$ saadaan $\hat{\alpha} = S_{01} \hat{\beta}_*$ (ks. yhtälö (7.7) ja sen jälkeinen keskustelu), joten yhtälöstä (7.14) seuraa

$$\hat{\alpha}' S_{00}^{-1} \hat{\alpha} = \hat{\beta}'_* S_{10} S_{00}^{-1} S_{01} \hat{\beta}_* = \text{diag}[\hat{\lambda}_1 \ \dots \ \hat{\lambda}_r].$$

Ominaisarvojen $\hat{\lambda}_1, \dots, \hat{\lambda}_r$ voidaan näin ollen tulkita mittaavan estimaattorin $\hat{\alpha}$ kokoa. Toisin sanoen, jos oikea yi-aste on $r_1 > r$ ja oikea α on siten dimensiota $n \times r_1$, on odotettavissa, että ainakin jotkut ominaisarvoista $\hat{\lambda}_{r+1}, \dots, \hat{\lambda}_n$ tulevat selvästi nollaa suuremmiksi ja siten myös testisuureen $\text{LR}(r)$ arvo tulee suureksi.

Kuten edellisessä jaksossakin, päädytään deterministisiä termejä sisältävissä malleissa samanlaiseen testisuureeseen. Ainoa ero edellä esitettyyn on, että matriisit S_{ij} ja ominaisarvot $\hat{\lambda}_1, \dots, \hat{\lambda}_r$ täytyy määrittellä uudelleen ja SU-estimaattorit tulkita vastaavasti uudelleen (ks. s. 65-66).

²⁰ Vaihtoehtoinen esitys on $\text{LR}(r) = T [\log \det(\hat{\Omega}(r)) - \log \det(\hat{\Omega}(n))]$, jossa $\hat{\Omega}(k)$ on kovarianssimatriisin Ω SU-estimaattori, kun yi-asteeksi on oletettu k .

Uskottavuusosamäärätesti: Asymptoottinen jakauma. Jos edellä tarkasteltua virheenkorjausmallia tarkastellaan reaalisessa erikoistapauksessa (eli $n = 1$), on yi-aste r joko nolla tai yksi. Testattava hypoteesi on $H(0)$, jonka voimassa ollessa y_t on astetta $p - 1$ oleva autoregressiivinen yksikköjuuriprosessi. Vaihtoehtona on stationaarinen (tai stabiili) AR(p)-prosessi. Testausongelma kutistuu siten jakson 6 viimeisessä kohdassa tarkastellun autoregressiivisen yksikköjuuren testaukseen. Koska tässä erikoistapauksessa päädyttiin tavanomaisesta poikkeaviin asymptoottisiin jakaumiin, on selvää, ettei nyt tarkasteltavassa yleisemmässä tapauksessakaan voida vedota tavanomaiseen SU-teorian asymtotiikkaan.

Erikoistapauksessa $r = 0$ sujuu testisuureen $LR(r)$ asymptoottisen jakauman johtaminen samaan tapaan kuin aikaisemmin tarkastellussa yksikköjuuritestissä. Kun $r > 0$, tulee tilanne kuitenkin hankalammaksi. Todetaan vain, että hypoteesin $H(r)$ voimassa ollessa ja vakiollisen mallin (7.1) tapauksessa

$$LR(r) \xrightarrow{d} \text{tr} \left\{ \left(\int_0^1 F(u) dW(u)' \right)' \left(\int_0^1 F(u) F(u)' du \right)^{-1} \left(\int_0^1 F(u) dW(u)' \right) \right\}, \quad (7.15)$$

jossa $W(u) \sim \mathbf{BM}(I_{n-r})$ ja $F(u) = [W(u)' \ 1]'$. Rajajakaumassa esiintyvän Brownin liikkeen dimensio on siis $n - r$ eli nollahypoteesin olettamien yksikköjuurien tai yhteisten trendikomponenttien lukumäärä (vrt. yksikköjuuren testaus jakson 6 AR(p)-mallissa).

Edellä esitettyä testiä sanotaan yleensä *jälkitestiksi*. Vaihtoehtoisesti voidaan käyttää ns. *maksimiminaisarvotestiä*, joka perustuu testisuureeseen

$$LR_{\max}(r) = 2[l(H(r+1)) - l(H(r))] = -T \log(1 - \hat{\lambda}_{r+1}).$$

Kuten testisuureen lausekkeesta voidaan päätellä, on vaihtoehtona tässä tapauksessa $H(r+1) : r(\Theta_*) = r + 1$. Nollahypoteesin voimassa ollessa pätee

$$LR_{\max}(r) \xrightarrow{d} \lambda_{\max} \left\{ \left(\int_0^1 F(u) dW(u)' \right)' \left(\int_0^1 F(u) F(u)' du \right)^{-1} \left(\int_0^1 F(u) dW(u)' \right) \right\}, \quad (7.16)$$

jossa $\lambda_{\max}\{\cdot\}$ on aaltosuluissa olevan matriisin suurin ominaisarvo (tässä tapauksessa dimensioltaan $(n - r + 1) \times (n - r + 1)$).

Kuten autoregressiivisissä yksikköjuuritesteissä (ks. jakso 6), riippuvat testien asymptoottiset jakaumat myös yi-astetta testattaessa siitä mitä deterministisiä komponentteja malliin on sisällytetty. Uskottavuusosamäärätestien tapauksessa testisuureiden asymptoottiset jakaumat voidaan esittää muuttamalla funktion $F(u)$ määrittelyä tuloksissa (7.15) ja (7.16). Oheisessa taulukossa esitetään aikaisemmin esillä olleet mallivaihtoehdot. Taulukosta voidaan päätellä, että kaikissa tapauksissa testien asymptoottiset jakaumat ovat riippumattomia tuntemattomista parametreista, joten ne voidaan taulukoita (taulukkoita löytyy useista kirjoista ja tietokoneohjelmistoista).

Malli	$F(u)$
$\Delta y_t = \alpha \beta' y_{t-1} + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t$	$(W_1(u), \dots, W_{n-r}(u))$
$\Delta y_t = \alpha(\beta' y_{t-1} - \nu) + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t$	$(W_1(u), \dots, W_{n-r}(u), 1)$
$\Delta y_t = \nu_0 + \alpha(\beta' y_{t-1} - \nu_1(t-1)) + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t$	$(\bar{W}_1(u), \dots, \bar{W}_{n-r}(u), u - \frac{1}{2})$
$\Delta y_t = \nu_0 + \alpha \beta' y_{t-1} + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t, \quad \nu_0 \neq -\alpha \nu$	$(\bar{W}_1(u), \dots, \bar{W}_{n-r-1}(u), u - \frac{1}{2})$

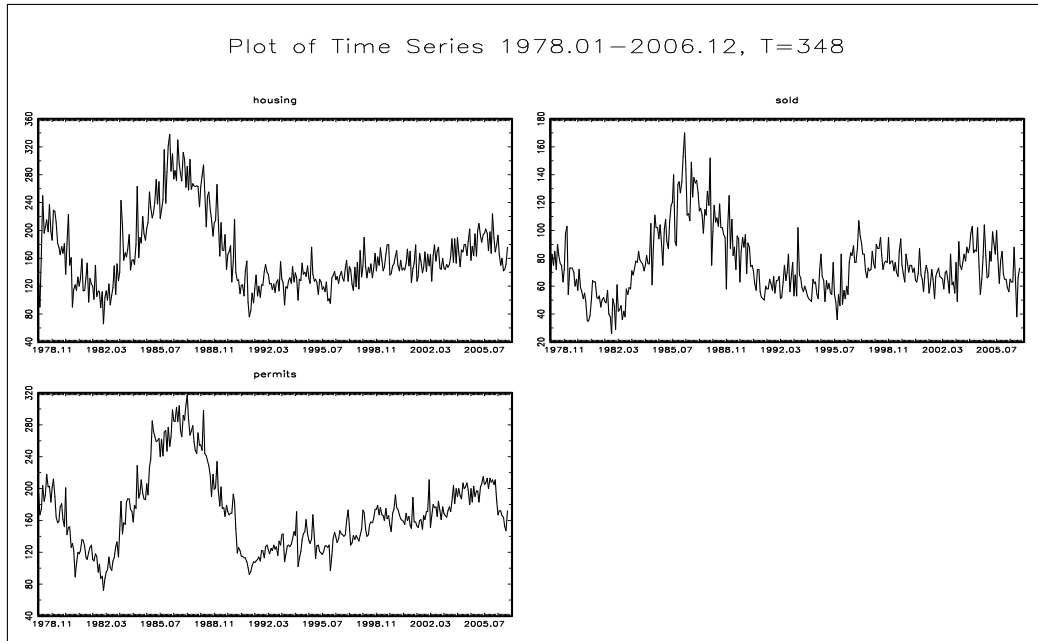
Taulukko 7.1. Testisuureiden $LR(r)$ ja $LR_{\max}(r)$ asymptoottisissa jakaumissa (7.15) ja (7.16) esiintyvän funktion $F(u)$ määrittely eri mallivaihtoehdoissa. $W(u) = (W_1(u), \dots, W_m(u))$ on standardi Brownin liike ja $\bar{W}_i(u)$ keskistetty Brownin liike.

Todetaan myös, että testien johtamisessa käytetty normaalisuusoletus ei ole kriittinen, sillä samat asymptoottiset jakaumat pätevät iid-oletuksella.

Taulukossa 7.1 viimeisenä esitetyn mallin (5.14) yhteydessä on syytä huomata, että se johdettiin yhtälöstä $y_t = \mu_0 + \mu_1 t + x_t$ ($t = 1, 2, \dots$) olettaen, että $\mu_1 = \beta_{\perp} \mu_*$ ($\mu_* \neq 0$). Koska tässä β_{\perp} on yi-matriisin ortogonaalinen komplementti, sisältää malli oletuksen yhteisintegroituvuuden olemassaolosta. Tästä syystä voidaan hypoteesin $H(n)$ testaamista pitää tässä mallissa epäloogisena. Koska vahvaa etukäteistietoa rajoitteen $\mu_1 = \beta_{\perp} \mu_*$ realistisuudesta on melko harvoin käytettävissä, suoritetaan testaus usein ilman tätä rajoitetta käyttäen yleisempää mallia (5.13). Kun yi-aste on saatu täsmennetyksi, voidaan mallin (5.14) sopivuutta testata testaamalla mallissa (5.13) hypoteesia $\nu_1 = 0$ (tähän palataan myöhemmin).

Etenkin useampiulotteisissa tapauksissa on mahdollisesta yi-asteesta usein niukalti etukäteistietoa eikä luontevaa nollahypoteesissa käytettävää astetta ole mahdollista spesifioida. Tällöin on tavallista käyttää edellä esitettyjä testejä peräkkäin. Ensin testataan hypoteesia $H(0)$ ja, jos se jää voimaan, valitaan $r = 0$. Jos $H(0)$ hylätään, testataan hypoteesi $H(1)$ ja jos se jää voimaan, valitaan $r = 1$. Muutoin jatketaan hypoteesiin $H(2)$ jne. Kuten stationaarisen VAR(p)-mallin astetta testattaessa on tässäkin syytä tiedostaa peräkkäisten testien p-arvoihin liittyvät hankaluudet.

Empiirinen esimerkki. Tarkastellaan kolmea kuukausittain havaittua aikasarjaa, jotka kuvaavat USA:n luoteisella alueella aloitettujen uusien asuntorakennushankkeiden määrää (y_{1t}), myytyjen uusien yhden perheen talojen määrää (y_{2t}) ja myönnettyjen uusien yksityisasuntojen rakennuslupien määrää (y_{3t}). Aineisto on ajanjaksolta 1978I - 2006XII ja esitetään graafisesti Kuviossa 7.1. Kaikissa sarjoissa on I(1)-tyyppisiä piirteitä. Aloitettujen asuntorakennushankkeiden määrä ja myönnettyjen rakennuslupien määrä muistuttavat yleisilmeeltään paljon toisiaan ja kummassakin



Kuvio 7.1. USA:n luoteisella alueella aloitettujen uusien asuntorakennushankkeiden määrä (y_{1t} , ylh. vas.), myytyjen uusien yhden perheen talojen määrä (y_{2t} , ylh. oik.) ja myönnettyjen uusien yksityisasuntojen rakennuslupien määrä (y_{3t} , alh.) tuhansina ajanjaksolta 1978I - 2006XII. Aikasarjat ovat kausipuhdistettuja.

voisi tulkita olevan lievän nousevan trendin. Myytyjen yhden perheen talojen määrä on muuten samankaltainen, mutta trendiä ei näyttäisi olevan.

