Stationaariset aikasarjat sl 2014, HT 6, viikko 45

1. Tarkastellaan regressiomallia $y_t = \boldsymbol{z}_t'\boldsymbol{\varphi} + u_t$, $t = 1, ..., T$, jossa $\boldsymbol{z}_t$ $(m \times 1)$ on ei-satunnaisten selittävien muuttujien vektori, $\boldsymbol{\varphi}$ $(m \times 1)$ on vastaava tuntematon kerroinvektori ja normaalinen ARMA(p,q)-prosessi $u_t = \phi_1 u_{t-1} + \cdots + \phi_p u_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$, $\varepsilon_t \sim \mathsf{nid}\left(0, \sigma^2\right)$, toteuttaa stationaarisuus-, käännettävyys- ja identifioituvuusehdon. Malli voidaan esittää matriisimerkinnöin

$$\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\varphi} + \boldsymbol{u},$$

jossa $\boldsymbol{y} = [y_1 \ \cdots \ y_T]'$, $\boldsymbol{Z} = [\boldsymbol{z}_1 : \cdots : \boldsymbol{z}_T]'$ $(T \times m)$ ja $\boldsymbol{u} = [u_1 \ \cdots \ u_T]'$. Matriisin $\boldsymbol{Z}$ asteen oletetaan olevan m.

Yleistä monisteen s. 46-47 esitetty tarkastelu edellä kuvattuun tilanteeseen ja esitä (a) satunnaisvektorin $\mathbf{y}$ yhteistiheysfunktio, (b) parametrin $\left(\boldsymbol{\varphi}, \boldsymbol{\beta}, \sigma^2\right)$ log-uskottavuusfunktio ja (c) profiiliuskottavuusfunktio. Konkreettisina esimerkkeinä vektorista $\boldsymbol{z}_t$ voidaan mainita $\boldsymbol{z}_t = 1$ ja $\boldsymbol{z}_t = [1 \ \ t]'$.

Vihje: Satunnaisvektorin $\boldsymbol{u}$ jakauma on sama kuin monisteen s. 46 esitetty satunnaisvektorin $\mathbf{y}$ jakauma, joten koska $\boldsymbol{Z}\boldsymbol{\varphi}$ on ei-satunnainen, niin ...

2. Tarkastellaan stationaarista ja käännettävää ARMA(1,1)-mallia $y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$, $\varepsilon_t \sim \mathsf{nid}\left(0, \sigma^2\right)$ ($|\phi| < 1$ ja $|\theta| < 1$), ja siihen liittyen monisteen tuloksessa (4.5) olevaa parametrin $\boldsymbol{\beta} = (\phi, \theta)$ Fisherin informaatiomatriisia $\mathbf{V}(\boldsymbol{\beta})$ (ks. monisteen s. 48 ja myös alla oleva huomautus). Kuten monisteen s. 48 todetaan, pätee $\mathbf{V}(\boldsymbol{\beta}) = T\sigma^{-2}\mathsf{Cov}(\mathbf{x})$, jossa nyt tarkasteltavan ARMA(1,1)-mallin tapauksessa $\mathbf{x} = \left(y_1^+, y_1^*\right) \in \mathbb{R}^2$ ja y_1^+ ja y_1^* saadaan AR-prosesseista $y_t^+ = \phi y_{t-1}^+ + \varepsilon_t$ ja $y_t^* = -\theta y_{t-1}^* + \varepsilon_t$, $\varepsilon_t \sim \mathsf{nid}\left(0, \sigma^2\right)$. Oletetaan, että $\phi = -\theta$ eli ettei tarkasteltava ARMA(1,1)-malli toteuta tuloksen (4.5) olettamaa identifioituvuusehtoa (ks monisteen s. 32).

