REGRESSIOANALYYSIN JATKOKURSSI, 5–10 OP (aine- ja syventävät opinnot). 15.9.–16.12.2011. Kirjallisuus: Russell Davidson ja James MacKinnon: Econometric Theory and Methods. Luennoi: yliopistonlehtori Pekka Pere.

Harjoitukset 12 (pe 9.12.)

1. Tehtävä liittyy lineaarisen mallin estimointiin yleistetyllä PNS:llä, kun jäännös noudattaa AR(1)-prosessia

$$u_t = \rho u_{t-1} + \varepsilon_t. \tag{1}$$

Yllä $|\rho| < 1$ ja $\varepsilon_t \sim \mathsf{IID}(0, \sigma^2)$. Vektorin $\mathbf{u} = [u_1 \dots u_n]'$ kovarianssimatriisi on (kirjan kaava (7.32))

$$\mathbf{\Omega}(\rho) = \frac{\sigma_\varepsilon^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & & \rho^{n-2} \\ \rho^2 & \rho & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \rho^{n-2} & \rho^{n-3} & & \ddots & 1 & \rho \\ \rho^{n-1} & \rho^{n-2} & & \cdots & \rho & 1 \end{bmatrix}$$

a) Osoita, että

$$\mathbf{\Omega}^{-1}(\rho) = \sigma_\varepsilon^{-2} \begin{bmatrix} 1 & -\rho & 0 & \cdots & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1+\rho^2 & -\rho & 0 \\ \vdots & & \ddots & -\rho & 1+\rho^2 & -\rho \\ 0 & \cdots & \cdots & 0 & -\rho & 1 \end{bmatrix}$$

b) Kaavan (7.60) mukaan

$$\mathbf{\Psi}(\rho) = \begin{bmatrix} \sqrt{1-\rho^2} & -\rho & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 & -\rho \\ 0 & \cdots & \cdots & 0 & 0 & 1 \end{bmatrix}.$$

Osoita, että $\sigma_\varepsilon^{-1}\mathbf{\Psi}(\rho)\mathbf{\Psi}(\rho)'\sigma_\varepsilon^{-1} = \mathbf{\Omega}^{-1}(\rho)$.

2. Olkoon regressiomalli lineaarinen

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \tag{7.42}$$

havainnot aikajärjestyksessä ja jäännös prosessin (1) mukainen. Hypoteesia $\rho = 1$ voidaan (asymptotiikan perusteella) testata kertoimen b_ρ PNS-estimaatin t-arvolla regressiossa

$$\widetilde{\mathbf{u}} = \mathbf{X}\mathbf{b} + b_\rho \widetilde{\mathbf{u}}_1 + \text{residuaalivektori}. \tag{7.43}$$

Yllä $\widetilde{\mathbf{u}}$ on yhtälön (7.42) PNS–residuaalivektori ja $\widetilde{\mathbf{u}}_1$ on $\widetilde{\mathbf{u}}$ yhdellä viivästettynä ja 0 lisättynä ensimmäisen havainnon paikalle ($\widetilde{u}_1$ yhdellä viivästettynä on tuntematon $\widetilde{u}_0$, joka korvataan 0:lla). Perustele huolellisesti, miksi täsmälleen sama testisuure voidaan laskea regressiosta

$$\mathbf{y} = \mathbf{X}\mathbf{b} + b_\rho \widetilde{\mathbf{u}}_1 + \text{residuaalivektori}. \tag{7.44}$$

3. Tutkitaan mallin (7.57)

$$y_t = \mathbf{X}_t\boldsymbol{\beta} + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathsf{IID}(0, \sigma_\varepsilon^2)$$

estimointia ($t = 1, \ldots, n$). Oletetaan ensin, että regressorit ovat eksogeenisia. Normeerattujen epälineaaristen PNS-estimaattoreiden $n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ ja $n^{1/2}(\hat{\rho} - \rho)$ kovarianssimatriisi on suurilla havaintomäärillä

$$\begin{bmatrix} \sigma_\varepsilon^2[n^{-1}(\mathbf{X} - \rho_0\mathbf{X}_1)'(\mathbf{X} - \rho_0\mathbf{X}_1)]^{-1} & \mathbf{0} \\ \mathbf{0} & 1 - \rho_0^2 \end{bmatrix}$$

(kirjan kaava (7.66)). Yleistetyn normeeratun PNS-estimaattorin $n^{1/2}(\hat{\boldsymbol{\beta}}_Y - \boldsymbol{\beta})$ kovarianssimatriisi on suurilla havaintomäärillä sama kuin yllä

$$\sigma_\varepsilon^2[n^{-1}(\mathbf{X} - \rho_0\mathbf{X}_1)'(\mathbf{X} - \rho_0\mathbf{X}_1)]^{-1}$$

(kirjan kaava (7.67)).