Tutkitaan yhteisintegroituvuutta jälkitestiä käyttäen. Koska lineaarinen (nou-seva) aikatrendi vaikuttaa mahdolliselta, valitaan yleisin tarkasteltu malli (7.12), jossa on aikatrendi pitkän aikavälin relaatiossa. Jos yhteisintegroitusrajoitteita ei oteta huomioon, voidaan malli kirjoittaa (käyttäen merkintöjä ν_0 ja ν_1 aikaisemmasta hie-man poikkeavalla tavalla)

$$\Delta y_t = \nu_0 + \nu_1 t + \Theta y_{t-1} + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1, \dots, T, \quad \varepsilon_t \sim \text{iid}(0, \Omega).$$

Seuraavaksi on valittava mallin aste p . Tässä yhteydessä voidaan käyttää samoja menettelyjä kuin stationaarisessa tapauksessa (ks. jakso 4.4). Tosin sanoen, valitsemalla aste alustavasti ”tarpeeksi suureksi” voidaan tutkia hypoteesia $\Lambda_{p-1} = 0$ ja, jos se jää voimaan, seuraavaksi hypoteesia $\Lambda_{p-2} = 0$ jne., kunnes saadaan ensimmäinen hylkäävä tulos. Voidaan osoittaa, että huolimatta y_t :n mahdollisesta I(1)-luonteesta uskottavuusosamäärätestiä ja Waldin testiä voidaan soveltaa tavanomaista (asymptoottista) χ^2 -jakaumaa käyttäen. Toinen vaihtoehto asteen valinnassa on käyttää mallinvalintakriteerejä. Kun mallin aste on tavalla tai toisella valittu, on

hyvä tutkia myös residuaalien ominaisuuksia ja erityisesti niiden mahdollista korreloituneisuutta.²¹

Mallinvalintakriteereistä AIC ja HQ suosittelevat asteeksi kahta ja BIC yhtä. Astetta kaksi kokeiltaessa residuaaleissa oli muutama korrelaatio, jotka ylittivät niukasti ”kriittisen rajan” $2/\sqrt{T} \approx 0.10$. Asteen nostaminen kolmeen tai neljään pienensi nämä korrelaatiot, mutta ei vaikuttanut residuaalien neliöissä havaittavaan hieman vahvempaan korrelaatioon, mikä viittaa ehdolliseen heteroskedastisuuteen. Jälkimmäinen seikka ei ole kuitenkaan yi–asteen testauksen kannalta kriittinen, sillä testi on osoitettu robustiksi (esimerkiksi GARCH–tyyppiselle) ehdolliselle heteroskedastisuudelle.

Asteen valinta edellä mainittujen vaihtoehtojen välillä ei vaikuttanut testituloksiin. Astetta kaksi käytettäessä saadaan jälkitestillä seuraavat tulokset (esimerkiksi sarakkeessa 95% olevat luvut ovat testisuuren asymptoottisen jakauman 95%:n prosenttipisteitä).

r	LR(r)	90%	95%	99%
0	254.67	39.73	42.77	48.87
1	81.85	23.32	25.73	30.67
2	4.69	10.68	12.45	16.22
Aste $p = 2$				

Tuloksista havaitaan, että hypoteesit $H(0)$ ja $H(1)$ hylätään selvästi, mutta hypoteesi $H(2)$ jää voimaan. Viimeksi mainittu tulos on varsin odotettu Kuvion 7.1 perusteella. Sama tulos saatiin kokeiltaessa myös vakiollisia malleja (7.1) ja (7.13).

7.3 SU–estimaattorien asymptoottinen jakauma

Edellä on todettu, että I(1)–prosesseilla tavanomaiset raja-arvolauseet eivät päde ja että autoregressiivisen yksikköjuuren testauksessa ja yleisemmin yi–asteen testauksessa päädytään tavanomaisista poikkeaviin asymptoottisiin jakaumiin. Jaksossa 7.1 johdettujen yi–vektoreiden SU–estimaattorien asymptoottiset jakaumat ovat samoin normaalista poikkeavia. Tätä taustaa vasten saattaa tuntua yllättävältä, että yi–vektoreita koskevien hypoteesien testaus voidaan kuitenkin perustaa tavanomaiseen tapaan normaali- tai χ^2 –jakaumaan. Selitys piilee siinä, että yi–vektoreiden SU–estimaattorien asymptoottisilla jakaumilla on ”tarpeeksi läheinen” yhteys normaali-jakaumaan. Seuraavassa tätä pyritään havainnollistamaan yksinkertaista esimerkkiä käyttäen.

²¹Residuaaleihin perustuvissa tarkasteluissa, joihin mallinvalintakriteerienkin käyttö kuuluu, saadaan identtiset tulokset käyttäen parametrintia $y_t = \nu_0 + \nu_1 t + A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t$.

Yksinkertainen esimerkkimalli. Tarkastellaan n -ulotteista prosessia $x_t = (z_t, q_t)$, $t = 1, 2, \dots$, jossa z_t on reaalinen. Oletetaan, että x_t :llä on rakenne

$$\begin{aligned} z_t &= \alpha' q_{t-1} + \eta_t \\ q_t &= q_{t-1} + \xi_t, \end{aligned}$$

jossa $(\eta_t, \xi_t) \sim \text{nid}(0, \Omega)$, $\Omega = \text{diag}[\omega^2 \Sigma]$ (positiivisesti definiitti) ja alkuarvo q_0 on kiinteä. Vähentämällä ensimmäisessä yhtälössä puolittain z_{t-1} ja järjestämällä termejä saadaan

$$\Delta x_t = \begin{bmatrix} -1 \\ 0 \end{bmatrix} [1 \quad -\alpha'] x_{t-1} + \varepsilon_t,$$

jossa $\varepsilon_t = (\eta_t, \xi_t)$. Kysymyksessä on siis normaalin VAR(1)-prosessi, jonka virheenkorjausesityksessä $\alpha = (-1, 0)$ on tunnettu. Osoittautuu, että yi-vektorin tai parametrin α estimointi voidaan kuitenkin perustaa regressioyhtälöön $z_t = \alpha' q_{t-1} + \eta_t$.

Oletetaan havainnot (z_t, q_{t-1}) , $t = 1, \dots, T$, ja tarkastellaan parametrin α PNS-estimaattoria

$$\hat{\alpha} = \left(\sum_{t=1}^T q_{t-1} q_{t-1}' \right)^{-1} \sum_{t=1}^T q_{t-1} z_t = (Q'Q)^{-1} Q'z,$$

jossa $z = [z_1 \dots z_T]'$, $Q = [q_0 \dots q_{T-1}]'$ ($T \times (n-1)$ ja $T > n-1$). Tehdyistä oletuksista seuraa, että prosessit $\{\eta_t\}$ ja $\{\xi_t\}$ ovat riippumattomia ja siten matriisi Q on riippumaton virhevektorista $\eta = [\eta_1 \dots \eta_T]'$. Tästä seuraa edelleen (miten?), että $\hat{\alpha}$ on parametrin α SU-estimaattori.

Tarkastellaan nyt estimaattorin $\hat{\alpha}$ jakaumaa. Kirjoitetaan

$$\hat{\alpha} - \alpha = (Q'Q)^{-1} Q'\eta$$

ja tutkitaan oikealla olevan satunnaisvektorin jakaumaa. Koska $\eta \parallel Q$ ja $\eta \sim \text{N}(0, \omega^2 I_T)$, on intuitiivisesti selvää, että ehdolla $Q = Q_*$ tarkasteltava jakauma on sama kuin tapauksessa, jossa Q :n paikalla on kiinteä Q_* .²² Toisin sanoen,

$$(\hat{\alpha} - \alpha) | (Q = Q_*) \sim \text{N}\left(0, \omega^2 (Q_*' Q_*)^{-1}\right).$$

Koska $Q'Q$ on positiivisesti definiitti, sillä on (yksikäsitteinen) positiivisesti definiitti neliöjuuri $(Q'Q)^{1/2}$, joka siis toteuttaa $(Q'Q)^{1/2} (Q'Q)^{1/2} = Q'Q$.²³ Merkitsemällä

²²Matemaattinen perustelu saadaan satunnaisvektorien muunnoksia koskevista tuloksista. Tilanne on kuitenkin tavanomaista mutkikkaampi mm. siksi, että muunnetun sv:n $(Q'Q)^{-1} Q'\eta$ dimensio on pienempi kuin sv:n η .

²³Jos A ($m \times m$) on mielivaltainen positiivisesti semidefiniitti matriisi, sillä on pääakseliha-jotelma $A = RDR'$, jossa R on ortogonaalinen (eli $R'R = I_m$) ja $D = \text{diag}[d_1 \dots d_m]$. Koska D :n diagonaali-alkiot ovat ei-negatiivisia, voidaan kirjoittaa $A = RD^{1/2}R'D^{1/2}R'$, jossa $D^{1/2} = \text{diag}[\sqrt{d_1} \dots \sqrt{d_m}]$. Tämä osoittaa, että $A^{1/2} = RD^{1/2}R'$ (yksikäsitteisyys jää pohdit-tavaksi). Kun A on positiivisesti definiitti, ovat $(A$:n ominaisarvot) d_1, \dots, d_m positiivisia.

$((Q'Q)^{1/2})^{-1} = (Q'Q)^{-1/2}$ voidaan siten todeta kuten edelläkin, että

$$(Q'Q)^{-1/2} Q'\eta \mid (Q = Q_*) \sim \mathbf{N}(0, \omega^2 I_{n-1}).$$

Koska ehdollinen jakauma ei riipu ehtomuuttujan saamasta arvosta, pätee riippumattomuus $\zeta := (Q'Q)^{-1/2} Q'\eta \perp\!\!\!\perp Q$ ja voidaan kirjoittaa

$$\hat{\mathbf{a}} - \mathbf{a} \sim (Q'Q)^{-1/2} \zeta, \text{ jossa } \zeta \sim \mathbf{N}(0, \omega^2 I_{n-1}) \text{ ja } \zeta \perp\!\!\!\perp Q.$$

Estimaattorin $\hat{\mathbf{a}}$ jakauma on siis normaalin, kun ehdollistetaan selittävien muuttujien havaintojen suhteen. On kuitenkin syytä huomata, että ehdollistamaton jakauma ei ole normaalin. Tämän tyyppistä jakaumaa sanotaan usein *sekoitetuksi normaalijakaumaksi* tai täsmällisemmin *(multi)normaalijakaumien kovarianssimatriisisekoitukseksi*. Tilastollisessa päättelyssä tarvitaan kuitenkin vain ehdollista jakaumaa. Tarkastellaan esimerkiksi hypoteesin $\mathbf{a} = c$ (tunnettu) testaamista ja todetaan aluksi, että

$$(\hat{\mathbf{a}} - c)' Q'Q (\hat{\mathbf{a}} - c) / \omega^2 \underset{H_0}{\sim} \zeta'\zeta / \omega^2 \sim \chi_{n-1}^2.$$

Tässä tapauksessa voitaisiin käyttää tavanomaista F-testiä, mutta tulevaa ajatellen tarkastellaan asymptoottista testiä, jossa edellä varianssi ω^2 korvataan tarkentuvalla estimaattorilla $\hat{\omega}^2$, joka voi olla joko tavanomainen lineaarisen mallin harhaton estimaattori tai SU-estimaattori. Tämä korvaaminen ei vaikuta (tarkentuvuuden nojalla) asymptoottisesti, joten saatavalle Waldin testisuurelle pätee

$$W = (\hat{\mathbf{a}} - c)' Q'Q (\hat{\mathbf{a}} - c) / \hat{\omega}^2 \xrightarrow{H_0} \chi_{n-1}^2.$$

Vastaava tulos voidaan johtaa myös yleisemmille lineaarisille hypoteeseille.

