Stationaariset aikasarjat sl 2016, 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 (i) satunnaisvektorin $\mathbf{y}$ yhteistiheysfunktio, (ii) parametrin $\left(\boldsymbol{\varphi}, \boldsymbol{\beta}, \sigma^2\right)$ log-uskottavuusfunktio ja (iii) 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 käyttäen multinormaalijakauman ominaisuuksia lineaarisissa (tai affiineissa) muunnoksissa saadaan ...

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 $\mathbf{V}(\boldsymbol{\beta})$. 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 päteä. Tämä havainnollistaa ARMA-mallin identifiointiehdon tarpeellisuutta tilastollisen päättelyn kannalta. Vastaava tulos pätee yleisesti ARMA(p,q)-malleilla.

Vihje.: Ratkaisussa voi voi käyttää monisteen s. 14 esitettyä AR(1)-prosessin varianssin kaavaa.

Seuraavat tehtävät on tarkoitus ratkaista käyttäen kurssisivulta löytyvää R-koodia (R-koodi_2) ja aineistoja Lake ja Drought.

3. HT:n 4.2 malliratkaisussa todetaan, että estimoidun autokorrelaatio- ja osittaisautokorrelaatiofunktion perusteella AR(2)-malli voisi olla sopiva Lake-sarjalle.

(i) Tutki ARMA-mallin asteiden valintaa R-koodi_2:ssa olevan auto.arima-koodin avulla ja estimoi joko auto.ariman ehdottaman tai muilla perusteilla valitsemasi mallin parametrit SU-menetelmällä.

(ii) Tutki estimoimasi mallin riittävyyttä monisteen s. 50 esitettyjä residuaalitarkasteluja käyttäen (voit myös verrata valittua mallia joihinkin "naapurimalleihin" tilastollisten testien avulla). Jos valintasi osoittautuu puutteelliseksi, yritä löytää sopivampi vaihtoehto.

(iii) Laske ennusteet valitsemallesi määrälle sarjan tulevia arvoja.

Huom.: R-koodi_2:ssa oleva auto.arima-koodi tulostaa AIC- ja BIC-kriteerien arvojen lisäksi AICC-kriteerin arvon, jossa sakkofunktio $g(T) = 2(p+q+1)/(T-p-q-2)$, kun määritelmässä noudatetaan kurssilla käytettyä tapaa. AICC-kriteeriä tarkastellaan lähemmin kurssisivulla mainitun Brockwell'in ja Davis'in kirjan luvuissa 5.2 ja 5.5.2. Huomaa, että Brockwell'in ja Davis'in (kuten joidenkin muidenkin) mallinvalintakriteerit saadaan kurssilla esitetyistä kertomalla havaintojen lukumäärällä.

4. HT:n 4.3 malliratkaisussa todetaan, että estimoidun autokorrelaatio- ja osittaisautokorrelaatiofunktion perusteella AR(5)-malli voisi olla sopiva Drought-sarjalle, joskin myös AR(1) "voisi olla hyvä arvaus". Rakenna ARMA-malli Drought-sarjalle menettelemällä kuten HT:ssä 3.