Oletetaan seuraavaksi, että regressorit eivät ole eksogeenisia. Tällöin yleistetty PNS ei ole sopiva estimointimenetelmä mutta epälineaarinen PNS on. Estimaattoreiden $n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ ja $n^{1/2}(\hat{\rho} - \rho)$ kovarianssimatriisi ei ole kuitenkaan tällöin lohkodiagonaalinen. Osoita, että tällöin estimaattorin $n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ kovarianssimatriisi on (kirjan s:lla 105 esitetyssä mielessä) vähintään yhtä suuri kuin $\sigma_\varepsilon^2[n^{-1}(\mathbf{X} - \rho_0\mathbf{X}_1)'(\mathbf{X} - \rho_0\mathbf{X}_1)]^{-1}$. (Vihjeitä: HT:t 1.5, 5.1 ja 5.4.)

4.

a) Olkoon paneeliaineiston selitettävät havainnot $\mathbf{y}$ $(n \times 1)$ ryhmitelty siten, että 1. ryhmän havainnot ovat ylimpänä, 2. ryhmän havainnot seuraavana jne. Ryhmiä on m kappaletta, ja kussakin on T havaintoa $(n = mT)$. Selitä yksityiskohtaisesti dummy-matriisin rakenne kiinteiden vaikutusten (fixed effects) mallissa (7.85)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\eta} + \boldsymbol{\varepsilon},$$

kun aineisto on ryhmitelty edellä kuvatulla tavalla.

b) Merkitään $\mathbf{M_D} = \mathbf{I}_n - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$ $(n \times n)$. Osoita, että $\{\mathbf{M_D x}\}_{it} = x_{it} - \bar{x}_i$, jossa $\{\}_{it}$ on aaltosulkujen sisällä olevan matriisin it. alkio, $\mathbf{x}$ on n-vektori, jonka havainnot on ryhmitelty edellä kuvatulla tavalla ja $\bar{x}_i = T^{-1}\sum_{t=1}^{T} x_{it}$ on vektorin i. ryhmän keskiarvo.

5. Osoita, että virhetermin varianssi on $\sigma_v^2 + \sigma_\varepsilon^2/T$ regressiossa (7.89)

$$\mathbf{P_D y} = \mathbf{P_D X}\boldsymbol{\beta} + \text{residuaalit}.$$

Yllä $\mathbf{P_D} = \mathbf{I} - \mathbf{M_D}$.

6. Yhtälön (7.88) mukaan

$$\boldsymbol{\Sigma} = \sigma_\varepsilon^2 \mathbf{I}_T + \sigma_v^2 \boldsymbol{\iota}\boldsymbol{\iota}'.$$

a) Osoita, että

$$\boldsymbol{\Sigma}^{-1/2} = \frac{1}{\sigma_\varepsilon}(\mathbf{I}_T - \lambda \mathbf{P}_{\boldsymbol{\iota}}),$$

jossa $\mathbf{P}_{\boldsymbol{\iota}} = \boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{\iota})^{-1}\boldsymbol{\iota}' = T^{-1}\boldsymbol{\iota}\boldsymbol{\iota}'$ ja $\lambda = 1 - (T\sigma_v^2/\sigma_\varepsilon^2 + 1)^{-1/2}$. (Vihje: Kokeile, voisiko $\mathbf{V}^{-1/2}$ olla vakiolla kertomista vaille muotoa $\mathbf{I}_T - \lambda \mathbf{P}_{\boldsymbol{\iota}}$ ja ratkaise syntyvä toisen asteen yhtälö.)

b) Osoita, että satunnaisten vaikutusten (random effects) mallissa $\boldsymbol{\beta}$:n yleistetty PNS-estimaattori saadaan laskemalla regressio (7.92)

$$(\mathbf{I} - \lambda \mathbf{P_D})\mathbf{y} = (\mathbf{I} - \lambda \mathbf{P_D})\mathbf{X}\boldsymbol{\beta} + \text{residuaalit}.$$

Mikä on yleistetyn PNS-estimaattorin kovarianssimatriisi?