Tarkastellaan vielä estimaattorin $\hat{\mathbf{a}}$ asymptoottista jakaumaa käyttäen samanlaista tekniikkaa kuin yksikköjuuren testauksen yhteydessä jaksossa 6. Kirjoitetaan

$$T(\hat{\mathbf{a}} - \mathbf{a}) = \left(\frac{1}{T^2} \sum_{t=1}^T q_{t-1} q'_{t-1} \right)^{-1} \frac{1}{T} \sum_{t=1}^T q_{t-1} \eta_t$$

ja todetaan, että tulosten (6.2) ja (6.3) vastineiksi saadaan nyt (jakaumakonvergenssit yhteisjakaumien tasolla)

$$\frac{1}{T^2} \sum_{t=1}^T q_{t-1} q'_{t-1} \xrightarrow{d} \int_0^1 B(u) B(u)' du, \quad B(u) \sim \mathbf{BM}(\Sigma)$$

ja

$$\frac{1}{T} \sum_{t=1}^T q_{t-1} \eta_t \xrightarrow{d} \int_0^1 B(u) dB_0(u), \quad B_0(u) \sim \mathbf{BM}(\omega^2).$$

Jälkimmäisessä on kaksi eri Brownin liikettä, koska prosessin q_t innovaatio on ξ_t eikä η_t . Huomaa myös, että parametrin \mathbf{a} havaittu informaatiomatriisi on $\omega^{-2} \sum_{t=1}^T q_{t-1} q'_{t-1}$, joten edellinen tulos osoittaa, että tavanomaisista malleista poiketen (sopivasti skaalattu) havaittu informaatio on asympotoottisestikin satunnainen. Edellä esitetyistä tuloksista saadaan (ks. Liite B, Lause B.2)

$$T(\hat{\mathbf{a}} - \mathbf{a}) \xrightarrow{d} \left(\int_0^1 B(u) B(u)' du \right)^{-1} \int_0^1 B(u) dB_0(u).$$

Koska estimaattorin $\hat{\mathbf{a}}$ tarkka jakauma on sekoitettu normaalijakauma, tuntuu selvältä, että saman täytyy päteä myös asympotoottisesti. Tähän liittyen voidaan osoittaa, että prosessien $\{q_t\}$ ja $\{\eta_t\}$ tapaan Brownin liikkeet $\{B(u)\}$ ja $\{B_0(u)\}$ ovat (prosesseina) riippumattomia ja edelleen, että

$$\left(\int_0^1 B(u) B(u)' du \right)^{-1/2} \int_0^1 B(u) dB_0(u) \sim \left(\int_0^1 B(u) B(u)' du \right)^{-1/2} \zeta_0,$$

jossa $\zeta_0 \sim \mathbf{N}(0, \omega^2 I_{n-1})$ ja $\zeta_0 \perp \int_0^1 B(u) B(u)' du$. Toisin sanoen, standardoidun estimaattorin $T(\hat{\mathbf{a}} - \mathbf{a})$ asympotoottinen jakauma on sekoitettu normaalijakauma. Edellä esitetyn Waldin testisuureen asympotoottinen jakauma voidaan perustella myös käyttäen tätä ja edellä mainittuja asympotoottisia jakaumatuloksia. Kaikki tässä esitetyt asympotoottiset tulokset pätevät lisäksi ilman tehtyä normaalisuusoletusta ja voidaan yleistää tapaukseen, jossa yi-relaatiossa on deterministisiä termejä.

Yleinen tapaus. Yleistäen edellä esitettyä huomattavasti voidaan johtaa jaksossa 7.1 johdetun yi-matriisin β normalisoidun version SU-estimaattorin asympotoottinen jakauma. Johdosta tulee algebrallisesti varsin tekninen, mutta taustalla oleva periaate on sama kuin edellä tarkastellussa erikoistapauksessa. Seuraavassa esitetään vain lopputulos käyttäen samaa vakiollista mallia (7.1) kuin jaksossa 7.1.

Kuten edellä mainittiin, täytyy yi-matriisin toteuttaa jotkin normalisointi- tai identifiointirajoitteet, jotta SU-estimaattorille voidaan johtaa asympotoottinen jakauma (vrt. parametrin identifioituvuusvaatimus tilastollisen päättelyn kurssilla). Oletetaan aiemmin tarkasteltu yksinkertainen normalisointi $\beta'_* = [I_r : A : -\nu]$. Merkitään $\mathbf{a} = \text{vec}(A')$ eli parametrivektori \mathbf{a} sisältää matriisin A rivit allekkain (vrt. π jaksossa 4.1). Tällöin \mathbf{a} :n SU-estimaattorille $\hat{\mathbf{a}}$ pätee

$$T(\hat{\mathbf{a}} - \mathbf{a}) \xrightarrow{d} \left(\alpha' \Omega^{-1} \alpha \otimes \int_0^1 \bar{B}(u) \bar{B}(u)' du \right)^{-1/2} \bar{\zeta}, \quad (7.17)$$

jossa $\bar{\zeta} \sim \mathbf{N}(0, I_{r(n-r)})$, $\bar{\zeta} \perp \int_0^1 \bar{B}(u) \bar{B}(u)' du$ ja $\bar{B}(u)$ on $(n-r)$ -ulotteisen (ja $\bar{\zeta}$:sta riippumattoman) Brownin liikkeen $B(u) \sim \mathbf{BM}(\Sigma)$ keskistetty versio. Edelleen, kovarianssimatriisi $\Sigma = c'_\perp C \Omega C' c_\perp$, jossa $C = \beta_\perp \left(\alpha'_\perp \sum_{j=1}^{p-1} \Lambda_j \beta_\perp \right)^{-1} \alpha'_\perp$ on kuten

Grangerin esityslauseessa ja $c'_\perp = [0 : I_{n-r}]$. Keskistetyn Brownin liikkeen $\bar{B}(u)$ ilmaantuminen tulokseen (7.17) johtuu vakiotermistä ν (vrt. testisuureen τ_μ asymptoottinen jakauma jakson 6 lopussa). Vakion ν SU-estimaattorin $\hat{\nu}$ asymptoottista jakaumaa ei tarkastella. Todetaan vain, että se riippuu Brownin liikkeestä $B(u)$ eikä ole normaalin.

Estimaattorin $\hat{\alpha}$ asymptoottiselle jakaumalle saatu esitys on sangen monimutkainen, mikä kuvastuu myös sen (sivuutetussa) johdossa. Olennaisempaa kuin jakauman esityksen yksityiskohdat on kuitenkin huomata, että tuloksena saadaan sekoitettu normaalijakauma kuten edellä tarkastellussa yksinkertaisessa esimerkissäkin (koskee myös $\hat{\nu}$:n asymptoottista jakaumaa). Tämä selittää sen miksi parametrivektoria α (ja ν) koskevissa Waldin ja uskottavuusosamäärätesteissä voidaan käyttää tavanomaiseen tapaan asymptoottista χ^2 -jakaumaa tai normaalijakaumaa.

Parametrien α ja $\Lambda = [\Lambda_1 : \dots : \Lambda_{p-1}]$ eli ns. lyhyen aikavälin parametrien SU-estimaattorit voidaan muodostaa soveltamalla PNS:ää malliyhtälössä

$$\Delta y_t = \alpha \hat{e}_{t-1} + \sum_{j=1}^{p-1} \Lambda_j \Delta y_{t-j} + \varepsilon_t^*, \quad t = 1, \dots, T,$$

jossa $\hat{e}_{t-1} = z_{t-1} + \hat{A}q_{t-1} - \hat{\nu}$ on estimoitu pitkänaikavälin relaatio. Tämä seuraa jakson 7.1 vaiheittaisen SU-estimoinnin vaiheista 2 ja 3 tai johtamalla parametrien α ja Λ profiiliuskottavuusfunktio ja vetoamalla profiiliuskottauusfunktion periaatteeseen. Voidaan osoittaa, että SU-estimaattorin $(\hat{\alpha}, \hat{\Lambda})$ asymptoottinen jakauma ei muutu, vaikka edellä esitetyssä yhtälössä \hat{e}_{t-1} korvataan teoreettisella vastineellaan $e_{t-1} = z_{t-1} + Aq_{t-1} - \nu$ (ks. malli (7.1)). Käyttäen tätä tulosta ja menettelemällä kuten stationaarisen VAR(p)-mallin tapauksessa jaksossa 4.1 voidaan johtaa parametrin (α, Λ) Fisherin informaatiomatriisi ja perustella estimaattorin $(\hat{\alpha}, \hat{\Lambda})$ asymptoottinen jakauma stationaaristen mallien teorian avulla. Jos merkitään $\Pi = [\alpha : \Lambda]$, $\pi = \text{vec}(\Pi')$ ja vastaavasti $\hat{\Pi}$ ja $\hat{\pi}$ saadaan tulokseksi (yksityiskohdat jätetään tehtäväksi)

$$\sqrt{T}(\hat{\pi} - \pi) \xrightarrow{d} \mathbf{N}(0, \Omega \otimes \Gamma_v^{-1}), \quad (7.18)$$

jossa $\Gamma_v = \text{Cov}(v_t)$ ja $v_t = [e'_{t-1} \ \Delta y'_{t-1} \ \dots \ \Delta y'_{t-p+1}]'$.

Edellä sanottu pätee pääpiirteissään myös tapauksessa, jossa havainnoissa on aikatrendi eli malleissa (7.12) ja (7.13). Tuloksen (7.17) vastineeseen tulee keskistetyn Brownin liikkeen $\bar{B}(u)$ paikalle tällöin toisenlainen muunnos Brownin liikkeestä $B(u)$ (mallissa, jossa ei ole lainkaan deterministisiä termejä $\bar{B}(u)$:n paikalle tulee $B(u)$ itse). Olennaista on jälleen, että asymptoottinen jakauma on sekoitettu normaalijakauma, joten uskottavuusteorian testeissä voidaan käyttää tavanomaiseen tapaan asymptoottista χ^2 -jakaumaa tai normaalijakaumaa. Kuten edelläkin, saadaan parametrin π SU-estimaattorille asymptoottinen normaalijakauma, mutta rajajakauman kovarianssimatriisi ei ole yleisesti kuten tuloksessa (7.18).

Edellä esitetty voidaan yleistää tapaukseen, jossa Y -matriisi toteuttaa lineaarisia rajoitteita. Jos rajoitteet koskevat parametrivektoria α on tilanne suoraviivainen. Tällöin voidaan tarkastella rajoitteita $\alpha = H\delta + a$, jossa matriisi H ja vektori a ovat tunnettuja ja sopivaa dimensiota ja H on täyttä sarakeastetta (vrt. jakso 4.2). Saatavan parametrin δ SU-estimaattorin asymptoottinen jakauma on sekoitettu normaali-jakauma, joten parametrivektoria δ koskevissa Waldin ja uskottavuusosamäärätesteissä voidaan käyttää tavanomaiseen tapaan (asymptoottista) χ^2 -jakaumaa. Parametrivektorin π SU-estimaattorille pätee tässäkin tapauksessa tulos (7.18). Myös parametrille π on mahdollista spesifioida lineaarisia rajoitteita. Saatava SU-estimaattori on tällöinkin asymptoottisesti normaalin. Tässä tarkasteltujen yleisten lineaaristen rajoitteiden tapauksessa jaksossa 7.1 esitetty uskottavuusfunktion vaiheittainen maksimointimenetelmä ei kuitenkaan toimi (ainakaan esitettyssä muodossaan). Uskottavuusfunktion maksimointi ”suoraan” numeerisia menetelmiä käyttäen sen sijaan onnistuu (vakiollisen mallin tapauksessa käytetään log-uskottavuusfunktion (7.3) lauseketta).

Y -matriisin rajoittaminen jakson 5.3 lopussa mainittuja lineaarisia rajoitteita käyttäen on myös mahdollista, mutta kuten aiemmin mainittiin vaaditaan tällöin lisäehtoja Y -vektoreiden yksikäsitteisyyden ja toimivan asymptoottisen estimointiteorian aikaan saamiseksi. Näiden lisäehtojen voimassa ollessa pätee edellä sanottu tässäkin tapauksessa.

7.4 Hypoteesien testaus

Kuten edellisessä jaksossa todettiin (konkreettisesti jakson yksinkertaisessa esimerkissä), voidaan sekä Waldin testiä että uskottavuusosamäärätestiä soveltaa tavanomaiseen tapaan testattaessa yhteisintegroituneen VAR-mallin virheenkorjausesityksen parametreja koskevia hypoteeseja. Tämä edellyttää kuitenkin, että Y -aste on oikein spesifioitu. Testattava hypoteesi ei saa myöskään koskea suoraan Y -astetta tai olla sellainen, että Y -aste riippuu siitä onko testattava hypoteesi tosi vai ei. Esimerkki edellisestä on Y -astetta konkreettisesti koskeva hypoteesi (ks. jakso 7.2) ja jälkimmäisestä hypoteesi, jonka voimassa ollessa parametrimatriisin α aste on pienempi kuin spesifioitu Y -aste r .