Osoita, että edellä kuvatussa tilanteessa kovarianssimatriisi $\mathsf{Cov}(\mathbf{x})$ on singulaarinen, joten tuloksessa (4.5) oleva käänteismatriisi $\mathbf{V}(\boldsymbol{\beta})^{-1}$ ei ole määritelty eikä siinä esitetty $\boldsymbol{\beta}$:n SU-estimaattorin asymptoottinen normaalisuus voi siten pätea. Tämä havainnollistaa ARMA-mallin identifiointiehdon tarpeellisuutta tilastollisen päättelyn kannalta. Vastaava tulos pätee yleisesti ARMA(p,q)-malleilla.

Huom.: Monisteen tuloksessa (4.5) olevan matriisin $\mathbf{V}(\boldsymbol{\beta})$ määritelmässä on painovirhe. Oikea määritelmä on $\mathbf{V}(\boldsymbol{\beta}) = \mathsf{E}\left[-\partial^2 l(\boldsymbol{\beta}, \sigma^2)/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'\right]$ eli log-uskottavuusfunktion argumenttiin pitää lisätä σ^2. Merkintä $\mathbf{V}(\boldsymbol{\beta})$ on silti perusteltu, koska σ^2 supistuu odotusarvon ottamisen myötä pois. Tähän liittyen, sivun puolivälissä oleva yhtälö $\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}) = -\partial^2 l(\hat{\boldsymbol{\beta}})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'$ on korjattava muotoon $\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}, \sigma^2) = -\partial^2 l(\hat{\boldsymbol{\beta}}, \sigma^2)/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'$ (tässä ei oteta odotusarvoa, joten σ^2 ei supistu pois).

Vihje.: Ratkaisussa voi voi käyttää monisteen s. 16 esitettyä AR(1)-prosessin varianssin kaavaa.

Kaksi seuraavaa tehtävää on tarkoitus ratkaista käyttäen joko JMulti- tai R-ohjelmistoa, jotka saa käyttöön kurssisivulla mainituista linkeistä. Aineistoina olevat aikasarjat US_GDP ja OscInd löytyvät myös kurssisivulta samoin kuin R-ohjelmistoa käytettäessä tarvittavat koodit ohjeineen (R-koodi_1 ja R-koodi_2). Ratkaisut (eli empiiristen analyysien tulokset) tulostetaan ja palautetaan harjoitustilaisuudessa.

Asialla ei välttämättä ole merkitystä alla olevien tehtävien ratkaisujen kannalta, mutta kurssisivulla olleen OscInd-tiedoston otsikossa oli sarjan päättymisajankohdassa painovirhe. Oikea ajankohta on 1994XII (eikä 1995XII) ja virhe on korjattu myös kurssisivulla ollevaan tiedostoon.

3. HT:n 4.3 malliratkaisussa todetaan, että estimoidun autokorrelaatio- ja osittaisautokorrelaatiofunktion perusteella AR(2)-malli voisi olla sopiva US_GDP-sarjalle.

(i) Estimoi AR(2)-mallin parametrit SU-menetelmällä.

(ii) Tutki estimoimasi mallin riittävyyttä monisteen s. 50 esitettyjä residuaalitarkasteluja käyttäen ja vertaa valittua mallia myös joihinkin "naapurimalleihin" SU-estimointiin perustuvien mallinvalintakriteerien ja tilastollisten testien avulla.

4. HT:n 4.4 malliratkaisussa todetaan, että estimoidun autokorrelaatio- ja osittaisautokorrelaatiofunktion avulla on vaikea valita sopivaa mallia OscInd-sarjalle. Koska molemmat vaimenevat melko nopeasti nollaan, voidaan pieniasteisia malleja pitää sopivina lähtökohtina.

(i) Kokeile edellä sanotun perusteella ARMA(p,q)-malleja, joissa p ja q ovat korkeintaan kaksi, valitse niistä sopivimmalta tuntuva ja estimoi valitsemasi mallin parametrit.

(ii) Tutki estimoimasi mallin riittävyyttä monisteen s. 50 esitettyjä residuaalitarkasteluja käyttäen ja, jos se osoittautuu riittämättömäksi, yritä löytää sopivampi vaihtoehto.