

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.

## 2. välikoe (pe 16.12.2011)

*Vastaa kolmeen kysymykseen. Jos vastaat useampaan, merkitse selkeästi, mitkä vastauksistasi haluat arvosteltavan. Käytä vastauksissasi aina kaavoja, selitä huolellisesti niiden merkinnät, merkitys ja perustelut. Ylipäänsäkin perustele kaikki vastauksesi huolellisesti.*

*Toivotan Oikein Hyvää ja Rauhaista Joulua kaikille kurssilaisilleni!*

1. Tutkitaan lineaarista mallia

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad (1)$$

jossa  $\mathbf{y} = [y_1 \dots y_n]'$  on selitettävistä havainnoista koostuva  $n$ -vektori,  $\mathbf{X}$  on eksogeenisista selittäjistä koostuva täysiasteinen  $n \times k$  -matriisi ( $n > k$ ),  $\boldsymbol{\beta}$  on  $k$ -parametrivektori,  $\mathbf{u} = [u_1 \dots u_n]'$  on  $n$ -virhevektori,  $E(\mathbf{u}) = \mathbf{0}$ ,  $E(\mathbf{u}\mathbf{u}') = \boldsymbol{\Omega}$ , ja  $\boldsymbol{\Omega}$  on positiivisesti definiitti  $n \times n$  -kovarianssimatriisi.

a) Mikä on parametrivektorin  $\boldsymbol{\beta}$  pienimmän neliösumman (PNS) -estimaattorin ( $\hat{\boldsymbol{\beta}}$ ) kovarianssimatriisi?

b) Liittyykö kovarianssimatriisin  $\boldsymbol{\Omega}$  estimointiin ongelmia? Selitä ja perustele.

c) Miten PNS-estimaattorin  $\hat{\boldsymbol{\beta}}$  kovarianssimatriisi voidaan estimoida, kun  $\boldsymbol{\Omega}$  on diagonaalimatriisi?

2. Epälineaarinen regressiomalli on

$$y_t = x_t(\boldsymbol{\beta}) + u_t, \quad (2)$$

jossa  $y_t$  on selitettävä muuttuja,  $x_t(\boldsymbol{\beta})$  on regressiofunktio,  $\boldsymbol{\beta}$  on  $k$ -parametrivektori, jäännös  $u_t \sim \text{IID}(0, \sigma^2)$  ja  $t = 1, \dots, n$ . Vektorimerkinnöillä malli on

$$\mathbf{y} = \mathbf{x}(\boldsymbol{\beta}) + \mathbf{u},$$

jossa  $\mathbf{y} = [y_1 \dots y_n]'$ ,  $\mathbf{x}(\boldsymbol{\beta}) = [x_1(\boldsymbol{\beta}) \dots x_n(\boldsymbol{\beta})]'$  ja  $\mathbf{u} = [u_1 \dots u_n]' \sim \text{IID}(0, \sigma^2 \mathbf{I})$ .

a) Määrittele  $\boldsymbol{\beta}$ :n epälineaarinen PNS (EPNS) -estimaattori ( $\hat{\boldsymbol{\beta}}$ ).

b) Mikä on normeeratun  $\hat{\boldsymbol{\beta}}$ :n asymptoottinen jakauma? Mikä on sopiva normeeraus?

c) Ositetaan parametrivektori:  $\boldsymbol{\beta} = [\boldsymbol{\beta}'_1 \quad \boldsymbol{\beta}'_2]'$ . Selitä huolellisesti, miten nol-lahypoteesia  $\boldsymbol{\beta}_2 = \mathbf{0}$  voidaan testata. (Kaksi vaihtoehtoa; selitä toinen.)

d) Määrittele malliin liittyvä Gauss–Newton (GN) -regressio.

e) Oletetaan, että olet johtanut edellä mainitun GN-regression. Kehitä se pisteessä  $\hat{\beta}$  ja estimoi regression parametrit PNS:llä (olettaen, että regressio-funktio ja tilasto-ohjelmisto ovat sellaiset, että jälkimmäinen löytää  $\hat{\beta}$ :n). Mitkä ovat saamiesi estimaattien numeeriset arvot? Todista edellinen vastauksesi.

f) Tilasto-ohjelmisto laskee edellä mainittujen PNS-estimaattien lisäksi niihin liittyvän kovarianssimatriisin. Esitä kaava tälle kovarianssimatriisille. Perustele sen rakenne huolellisesti. Mikä on tämän kovarianssimatriisin EPNS-estimoinnin kannalta relevantti tulkinta?

3. Olkoon nollahypoteesi, että malli on lineaarinen

$$y_t = \mathbf{X}_t \boldsymbol{\beta} + u_t.$$

Yllä  $\mathbf{X}_t'$  ja  $\boldsymbol{\beta}$  ovat  $k$ -vektoreita, selittäjät ( $x_{ti}$ ) ovat ennaltamääräytyneitä jäännösten  $u_t \sim \text{IID}(0, \sigma^2)$  suhteen ja  $t = 1, \dots, n$ . Vastahypoteesin pätiessä malli on

$$y_t = \mathbf{X}_t \boldsymbol{\beta} (1 + \theta \mathbf{X}_t \boldsymbol{\beta}) + u_t, \quad (3)$$

jossa parametri  $\theta \neq 0$ .

a) Mitkä ovat mallin (3) selittävät muuttujat ja parametrit vastahypoteesin pätiessä? Selitä huolellisesti, miksi malli (3) on epälineaarinen vastahypoteesin pätiessä. (1,5 p)

b) Johda Gauss–Newton-regression avulla testisuure epälineaarisuudelle. (Kysessä on ns. RESET-testi; Regression Specification Error Test.) (4,5 p)

4. Tutkitaan mallia

$$y_t = \mathbf{X}_t \boldsymbol{\beta} + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t, \quad (4)$$

jossa  $\mathbf{X}_t'$  ja  $\boldsymbol{\beta}$  ovat  $k$ -vektoreita,  $|\rho| < 1$ ,  $\rho \neq 0$ ,  $\varepsilon_t \sim \text{IID}(0, \sigma_\varepsilon^2)$  ja  $t = 1, \dots, n$ . Estimoidaan  $\boldsymbol{\beta}$  (ei epälineaarilla vaan tavallisella) PNS-menetelmällä (huomioimatta jäännöksen  $u_t$  mahdollista autokorrelaatiota).

Vastaa seuraaviin kysymyksiin kohtien a) ja b) tilanteissa: Onko PNS-estimaattori  $\hat{\boldsymbol{\beta}}$  tarkentuva? Onko se harhaton? Onko se tehokas? Jos ei ole, niin onko parempaa estimaattoria olemassa ja jos on, niin mikä se on? Onko PNS-estimaattorin  $\hat{\boldsymbol{\beta}}$  todellinen kovarianssimatriisi analyttisesti laskettavissa? Jos on, laske se. Onko se sama kuin  $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$  (ilmeisin merkinnöin ja oletuksin)? Perustele kaikki vastauksesi huolellisesti.

a) Selittäjät ( $\mathbf{X}_t$ ) ovat eksogeenisiä.

b) Selittäjät ( $\mathbf{X}_t$ ) ovat ennaltamääräytyneitä, ja yksi niistä on selitettävän muuttujan viive  $y_{t-1}$ .