Waldin testi. Tarkastellaan jälleen vakiollista mallia (7.1) ja kootan sen parametrin vektoriin $\theta = (\alpha, \nu, \pi, \omega)$, jossa ω sisältää (symmetrisen) kovarianssimatriisin Ω erisuuret alkiot (ω otetaan mukaan pelkästään esityksen yksinkertaistamiseksi). Tarkastellaan lineaarista hypoteesia (vrt. jakso 4.3)

$$H_0 : R\theta = b,$$

jossa R ($q \times \dim(\theta)$) ja b ($q \times 1$) ovat tunnettuja ja $r(R) = q$. Olkoon

$$\mathcal{J}_{\theta\theta}(\theta) = -\frac{\partial^2}{\partial\theta\partial\theta'} l(\theta)$$

parametrin θ havaittu informaatiomatriisi ja tarkastellaan tilannetta, jossa hypoteesi rajoittaa vain yi-vektoreita eli parametria α . Tällöin $R = [R_\alpha : 0]$, jossa R_α on $q \times r(n-r)$ ja $r(n-r)$ on α :n dimensio. Hypoteesin H_0 voimassa ollessa pätee Waldin testisuurelle tällöin

$$W = (R\hat{\theta} - b)' \left(R\mathcal{J}_{\theta\theta}^{-1}(\hat{\theta})R' \right)^{-1} (R\hat{\theta} - b) \xrightarrow{d} \chi_q^2. \quad (7.19)$$

Testaus tai P-arvojen laskeminen sujuu siten tavanomaiseen tapaan (ks. jakso 4.3). Koska $R = [R_\alpha : 0]$, on $R\hat{\theta} - b = R_\alpha\hat{\alpha} - b$ ja testisuureessa oleva matriisi $R\mathcal{J}_{\theta\theta}^{-1}(\hat{\theta})R'$ riippuu vain matriisin $\mathcal{J}_{\theta\theta}^{-1}(\hat{\theta})$ vasemman yläkulman ositteesta. Yleensä matriisin $R\mathcal{J}_{\theta\theta}^{-1}(\hat{\theta})R'$ paikalla käytetään matriisia, joka muodostetaan käyttäen vain havaitun informaatiomatriisin parametreihin α , ν ja π liittyvää osaa. Tämä modifikaatio ei vaikuta testisuureen asympotoottiseen jakaumaan, mikä selittyy parametrien ω (tai Ω) ja (α, ν, π) ortogonaalisuudella. Kuten Waldin testissä yleisestikin, voidaan havaittu informaatiomatriisi korvata millä tahansa asympotoottisesti yhtäpitävällä vaihtoehdolla (yleensä tällä tavoitellaan laskennallista helpoutta).

Jos $R\theta = \alpha_i$ eli hypoteesi koskee vain yksittäistä α :n komponenttia, voidaan käyttää myös tulosta

$$(\hat{\alpha}_i - b)/\text{s.e.}(\hat{\alpha}_i) \xrightarrow{d} \text{N}(0, 1) \quad \text{tai} \quad (\hat{\alpha}_i - b)/\text{s.e.}(\hat{\alpha}_i) \underset{as}{\sim} \text{N}(0, 1),$$

jossa $\text{s.e.}(\hat{\alpha}_i)$ on matriisin $R\mathcal{J}_{\theta\theta}^{-1}(\hat{\theta})R'$ i :n diagonaali-alkion neliöjuuri eli estimaattorin $\hat{\alpha}_i$ likimääräinen ”keskivirhe”. Tässä keskivirheen tulkinta on tavanomaisesta sikäli poikkeava, että (sopivasti standardoitu) havaittu informaatiomatriisi on asympotoottisestikin satunnainen (ks. edellisen jakson yksinkertainen esimerkki). Edellä esitetyt ”t-suhteet” ja niiden tulkinta voidaan kuitenkin liittää tavanomaiseen tapaan hypoteesin $R\theta = \alpha_i$ testaamiseen. Parametrivektorin α komponenteille voidaan niin ikään muodostaa luottamusvälejä tavanomaiseen tapaan.

Tulos (7.19) pätee myös, kun hypoteesi rajoittaa lyhyen aikavälin parametrivektoria π tai vakiotermiä ν . Edellisessä tapauksessa rajoitematriisi on muotoa $R = [0 : R_\pi : 0]$ (R_π $q \times \dim(\pi)$) ja jälkimmäisessä $R = [0 : R_\nu : 0]$ (R_ν $q \times \dim(\nu)$). Tuloksen (7.19) jälkeen esitetyt huomautukset pätevät niin ikään (näihin tilanteisiin modifioituina). Koska parametrin π SU-estimaattorin asympotoottinen jakauma on tavanomainen normaalijakauma, on siihen liittyvän tilastollisen päättelyn asympotoottinen jakaumateoria aivan samanlainen kuin aiemmin tarkastellussa stationarisessa VAR-mallissa.

Kuten edellisen jakson estimointituloksetkin, pätee edellä Waldin testistä sanottu pääpiirteissään myös tapauksessa, jossa havainnoissa on aikatrendi eli malleissa (7.12) ja (7.13). Esimerkiksi mallissa (7.12) trendin tarpeellisuutta eli hypoteesia $\nu_1 = 0$ voidaan testata tavanomaista asympotoottista χ^2 -jakaumaa käyttäen.

Uskottavuusosamäärätesti. Uskottavuusosamäärätesti edellä tarkastelluille hypoteeseille voidaan muodostaa tavanomaiseen tapaan kertomalla rajoittamattoman ja rajoitetun mallin log-uskottavuuksien erotus kahdella. Käyttäen aikaisempia merkintöjä saadaan testisuureeksi

$$\text{LR} = 2[l(\hat{\theta}) - l(\tilde{\theta})] \xrightarrow{H_0} \chi_q^2,$$

jossa $\tilde{\theta}$ on parametrin θ rajoitettu SU-estimaattori (eli vakiollisen mallin (7.1) tapauksessa $\theta = (\mathbf{a}, \nu, \pi, \omega)$). Kun jakson 7.1 vaiheittaista uskottavuusfunktion maksimointimenettelyä voidaan käyttää, voidaan log-uskottavuusfunktion maksimiarvot laskea myös profiiliuskottavuusfunktion avulla. Seuraavassa esimerkki tästä.

Edellä tarkastellut estimointi- ja testauskysymykset ovat vaatineet yi-vektoreiden normalisoinnin (tai yksikäsitteisyyden). Tietyn tyyppisissä erikoistapauksissa uskottavuusosamäärätestiä voidaan kuitenkin käyttää ilman, että yi-vektoreita rajoitetaan yksikäsitteisiksi. Vakiollisen mallin (7.1) tapauksessa tämä koskee lineaarisia hypoteeseja, jota ovat tyyppiä

$$H_0 : \beta_* = H\delta,$$

jossa H $((n+1) \times m)$ on tunnettu, δ $(m \times r)$ tuntematon parametrimatriisi ja $\beta_* = [\beta' : -\nu]'$ kuten aikaisemminkin. Koska $r(\beta_*) = r(\beta) = r$, on oltava $m \geq r$ ja $r(\delta) = r$. Tämä hypoteesi on invariantti yi-vektoreiden normalisoinneille eli se rajoittaa vain yi-avaruutta (sillä $\beta_* = H\delta \Leftrightarrow \beta_*\xi = H\delta\xi$ kaikilla epäsingulaarisilla $r \times r$ matriiseilla ξ). Esimerkki tällaisesta hypoteesista saadaan asettamalla yksi matriisin H riveistä nollassi. Jos $m = n$ ja viimeinen H :n rivi on nolla, ei yi-relaatiossa ole vakiota tai yhtäpitävästi $E(\beta' y_t) = 0$. Jos joku muu H :n rivi on nolla, on vastaava yi-matriisin β rivi nolla eikä vastaava y_t :n komponentti esiinny yhdessäkään yi-relaatiossa. Tarkastellaan toisena esimerkkinä tapausta, jossa $n = 4$, $r = 1$ ja

$$\beta_* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \delta, \quad \delta \ (4 \times 1).$$

Jos δ :n (ja siten myös β :n) ensimmäinen komponentti oletetaan nollassa poikkeavaksi, seuraa tästä hypoteesista normalisoitu yi-relaatio $y_{1t} - y_{2t} + \beta_3 y_{3t} + \beta_4 y_{4t} - \nu$. Jos tämä hypoteesi saa tukea ja on tulkinallisia perusteita, voidaan tutkia hypoteesia

$$\beta_* = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \delta, \quad \delta \ (3 \times 1).$$

Jos edellä käytetty normalisointi oletetaan mahdolliseksi, on yi-relaatio tämän hypoteesin voimassa ollessa $y_{1t} - y_{2t} + \beta_3 (y_{3t} - y_{4t}) - \nu$, joten yi-relaatiossa esiintyvät vain erotukset $y_{1t} - y_{2t}$ ja $y_{3t} - y_{4t}$.

Uskottavuusosamäärätestiä varten täytyy johtaa rajoitettu SU-estimaattori. Testattavan hypoteesin voimassa ollessa tarkasteltava malli voidaan esittää sijoittamalla mallissa (7.2) β_* :n paikalle $H\delta$ (ks. jakson 7.1 alku). Tällöin saadaan

$$Z_{0t} = \alpha\delta'H'Z_{1t} + \Lambda Z_{2t} + \varepsilon_t, \quad t = 1, \dots, T, \quad \varepsilon_t \sim \text{nid}(0, \Omega),$$

joka on samaa tyyppiä kuin jaksossa 7.1, jossa tarkasteltavan mallin parametrien rajoittamattomat SU-estimaattorit johdettiin. Korvaamalla siellä käytetyssä vaiheittaisessa menettelyssä vektori Z_{1t} vektorilla $H'Z_{1t}$ saadaan vastaavat rajoitetut SU-estimaattorit. Erityisesti nähdään, että uskottavuusfunktion maksimointi johtaa parametrin δ osalta determinanttiyhtälöön

$$\det(\lambda H'S_{11}H - H'S_{10}S_{00}^{-1}S_{01}H) = 0.$$

Jos $1 > \tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_m > 0$ ovat tämän ratkaisut, niin kuten jaksossa 7.1 saadaan parametrin δ SU-estimaattoriksi $\tilde{\delta} = [\tilde{w}_1 \dots \tilde{w}_r]$ ($m \times r$), jossa \tilde{w}_i toteuttaa

$$\tilde{\lambda}_i H'S_{11}H = H'S_{10}S_{00}^{-1}S_{01}H\tilde{w}_i, \quad (i = 1, \dots, m).$$

Kuten jaksossa 7.1, käytetään tässäkin normalisointia $\tilde{\delta}'H'S_{11}H\tilde{\delta} = I_r$.

Edellä todetun perusteella parametrin β_* rajoitettu SU-estimaattori on $\tilde{\beta}_* = H\tilde{\delta}$. Sijoittamalla tämä rajoittamattoman SU-estimaattorin $\hat{\beta}_*$ paikalle jakson 7.1 vaiheittaisessa uskottavuusfunktion maksimointitarkastelussa nähdään, että

$$\begin{aligned} l(\tilde{\alpha}, \tilde{\beta}_*, \tilde{\Lambda}, \tilde{\Omega}) &= -\frac{T}{2} \log \det(\tilde{\Omega}(\tilde{\beta}_*)) - \frac{Tn}{2} \\ &= -\frac{T}{2} \log \det(S_{00}) - \frac{T}{2} \sum_{i=1}^r \log(1 - \tilde{\lambda}_i) - \frac{Tn}{2} \end{aligned}$$

on rajoitetun mallin log-uskottavuusfunktion maksimiarvo (kuten jaksossa 7.1, pätee $l(\tilde{\alpha}, \tilde{\beta}_*, \tilde{\Lambda}, \tilde{\Omega}) = l_3^{(p)}(\tilde{\beta}_*)$). Testisuureeksi saadaan siten

$$\text{LR} = T \sum_{i=1}^r \log \left(\frac{1 - \tilde{\lambda}_i}{1 - \hat{\lambda}_i} \right) \xrightarrow{H_0} \chi_{r(n+1-m)}^2.$$

Tässä asymptoottinen χ^2 -jakauma perustuu aikaisempaan tapaan siihen, että SU-estimaattorit noudattavat asymptoottisesti sekoitettua normaali-jakaumaa. Vapausasteluku $r(n+1-m)$ nähdään seuraavasti. Oletetaan normalisointi $\beta'_* = [I_r : A : -\nu]$, joka osoittaa, että matriisissa β_* on $r(n+1-r)$ vapaata parametria. Kun $\beta_* = H\delta$, niin $\alpha\beta'_* = \alpha\delta'H'$, joten δ ei ole yksikäsitteinen, mutta se voidaan β_* :n

tapaan normalisoida yksikäsitteiseksi vaatimalla $\delta' = [I_r : B]$, jossa B on $r \times (m - r)$ matriisi (tässä kuten edellisessäkin normalisoinnissa yksikkömatriisin sarakkeiden on oletettu yksinkertaisuuden vuoksi olevan ensimmäiset r saraketta). Nollahypoteesissa on siten rajoitteita $r(n + 1 - r) - r(m - r) = r(n + 1 - m)$.

Edellä johdettu uskottavuusosamäärätesti on sikäli kätevä, että se ei vaadi mitään erityistä yi-vektoreiden normalisointia eikä siten tietoa yi-relaatioiden rakenteesta (vrt. jakson 5.3 loppu). Sama periaate yleistyy ilmeisin muutoksin myös aikatrendin huomioon ottaviin malleihin (7.12) ja (7.13). Tähän liittyvää testiä voidaan soveltaa testattaessa jälkimmäistä mallia edellistä vastaan eli aikatrendin tarpeellisuutta yi-relaatioissa. Samantapainen menettely on mahdollinen myös, kun nollahypoteesi väittää, että jotkut matriisin β_* sarakkeet ovat tunnettuja (perustelu on kuitenkin hankalampi kuin edellä) samoin kuin testattaessa parametria α koskevia lineaarisia hypoteeseja $\alpha = H\varphi$, jossa H ($n \times m$) on tunnettu, φ ($m \times r$) on tuntematon ja lisäksi $r(H) = m$ ja $r(\varphi) = r$ (perustelu sivuutetaan).

7.5 Empiirinen esimerkki

Jatketaan jaksossa 7.2 tarkasteltua esimerkkiä, jossa aineisto koostui USA:n luoteisella alueella aloitettujen uusien asuntorakennushankkeiden määrästä (y_{1t}), myytyjen uusien yhden perheen talojen määrästä (y_{2t}) ja myönnettyjen uusien yksityisasuntojen rakennuslupien määrästä (y_{3t}). Jaksossa 7.2 nähtiin, että yi-aste $r = 2$ sai tukea aineistolta ja lineaarisen aikatrendin sisällyttäminen malliin tuntui varteenotettavalta vaihtoehdolta. Testin tulokset esitettiin VAR(2)-mallia käyttäen, jota käytetään myös seuraavissa estimointitarkasteluissa.

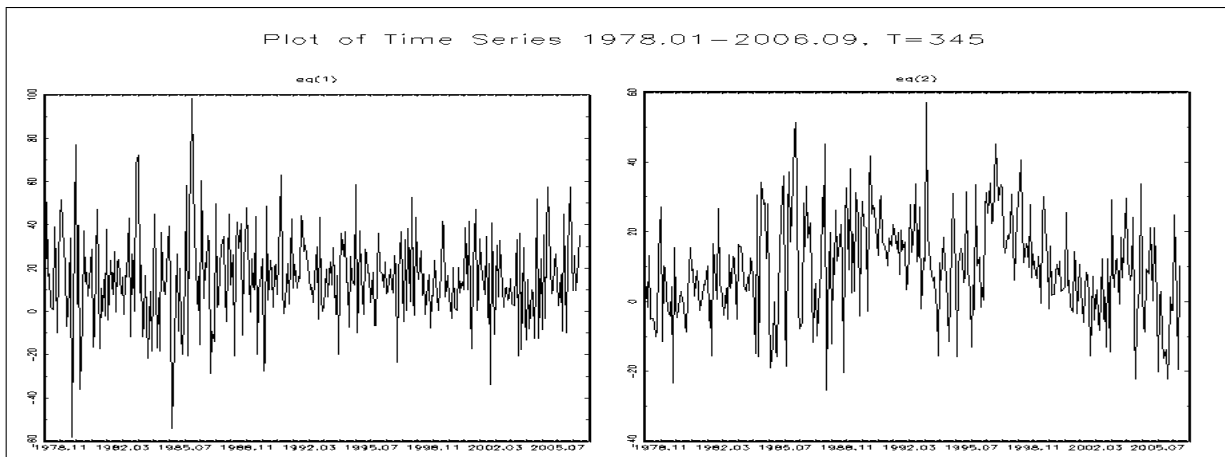
Kun yi-matriisi normalisoidaan kahden ensimmäisen muuttujan suhteen, saadaan jaksossa 7.1 esitettyä SU-menetelmää käyttäen estimoiduksi malliksi

$$\begin{aligned}
\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} &= \begin{bmatrix} -0.98 & 0.24 \\ (0.10) & (0.11) \\ 0.12 & -0.45 \\ (0.06) & (0.07) \\ 0.16 & 0.17 \\ (0.07) & (0.08) \end{bmatrix} \left(\begin{bmatrix} 1.0 & 0 & -0.98 \\ & & (0.02) \\ 0 & 1.0 & -0.40 \\ & & (0.03) \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} 0.08(t-1) \\ (0.01) \\ 0.01(t-1) \\ (0.01) \end{bmatrix} \right) \\
&+ \begin{bmatrix} 12.42 \\ (1.99) \\ 2.58 \\ (1.18) \\ -4.17 \\ (1.42) \end{bmatrix} + \begin{bmatrix} 0.05 & -0.09 & -0.35 \\ (0.08) & (0.11) & (0.10) \\ -0.05 & -0.21 & -0.03 \\ (0.05) & (0.07) & (0.06) \\ -0.04 & -0.07 & -0.27 \\ (0.06) & (0.08) & (0.07) \end{bmatrix} \begin{bmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{bmatrix} \\
&+ \begin{bmatrix} 0.01 & -0.12 & -0.18 \\ (0.06) & (0.09) & (0.08) \\ 0.01 & -0.12 & -0.04 \\ (0.03) & (0.06) & (0.05) \\ -0.04 & -0.08 & -0.12 \\ (0.04) & (0.07) & (0.06) \end{bmatrix} \begin{bmatrix} \Delta y_{1,t-2} \\ \Delta y_{2,t-2} \\ \Delta y_{3,t-2} \end{bmatrix} + \begin{bmatrix} \hat{\varepsilon}_{1t} \\ \hat{\varepsilon}_{2t} \\ \hat{\varepsilon}_{3t} \end{bmatrix}, \tag{7.20}
\end{aligned}$$

jossa estimaattien alla sulussa on edellisessä jaksossa esitetyt keskivirheet. Mallista

lasketut residuaalit ovat lievästi (samanaikaisesti) korreloituneita keskenään (korrelaatiot 0.2-0.4). Kuten jaksossa 7.2 tarkastellussa vastaavassa rajoittamattomassa VAR(2)-mallissakin, residuaalien auto- ja ristikorrelaatiot ovat kohtuullisen pieniä eivätkä viittaa selvästi korkeampiasteisen mallin tarpeellisuuteen. Formaalina testinä voidaan tässä yhteydessä käyttää jaksossa 4.4 esitettyä portmanteau-testiä, kunhan huomataan, että asymptoottisen χ^2 -jakauman vapausasteluku ei ole stationaarisessa tapauksessa soveltuva $n^2(K-p)$, vaan $n^2(K-(p-1)) - nr$ (olettaen mallin (7.20) tapaan, ettei parametreja α ja $\Lambda_1, \dots, \Lambda_{p-1}$ ole rajoitettu). Huolimatta kohtuullisen pienistä auto- ja ristikorrelaatioista testi hylkää nollahypoteesin selvästi (p-arvo 0.003, kun $K = 16$). Asteen kasvattaminen ei kuitenkaan vaikuta olennaisesti testin tulokseen. Tässä tapauksessa portmanteau-testiin on kuitenkin syytä suhtautua varauksin, sillä testi ei ole pätevä, mikäli (korreloimattomissa) virheissä on ehdollista heteroskedastisuutta. Tästä saatiin viitteitä jo jaksossa 7.2 ja sama näkyy myös mallin (7.20) residuaaleissa, joiden neliöissä on melko selvää auto- ja ristikorrelaatiota. Residuaalit ovat myös selvästi ei-normaalisia (jakauma on normaalijakaumaa paksuhäntäisempi). Kuten yi-asteen testauksessakin, virheiden ei-normaalisuus ja ehdollinen heteroskedastisuus ei vaikuta kuitenkaan yi-vektorien estimaattorien asymptoottisiin ominaisuuksiin. Tältä osin mallia (7.20) voidaan pitää tyydyttävänä.

Estimoiduista kahdesta pitkän aikavälin relaatiosta ensimmäinen liittyy toisiinsa aloitettujen uusien asuntorakennushankkeiden määrän ja myönnettyjen uusien yksityisasuntojen rakennuslupien määrän. Kun edellisen kerroin on normalisoitu ykköseksi, ei jälkimmäisen kerroin poikkea miinus ykkösestä ”merkittävästi”. Pitkän aikavälin relaatiossa olevan aikatrendin kerroin on kuitenkin selvästi nolasta poikkeava. Kerroimen etumerkistä voidaan päätellä, että aloitettujen uusien asuntorakennushankkeiden määrä on jäänyt estimointiperiodilla ajan mittaan systemaattisesti yhä enemmän



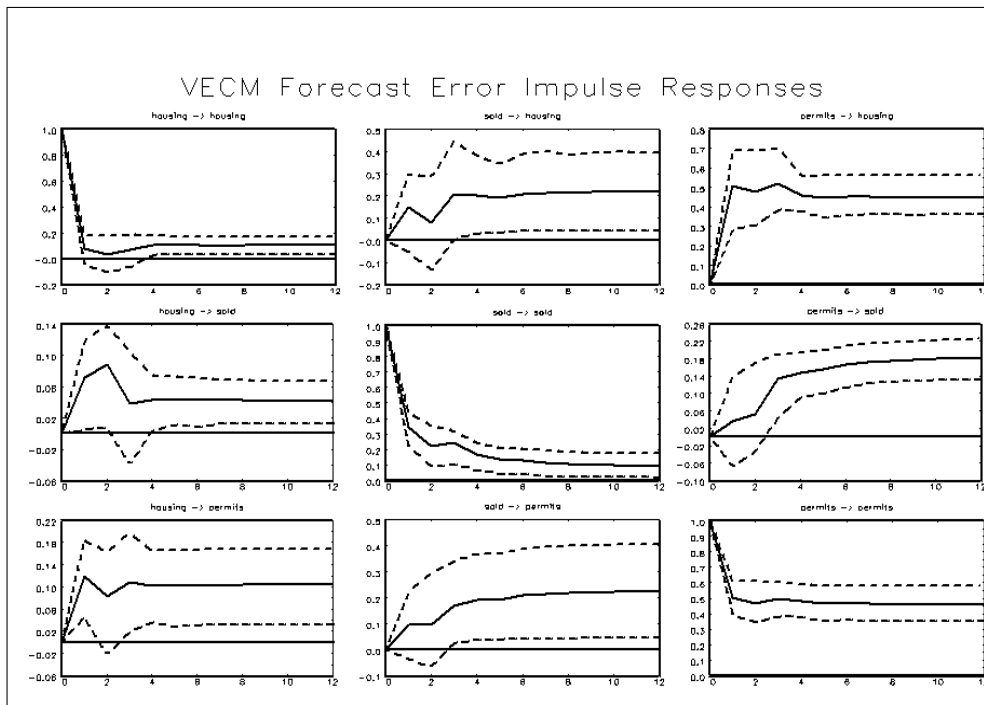
Kuvio 7.2. Estimoidun mallin (7.20) pitkän aikavälin relaatiot $y_{1,t-1} - 0.98y_{3,t-1} + 0.08(t-1)$ (vas.) ja $y_{1,t-1} - 0.40y_{3,t-1} + 0.01(t-1)$ (oik.).

myönnettyjen uusien yksityisasuntojen rakennuslupien määrää pienemmäksi. Trendin ympärillä uusien asuntorakennushankkeiden määrän kehitys noudattaa kuitenkin läheisesti (ja ehkä odotetusti) myönnettyjen uusien yksityisasuntojen rakennuslupien määrää.

Toinen estimoidun pitkän aikavälin relaatio liittyy toisiinsa myytyjen uusien yhden perheen talojen määrän ja myönnettyjen uusien yksityisasuntojen rakennuslupien määrän. Normalisointi on edellisen muuttujan suhteen. Tässä tapauksessa ”merkitsevästi” nollassa poikkeavaa aikatrendiä ei ole. Estimoitu yi-vektori (tai kerroin 0.40) osoittaa (ehkä jälleen odotetusti), että myytyjen uusien yhden perheen talojen määrä ei ole yhtä läheisessä yhteydessä myönnettyjen uusien yksityisasuntojen rakennuslupien määrään kuin aloitettujen uusien asuntorakennushankkeiden määrä. Estimoidut pitkän aikavälin relaatiot esitetään graafisesti Kuviossa 7.2.

Latausmatriisin α estimaatin kaikki alkiot voidaan keskivirheiden perusteella tulkita nollassa poikkeavaksi. Tiukkoja tulkintoja ajatellen on kuitenkin syytä muistaa, että mallin residuaalit viittasivat ehdolliseen heteroskedastisuuteen, minkä voimassa ollessa keskivirheiden taustalla oleva asymptoottinen jakumatulos (7.18) ei päde (tarkentuvuus kuitenkin pätee). Nämä estimaatit kuvaavat pitkän aikavälin relaatioissa ilmenevien poikkeamien vaikutuksia tarkasteltavien muuttujien muutoksiin. Esimerkiksi positiivinen muutos ensimmäisessä pitkän aikavälin relaatioissa (eli aloitettujen uusien asuntorakennushankkeiden määrän ja myönnettyjen uusien yksityisasuntojen rakennuslupien määrän erotuksen kasvaminen) aiheuttaa negatiivisen muutoksen (estimoitu kerroin -0.98) myönnettyjen uusien yksityisasuntojen rakennuslupien määrään ja positiivisen muutoksen (estimoitu kerroin 0.16) myönnettyjen uusien yksityisasuntojen rakennuslupien määrään. Vastaavanlainen tulkinta voidaan tehdä tarkasteltaessa toista pitkän aikavälin relaatiota.

Viivästettyjen differenssien estimoiduissa kerroinmatriiseissa on useita estimaatteja, jotka vaikuttavat pieniltä. Kuten latausmatriisin estimaattia tarkasteltaessa, on tässäkin syytä huomata ehdollisen heteroskedastisuuden mahdollinen vaikutus keskivirheisiin. Näitä estimaatteja on muutenkin hankala tulkita (vrt. vastaava impulssivasteiden yhteydessä esitetty keskustelu jaksossa 2.3). Kuviossa 7.3 esitetään tavanomaiset ennustevirheimpulssivastefunktiot. Kuvioon on piirretty myös eräällä bootstrap-menetelmän versiolla lasketut likimääräiset 95%:n luottamusvälit. Kuvioista nähdään, että toisin kuin stationaarisessa tapauksessa impulssivasteet eivät konvergoi nolnaan, vaan nollassa poikkeavaan vakioon. Impulssivasteet on laskettu vuoden periodilta, minkä kuluessa ne lähes vakioituvat. $I(1)$ -tyyppistä ajallista riippuvuutta lukuun ottamatta mallin riippuvuus on siten varsin lyhytkestoista. Tämä ilmenee myös rajoitetun mallin estimoidun polynomimatriisin $\alpha\beta'z + \left(I_n - \sum_{j=1}^{p-1} \Lambda_j z^j\right) \Delta$ determinantin ykköistä suuremmissa juurissa, joista lähinnä ykköstä oleva on itseisarvoltaan 1.54.



Kuvio 7.3. Estimoidusta mallista (7.20) lasketut impulssivasteet ($y_{1t} = \text{housing}$, $y_{2t} = \text{sold}$, $y_{3t} = \text{permits}$). Katkoviivat osoittavat likimääräiset 95%:n luottamusvälit, jotka on muodostettu eräällä bootstrap-menetelmän versiolla.

Liite

A Matriisilaskentaa

A.1 Matriisinormi

Jos $A = [a_{ij}]$ on dimensiota $n \times m$ oleva reaalinen matriisi, niin sen (Euklidinen) *normi* on määritelmän mukaan

$$\|A\| = \sqrt{\sum_{i=1}^n \sum_{j=1}^m a_{ij}^2}.$$

On selvää, että määritelmä voidaan esittää vaihtoehtoisesti muodostamalla matriisin A alkioista $nm \times 1$ vektori ja käyttämällä tavallista (Euklidisen) vektorinormin määritelmää. Tähän (ja muuhunkin) liittyen määritellään matriisin vektorointiopeeraattori $\text{vec}(\cdot)$ seuraavasti. Ositetaan $A = [a_1 \cdots a_m]$ sarakkeittain, jolloin siis a_i ($n \times 1$) on A :n i . sarake ja määritellään vektori (eli vektoroitu A)

$$\text{vec}(A) = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \quad (nm \times 1).$$

Tällöin on selvää, että $\|A\|$ on vektorin $\text{vec}(A)$ normi eli

$$\|A\| = \sqrt{\text{vec}(A)' \text{vec}(A)}.$$

Jälkimmäisestä määritelmästä ja vektorinormin tunnetuista ominaisuuksista nähdään, että matriisinormilla on vektorinormin tavanomaiset ominaisuudet

- (i) $\|A\| \geq 0$ ja $\|A\| = 0$ jos ja vain jos $A = 0$
- (ii) $\|cA\| = |c| \|A\|$ kaikilla $c \in \mathbb{R}$
- (iii) $\|A + B\| \leq \|A\| + \|B\|$ kaikilla $m \times n$ matriiseilla B (kolmioepäyhtälö).

Näiden lisäksi pätee

- (iv) $\|AB\| \leq \|A\| \|B\|$ kaikilla $m \times l$ matriiseilla B (ns. submultiplikatiivisuus).

Edellä määritellyn matriisinormin lisäksi käytetään useita muitakin matriisinormeja, joilta vaaditaan edellä mainitut neljä ehtoa (joissakin lähteissä neljättä ehtoa ei vaadita). Näitä matriisinormeja ei kuitenkaan tarvita tällä kurssilla.

Joissakin tapauksissa on kätevää esittää (Euklidinen) matriisinormi käyttäen neliömatriisin jälkeä eli diagonaalialkioiden summaa. Jos $B = [b_{ij}]$ on $n \times n$ matriisi, niin sen jälki $\text{tr}(B)$ on määritelmän mukaan

$$\text{tr}(B) = \sum_{i=1}^n b_{ii}.$$

Määritelmästä seuraa helposti, että matriisin jäljellä on seuraavat ominaisuudet

$$\text{tr}(B + C) = \text{tr}(B) + \text{tr}(C) \quad \text{ja} \quad \text{tr}(BD) = \text{tr}(DB).$$

Edellisessä yhtälössä C on $n \times n$ matriisi, mutta jälkimmäisessä B voi olla yleisemmin $n \times m$ matriisi ja D puolestaan $m \times n$ matriisi.

Jäljen määritelmää käyttäen voidaan todeta, että matriisin A ($n \times m$) normi voidaan kirjoittaa

$$\|A\| = \sqrt{\text{tr}(A'A)}.$$

Liitteessä A.3 nähdään miten matriisinormia voidaan käyttää tarkasteltaessa matriisijonojen ja sarjojen konvergenssia.

A.2 Matriisin ominaisarvot ja Jordanin hajotelma

Neliömatriisin A ($n \times n$) ominaisarvot $\lambda_1, \dots, \lambda_n$ toteuttavat determinanttiyhtälön $\det(A - \lambda I_n) = 0$. Vastaavat ominaisvektorit u_1, \dots, u_n ($u_i \neq 0$) saadaan yhtälöistä

$$(A - \lambda_i I_n) u_i = \lambda_i u_i \iff Au_i = \lambda_i u_i, \quad i = 1, \dots, n.$$

Jos A ($n \times n$) ja $\lambda_1, \dots, \lambda_n$ ovat kuten edellä, on olemassa epäsingulaarinen (eli kääntyvä) matriisi P ($n \times n$) siten, että

$$A = P\Lambda P^{-1}.$$

Tässä Λ ($n \times n$) on lohkodeagonaalinen ja täsmällisemmin

$$\Lambda = \begin{bmatrix} \Lambda_{n_1} & 0 & \cdots & 0 \\ 0 & \Lambda_{n_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Lambda_{n_k} \end{bmatrix} := \text{diag}[\Lambda_{n_1} \cdots \Lambda_{n_k}],$$

jossa $n_1 + \cdots + n_k = n$ ja

$$\Lambda_{n_i} = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_i & 1 & & & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & 0 & \lambda_i & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda_i \end{bmatrix}.$$

Matriisin Λ_{n_i} päädiagonaalilla on siis ominaisarvo λ_i , päädiagonaalin yläpuolella olevalla diagonaalilla 1 ja muualla 0 (sama ominaisarvo saattaa esiintyä yhtä useammassa Λ :n lohossa). Matriisin A edellä esitettyä esitystä sanotaan *Jordanin (kanoniseksi) muodoksi* tai *Jordanin hajotelmaksi*. Jos A :n ominaisarvot ovat erisuuria, on Λ diagonaalinen diagonaalialkioina A :n ominaisarvot ja matriisin P (linearisesti riippumattomat) sarakkeet ovat A :n ominaisvektoreita

Laskemalla voidaan todeta seuraava kurssin kannalta hyödyllinen tulos, jossa $\binom{k}{j} = 0$, jos $j > k$:

$$\Lambda_{n_i}^k = \begin{bmatrix} \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \cdots & \binom{k}{j}\lambda_i^{k-j} & \cdots & \binom{k}{n_i-1}\lambda_i^{k-n_i+1} \\ 0 & \ddots & \ddots & & \ddots & \\ \vdots & 0 & \ddots & \ddots & & \ddots \\ \vdots & & 0 & \ddots & \ddots & \\ \vdots & & & 0 & \ddots & \ddots \\ 0 & \cdots & \cdots & \cdots & 0 & \ddots \end{bmatrix},$$

jossa diagonaaleilla on aina sama alkio (ja päädiagonaalin alapuolella nolla).

A.3 Matriisijonot ja matriisisarjat

Matriisijonon $A_N = [a_{ij,N}]$, $N = 1, 2, \dots$, konvergenssi kohti (samaa dimensiota olevaa) matriisiä $A = [a_{ij}]$ voidaan määritellä vaatimalla, että $a_{ij,N} \rightarrow a_{ij}$ kaikilla i, j , kun $N \rightarrow \infty$. Tällöin merkitään $\lim_{N \rightarrow \infty} A_N = A$ tai $A_N \rightarrow A$ (kun $N \rightarrow \infty$).

Matriisiarvoinen sarja $S = \sum_{i=1}^{\infty} A_i$ määritellään osajonon $S_N = \sum_{i=1}^N A_i$ raja-arvona eli vaaditaan sellaisen matriisin S olemassaolo, että $S_N \rightarrow S$, jolloin siis $S = \sum_{i=1}^{\infty} A_i$.

Jos $A = [a_{ij}]$ on $n \times m$ matriisi, niin matriisinnormin määritelmästä seuraa $\max\{|a_{ij}|\} \leq \|A\| \leq \sqrt{nm} \max\{|a_{ij}|\}$, jossa maksimi on yli indeksien $i = 1, \dots, n$ ja $j = 1, \dots, m$. Käyttäen tätä ja reaaliulukujonon konvergenssin määritelmää nähdään, että matriisijonolle A_N pätee $A_N \rightarrow A$ jos ja vain jos $\|A_N - A\| \rightarrow 0$. Vastaava vaihtoehtoinen määritelmä voidaan tietysti esittää myös matriisisarjoille. Edellä todetun ja reaalisten sarjojen suppenemista koskevien tulosten perusteella on myös selvää, että matriisisarjan suppenemiselle on riittävää, että reaalin sarja $\sum_{i=1}^{\infty} \|A_i\|$ suppenee (eli itseinen suppeneminen pätee).

Olkoon nyt A ($n \times n$) neliömatriisi ja $\lambda_1, \dots, \lambda_n$ sen ominaisarvot. Oletetaan, että $\rho := \max_{1 \leq i \leq n} |\lambda_i| < 1$.²⁴ Jaksossa A.2 esitetyn Jordanin hajotelman avulla nähdään, että

$$A^N = (P\Lambda P^{-1})^N = P\Lambda^N P^{-1}.$$

²⁴Jos $\lambda \in \mathbb{C}$ eli $\lambda = x + iy$ ($i^2 = -1$), niin $|\lambda| = \sqrt{x^2 + y^2}$ on λ :n itseisarvo.

Koska tunnetusti $N^k \rho^N \rightarrow 0$, kun $N \rightarrow \infty$ kaikilla $k \geq 0$, saadaan jaksossa A.2 todetusta $\Lambda^N \rightarrow 0$, joten edellä todetusta yhtälöstä seuraa

$$A^N \rightarrow 0, \quad \text{kun } N \rightarrow \infty.$$

Valitaan nyt $0 < \bar{\lambda} < 1$ siten, että $\rho < \bar{\lambda} < 1$ (tällainen $\bar{\lambda}$ voidaan aina löytää). Tällöin $N^k \rho^N = N^k (\rho/\bar{\lambda})^N \bar{\lambda}^N$, jossa $N^k (\rho/\bar{\lambda})^N \rightarrow 0$. Näin ollen, $N^k \rho^N < \bar{\lambda}^N$ kaikilla tarpeeksi suurilla N :n arvoilla. Konvergenssi $\Lambda^N \rightarrow 0$ on siten geometrista eli on olemassa vakiot $C < \infty$ ja $r < 1$ siten, että matriisin Λ^N alkioille $[\Lambda^N]_{ij}$ pätee $|\Lambda^N_{ij}| \leq Cr^N$. Sama pätee siten matriisille $A^N = P\Lambda^N P^{-1}$. Matriisinormin submultiplikatiivisuudesta seuraa lisäksi, että

$$\|A^N\| = \|P\Lambda^N P^{-1}\| \leq \|P\| \|P^{-1}\| \|\Lambda^N\|,$$

joten myös $\|A^N\| \rightarrow 0$ geometrisesti (jos A :n ominaisarvot ovat erisuuria, on Λ diagonaalinen diagonaali-alkioina A :n itseisarvoltaan ykköstä pienemmät ominaisarvot ja tämän tuloksen perustelu yksinkertaistuu). Edellä todetusta seuraa, että sarja $\sum_{i=0}^{\infty} A^i$ suppenee ($A^0 = I_n$). Nähdään helposti, että (perustelu jätetään tehtäväksi)

$$\sum_{i=0}^{\infty} A^i = (I_n - A)^{-1}.$$

A.4 Polynomit, polynomimatriisit ja potenssisarjat

Astetta m oleva *polynomi* on funktio

$$p(z) = \sum_{k=0}^m a_k z^k,$$

jossa a_k ja z ovat reaali- tai kompleksilukuja ja $a_m \neq 0$. Kompleksitapauksessa $z \in \mathbb{C}$ voidaan kirjoittaa *algebran peruslauseen* mukaan

$$p(z) = a_m (z - \zeta_1) \cdots (z - \zeta_m),$$

jossa ζ_1, \dots, ζ_m ovat $p(z)$:n nollakohdat tai juuret eli pätee $p(\zeta_k) = 0$. Tapaus $\zeta_k = \zeta_l$ ($k \neq l$) on mahdollinen ja reaalikertoimisessa tapauksessa $a_k \in \mathbb{R}$ kaikilla k kompleksiset juuret esiintyvät liittolukuina eli jos $\zeta_k = x_k + iy_k$ ($i^2 = -1$), niin jollain $l \neq k$ pätee $\zeta_l = x_l - iy_l =: \bar{\zeta}_l$.

Polynomimatriisi on matriisi, jonka alkiot ovat polynomeja:

$$P(z) = [p_{ij}(z)], \quad p_{ij}(z) = \sum_{k=0}^{m_{ij}} a_{ij,k} z^k.$$

Asettamalla $m = \max \{m_{ij}\}$ voidaan vaihtoehtoisesti kirjoittaa

$$P(z) = \sum_{k=0}^m A_k z^k,$$

jossa $A_k = [a_{ij,k}]$ ja $a_{ij,k} = 0$, kun $k > m_{ij}$. Tämän esitystavan perusteella voidaan puhua myös *matriisipolynomeista* eli polynomeista, joiden kertoimet ovat matriiseja.

Laskutoimitukset $A(z)+B(z)$ ja $A(z)B(z)$ sekä determinantti $\det(A(z))$ määritellään kuten tavanomaisilla reaalilukumatriiseilla. Samoin käänteismatriisi eli polynomimatriisin $A(z)$ ($m \times m$) käänteismatriisi $A(z)^{-1}$ toteuttaa $A(z)^{-1}A(z) = A(z)A(z)^{-1} = I_m$. Kuten reaalilukumatriiseilla voidaan käänteismatriisi laskea myös kaavalla

$$A(z)^{-1} = \frac{1}{\det(A(z))} A(z)^* \quad (\det(A(z)) \neq 0),$$

jossa $A(z)^* = [a_{ij}^*(z)]$ on $A(z)$:n ns. adjungaattimatriisi eli $a_{ij}^*(z) = (-1)^{i+j}$ kertaa determinantti siitä $A(z)$:n alimatriisista, joka saadaan poistamalla $A(z)$:sta i . rivi ja j . sarake. Koska determinantti voidaan muodostaa summaamalla matriisin alkiosta laskettuja tuloja, ovat sekä $\det(A(z))$ että $a_{ij}^*(z)$ polynomeja. Käänteismatriisin $A(z)^{-1}$ alkiot ovat siten rationaalifunktioita eli tyyppiä $\sum_{j=0}^q a_j z^j / \sum_{j=0}^r b_j z^j$ ($q, r < \infty$).

Palataan nyt tavalliseen polynomiin $p(z) = \sum_{k=0}^m a_k z^k$. On helppo todeta, että $p(z)$ voidaan kirjoittaa

$$p(z) = p(1) + (1-z)q(z), \tag{A.1}$$

jossa

$$q(z) = \sum_{k=0}^{m-1} b_k z^k, \quad b_k = - \sum_{j=k+1}^m a_j \quad k = 1, \dots, m-1.$$

Vastaavanlainen vaihtoehtoinen esitys on

$$p(z) = p(1)z + (1-z)r(z), \tag{A.2}$$

jossa

$$r(z) = \sum_{k=0}^{m-1} c_k z^k, \quad c_0 = a_0 \quad \text{jä} \quad c_k = - \sum_{j=k+1}^m a_j, \quad k = 1, \dots, m-1.$$

Nämä esitykset voidaan yleistää edelleen potenssisarjoille eli tapaukseen

$$p(z) = \sum_{k=0}^{\infty} a_k z^k \quad (|z| \leq 1).$$

Jos oletetaan, että

$$\sum_{k=1}^{\infty} k |a_k| < \infty,$$

voidaan helposti todeta, että yhtälöt (A.1) ja (A.2) pätevät, kun $|z| \leq 1$ ja m korvataan symbolilla ∞ . Lisäksi pätee $\sum_{k=1}^{\infty} |b_k| < \infty$ ja $\sum_{k=1}^{\infty} |c_k| < \infty$.

Edellä esitetty yleistyy matriisitapaukseen. Olkoon

$$P(z) = [p_{ij}(z)] = \sum_{k=0}^m A_k z^k \quad (|z| \leq 1),$$

jossa $A_k = [a_{ij,k}]$ toteuttaa ehdon

$$\sum_{k=1}^{\infty} k \|A_k\| < \infty.$$

Soveltamalla aikaisempaa polynomeihin $p_{ij}(z)$ saadaan tuloksen (A.1) yleistys

$$P(z) = P(1) + (1-z)Q(z), \quad (\text{A.3})$$

jossa

$$Q(z) = \sum_{k=0}^{\infty} B_k z^k, \quad B_k = - \sum_{j=k+1}^{\infty} A_j \quad \text{ja} \quad \sum_{k=1}^{\infty} \|B_k\| < \infty.$$

Vastaavasti saadaan tuloksen (A.2) yleistys

$$P(z) = P(1)z + (1-z)R(z), \quad (\text{A.4})$$

jossa

$$R(z) = \sum_{k=0}^{\infty} C_k z^k, \quad C_0 = A_0, \quad C_k = - \sum_{j=k+1}^{\infty} A_j, \quad k \geq 1 \quad \text{ja} \quad \sum_{k=1}^{\infty} \|C_k\| < \infty.$$

Jos $P(z)$ on astetta m oleva polynomimatriisi eli $A_k = 0$ kaikilla $k > m$, niin $Q(z)$ ja $R(z)$ ovat astetta $m-1$ olevia polynomimatriiseja.

A.5 Ositetut matriisit ja Kroneckerin tulo

Ositettu matriisi on matriisi, jonka alkiot ovat matriiseja eli

$$A = [A_{ij}] = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nm} \end{bmatrix}.$$

Jos $A = [A_{ij}]$ ja $B = [B_{ij}]$, niin (olettaen, että dimensiot täsmäävät laskutoimitusten määritelmien kanssa)

$$A + B = [A_{ij} + B_{ij}] \quad \text{ja} \quad AB = \left[\sum_k A_{ik} B_{kj} \right].$$

Nämä tulokset seuraavat tavanomaisista matriisien yhteen- ja kertolaskukaavoista. Lisäksi, kun $\det(A) \neq 0$ ja $\det(A_{ii}) \neq 0$ ($i = 1, 2$),

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{bmatrix},$$

jossa $A_{i,j} = A_{ii} - A_{ij}A_{jj}^{-1}A_{ji}$.

Matriisien A ($n \times m$) ja B ($p \times q$) *Kroneckerin tulo* on $np \times mq$ matriisi

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{bmatrix} = [a_{ij}B].$$

Kroneckerin tulolla on seuraavat ominaisuudet, jotka voidaan todeta jokseenkin suoraan määritelmästä (olettaen, että laskutoimitukset ovat määriteltyjä).

- $(A \otimes B)(C \otimes D) = (AC \otimes BD)$
- $(A \otimes B)' = A' \otimes B'$
- $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$
- $A \otimes (B + C) = A \otimes B + A \otimes C$
- Jos matriisien A ($n \times n$) ja B ($m \times m$) ominaisarvot ovat $\lambda_1, \dots, \lambda_n$ ja μ_1, \dots, μ_m , niin matriisin $(A \otimes B)$ ominaisarvot ovat $\lambda_i \mu_j$ ($i = 1, \dots, n, j = 1, \dots, m$)
- $\text{vec}(ABC) = (C' \otimes A) \text{vec}(B)$

A.6 Ehdollisen odotusarvon ominaisuuksia

Tarkastellaan lyhyesti yleistä ennustustilannetta, jossa (reaalisen) satunnaismuuttujan Y arvoa halutaan ennustaa satunnaisvektorin (sv) X saaman arvon $X = x$ avulla. Todennäköisyyslaskennasta tiedetään, että keskineliövirheen mielessä optimaalinen ennuste on Y :n ehdollinen odotusarvo ehdolla $X = x$ eli pätee

$$\mathbb{E}[(Y - \mathbb{E}(Y|X = x))^2] \leq \mathbb{E}[(Y - g(x))^2]$$

olipa $g(x)$ mikä tahansa x :n funktio (tässä odotusarvojen äärellisyys oletetaan). Jatkuvien jakaumien tapauksessa määritellään ehdollinen odotusarvo kaavalla

$$\mathbb{E}(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y; x) dy = \int_{-\infty}^{\infty} y \frac{f_{Y,X}(y, x)}{f_X(x)} dy,$$

jossa $f_{Y,X}(y, x)$ on sv:n (Y, X) yhteistiheysfunktio, $f_X(x) = \int_{-\infty}^{\infty} f_{Y,X}(y, x) dy$ on X :n reunajakauman tiheysfunktio ja $f_{Y|X}(y; x) = f_{Y,X}(y, x) / f_X(x)$ on Y :n ehdollinen tiheysfunktio ehdolla $X = x$. Näissä määritelmissä, samoin kuin seuraavissa tarkasteluissa, Y voi olla vektoriarvoinen.

Kun x vaihtelee yli satunnaisvektorin X mahdollisten arvojen määrittelee x :n funktio $E(Y | X = x)$ satunnaismuuttujan, josta on luonteva käyttää merkintää $E(Y | X)$. Kehittyneessä todennäköisyyslaskennassa ehdollinen odotusarvo määritelläänkin suoraan satunnaismuuttujana ja siellä se yleistetään myös tapaukseen, jossa satunnaisvektori X voi olla ääretönulotteinen. Tämä tilanne tulee vastaan, kun tarkastellaan stokastisen prosessin ennustamista olettaen, että prosessin $\{y_t, t = 0, \pm 1, \dots\}$ kaikki ennusteaikakohtaa edeltävät arvot tunnetaan. Ehdollisesta odotusarvosta riittää kuitenkin tietää vain sen tavanomaiset ominaisuudet, jotka pätevät tällöinkin. Eri-tyisesti, että

EO1 $E(aY_1 + bY_2 | X) = aE(Y_1 | X) + bE(Y_2 | X)$, kun a ja b ovat vakioita (lineaarisuus)

EO2 $E(Y | X) = E(Y)$, kun $Y \perp\!\!\!\perp X$.

EO3 $E(Y) = E[E(Y | X)]$ (ns. iteroidun odotusarvon laki)

EO4 $E[g(X)Y | X] = g(X)E(Y | X)$ mille tahansa funktiolle g (jolla tulo $g(X)Y$ on määritelty ja sen odotusarvo on äärellisenä olemassa).

A.7 Derivointikaavoja

Kun $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on derivoituva, merkitään

$$\frac{\partial}{\partial x} f(x) = \left[\frac{\partial}{\partial x_1} f(x) \quad \cdots \quad \frac{\partial}{\partial x_n} f(x) \right]' \quad (n \times 1)$$

ja

$$\frac{\partial}{\partial x'} f(x) = \left[\frac{\partial}{\partial x_1} f(x) \quad \cdots \quad \frac{\partial}{\partial x_n} f(x) \right] \quad (1 \times n).$$

Kun $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ on derivoituva, merkitään

$$\frac{\partial}{\partial x} f(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f_1(x) & \cdots & \frac{\partial}{\partial x_n} f_m(x) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} f_1(x) & \cdots & \frac{\partial}{\partial x_n} f_m(x) \end{bmatrix}$$

On helppo todeta, että $\partial a'x / \partial x = a$ (a $n \times 1$) ja $\partial Ax / \partial x = A'$, kun A on $m \times n$. Edelleen, kun A on $n \times n$,

$$\begin{aligned} \frac{\partial}{\partial x} x'Ax &= (A + A')x \\ &= 2Ax, \quad \text{kun } A = A'. \end{aligned}$$

Kun $B = B'$, saadaan näiden avulla edelleen

$$\frac{\partial}{\partial x} (c - Ax)' B (c - Ax) = -2A'Bc + 2A'BAx.$$

Mainitaan lisäksi tulokset

$$\frac{\partial}{\partial A} x'Ax := \left[\frac{\partial}{\partial a_{ij}} x'Ax \right] = [x_i x_j] = xx' \quad (x \ n \times 1 \text{ ja } A \ n \times n)$$

ja

$$\frac{\partial}{\partial A} \log \det(A) = (A^{-1})' = (A')^{-1}.$$

Viimeksi mainittu perustuu determinantin tunnettuun kehityskaavan $\det(A) = \sum_k a_{ik} a_{ik}^*$, jossa A :n adjungaattimatriisin A^* (ks. Liite A.4) alkio a_{ik}^* ei riipu A :n alkioista a_{ij} .

Ketjusääntöä soveltaen saadaan näin ollen,

$$\frac{\partial}{\partial a_{ij}} \log \det(A) = \frac{1}{\det(A)} \frac{\partial \det(A)}{\partial a_{ij}} = \frac{1}{\det(A)} a_{ij}^* = [A^{-1}]_{ji} = [(A^{-1})']_{ij},$$

jossa viimeistä edellinen yhtälö perustuu käänteismatriisin laskukaavaan $A^{-1} = (\det(A))^{-1} A^*$.

B Asymptoottisia tuloksia

Seuraavassa esitetään lyhyesti stokastiseen konvergenssiin ja jakaumakonvergenssiin liittyviä tuloksia (mainitut konvergenssikäsitteet oletetaan tunnetuiksi). Ensimmäinen tulos kertoo miten stokastinen konvergenssi käyttäytyy jatkuvissa muunnoksissa. Tässä kuten myöhemminkin tarkoitetaan vakiolla ei-satunnaista (äärellistä) suuretta.

Lause B.1. Jos jono sv:ta X_1, X_2, \dots ($k \times 1$) konvergoi stokastisesti kohti vakiovektoria c eli $X_n \xrightarrow{p} c$ ja $g: \mathbb{R}^k \rightarrow \mathbb{R}^l$ jatkuva pisteessä c , niin $g(X_n) \xrightarrow{p} g(c)$.

Tulos pätee myös matriisitapauksessa, joka voidaan palauttaa lauseen tilanteeseen suorittamalla vektorointi (ks. Liite A.1). Sama pätee seuraaviin jatkuvan kuvauksen lauseisiin, joista ensimmäinen koskee jakaumakonvergenssia.

Lause B.2. Jos jono sv:ta X_1, X_2, \dots ($k \times 1$) konvergoi jakaumaltaan kohti sv:ia Z eli $X_n \xrightarrow{d} Z$ ja $g: \mathbb{R}^k \rightarrow \mathbb{R}^l$ on jatkuva, niin $g(X_n) \xrightarrow{d} g(Z)$.

Mainittakoon, että tämän lauseen jatkuvuusoletusta voidaan lieventää. Funktion g jatkuvuus riittää vaatia vain joukossa C_g , jolle pätee $\mathbf{P}\{Z \in C_g\} = 1$. Koska jakaumakonvergenssista kohti vakiota seuraa stokastinen konvergenssi, seuraa Lauseesta B.2 Lauseen B.1 tulos, kun $Z = c$ on vakio.²⁵ Seuraava lause sisältää edellä mainitut erikoistapauksenaan.

²⁵Stokastisessa konvergenssista (mahdollisesti kohti satunnaista suuretta) seuraa aina jakaumakonvergenssi, mutta käänteinen tulos pätee vain, kun jakaumakonvergenssissa raja-arvo on vakio.

Lause B.3. Oletetaan, että sv-jonoille X_1, X_2, \dots ($k \times 1$) ja Y_1, Y_2, \dots ($l \times 1$) pätee $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$, jossa c on vakiovektori. Tällöin $(X_n, Y_n) \xrightarrow{d} (Z, c)$.

Lauseiden 1.3 ja 1.4 suorina seurauksia saadaan mm. seuraavat tulokset.

Seuraus B.1. Oletetaan Lauseen B.3 tilanne eli $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$ (vakio) ja lisäksi että $k = l = 1$. Tällöin,

(i) $X_n + Y_n \xrightarrow{d} Z + c$

(ii) $X_n Y_n \xrightarrow{d} Zc$

(iii) $X_n/Y_n \xrightarrow{d} Z/c$, kun $c \neq 0$.

Tässä ensimmäinen tulos pätee tietysti myös vektoritapauksessa, jota tarkastellaan seuraavassa, jossa Y_n on $k \times l$ matriisi.

Seuraus B.2. Olkoon X_1, X_2, \dots ja Z ($k \times 1$) kuten Lauseessa 1.4 eli $X_n \xrightarrow{d} Z$ ja Y_1, Y_2, \dots jono $l \times k$ matriiseja, jolle pätee $Y_n \xrightarrow{p} A$, jossa A ($l \times k$) on vakiomatriisi. Tällöin,

(i) $Y_n X_n \xrightarrow{d} AZ$

Jos $k = l$, pätee myös

(ii) $X_n' Y_n X_n \xrightarrow{d} Z' A Z$

(iii) $X_n' Y_n^{-1} X_n \xrightarrow{d} Z' A^{-1} Z$, kun A on epäsingulaarinen (eli kääntyvä).

Nämä tulokset seuraavat funktioiden $(A, x) \mapsto Ax$, $(A, x) \mapsto x' Ax$ ja $A \mapsto A^{-1}$ (A epäsingulaarinen) jatkuvuudesta ja Lauseista B.2 ja B.3. Kun Z on vakio, voidaan jakaumakonvergenssi korvata (vahvemmalla) stokastisella konvergenssilla. Käänteismatriisin jatkuvuus voidaan perustella seuraavasti. Todetaan ensin, että determinanttifunktio $A \mapsto \det(A)$ on jatkuva, koska determinantti voidaan muodostaa summamalla matriisin alkioista laskettuja tuloja (katso 2×2 tapausta). Koska käänteismatriisin A^{-1} alkiot voidaan lausua A :n ja sen alimatriisien determinanttien osamäärien avulla (ks. Liite A.4 ja A.7), on funktio $A \mapsto A^{-1}$ jatkuvien funktioiden yhdisteenä jatkuva.

Seurauksessa B.1(iii) on implisiittisesti oletettu, että $Y_n \neq 0$ kaikilla n (vastaavasti Seurauksessa B.2(iii) on oletettu Y_n :n epäsingulaarisuus kaikilla n). Vaikka oletuksista $Y_n \xrightarrow{p} c$ ja $c \neq 0$ seuraa, että tapahtuman $\{Y_n \neq 0\}$ todennäköisyys lähestyy

ykköstä, ei ehdon $Y_n \neq 0$ toteutuminen äärellisillä n :n arvoilla ei ole kuitenkaan taattu ilman lisäoletuksia (kuten Y_n :n jakauman jatkuvuutta). Tämä teoreettinen hankaluus voitaisiin ottaa huomioon määrittelemällä $1/Y_n$ jollain (mielivaltaisella) tavalla ilman, että esitetty jakaumakonvergenssi muuttuisi mitenkään. Kuten kirjallisuudessa yleensäkin, ei tätä seikkaa oteta yksinkertaisuuden vuoksi eksplisiittisesti huomioon, vaan Seurauksen B.1(iii) kaltaisissa tilanteissa $Y_n \neq 0$ oletetaan, jos Y_n :n stokastinen raja-arvo on nolasta poikkeava. Vastaava tulkinta tehdään ilman eri mainintaa Seurauksen 1.2(iii) kaltaisissa tilanteissa.

Edellä esitettyjen tulosten ja väliarvolauseeseen avulla voidaan perustella ns. *deltamenetelmä*, joka reaaliarvoisessa tapauksessa on seuraava. Olkoon $\theta \in \Theta \subset \mathbb{R}$ estimoitava parametri ja θ_0 sen ”todellinen” arvo. Oletetaan (yksinkertaisuuden vuoksi), että parametrivaruus Θ on avoin väli ja tarkastellaan θ_0 :n asymptoottisesti normaalia estimaattoria $\hat{\theta}_n$ eli

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathbf{N}(0, \sigma^2(\theta_0)) \quad (\sigma^2(\theta_0) > 0).$$

Jos $h : \Theta \rightarrow \mathbb{R}$ on jatkuvasti derivoituva funktio ja $h'(\theta_0) \neq 0$, niin muunnoksen $h(\theta_0)$ estimaattorille $h(\hat{\theta}_n)$ pätee

$$\sqrt{n}(h(\hat{\theta}_n) - h(\theta_0)) \xrightarrow{d} \mathbf{N}(0, [h'(\theta_0)]^2 \sigma^2(\theta_0)).$$

Yleistys vektoriparametrin tapaukseen on seuraava. Olkoon $\Theta \subseteq \mathbb{R}^d$ avoin ja konvekksi ja $h : \Theta \rightarrow \mathbb{R}$ jatkuvasti derivoituva funktio. Oletetaan, että gradienttivektori $\partial h(\theta)/\partial \theta' = [\partial h(\theta)/\partial \theta_1 \ \dots \ \partial h(\theta)/\partial \theta_d]$ toteuttaa $\partial h(\theta_0)/\partial \theta \neq 0$ ja että estimaattoreille $\hat{\theta}_n$ ($d \times 1$) pätee

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathbf{N}_d(0, \Sigma(\theta_0)), \quad \Sigma(\theta_0) \text{ positiivisesti definiitti.}$$

Tällöin,

$$\sqrt{n}(h(\hat{\theta}_n) - h(\theta_0)) \xrightarrow{d} \mathbf{N}(0, V(\theta_0)), \quad V(\theta_0) = [\partial h(\theta)/\partial \theta'] \Sigma(\theta_0) [\partial h(\theta)/\partial \theta].$$

Deltamenetelmä voidaan yleistää edelleen tapaukseen, jossa h on vektoriarvoinen. Väliarvolauseetta täytyy tällöin soveltaa erikseen funktion h kuhunkin komponenttiin.