Stationaariset aikasarjat sl 2014, HT 5, viikko 41

Tehtävät 1 ja 2 liittyvät monisteen s. 34 ja 35 taitteessa mainittuun mahdollisuuteen ennustaa aikasarjan tulevia arvoja äärellisen realisaation avulla.

1. Tarkastellaan satunnaisvektoria $\mathbf{Z} = (Y, \mathbf{X})$, jossa Y on reaalinen ja $\mathbf{X}$ on $n \times 1$ vektori. Ositetaan $\mathbf{Z}$:n odotusarvo ja kovarianssimatriisi vastaavasti

$$\boldsymbol{\mu}_{\mathbf{Z}} = \begin{bmatrix} \mu_Y \\ \boldsymbol{\mu}_{\mathbf{X}} \end{bmatrix} \quad \text{ja} \quad \boldsymbol{\Sigma}_{\mathbf{Z}} = \begin{bmatrix} \sigma_Y^2 & \sigma_{Y\mathbf{X}} \\ \sigma_{\mathbf{X}Y} & \boldsymbol{\Sigma}_{\mathbf{X}} \end{bmatrix} \quad (\sigma_{Y\mathbf{X}} = \sigma'_{\mathbf{X}Y}).$$

Oletetaan $\boldsymbol{\Sigma}_{\mathbf{X}}$ positiivisesti definiitiksi ja määritellään $U = Y - \mu_Y - \sigma_{Y\mathbf{X}}\boldsymbol{\Sigma}_{\mathbf{X}}^{-1}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})$, jolloin $Y = \mu_Y + \sigma_{Y\mathbf{X}}\boldsymbol{\Sigma}_{\mathbf{X}}^{-1}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) + U$. Osoita seuraavat tulokset:

(a) $\mathsf{E}(U) = 0$ ja $\mathsf{Var}(U) = \sigma_Y^2 - \sigma_{Y\mathbf{X}}\boldsymbol{\Sigma}_{\mathbf{X}}^{-1}\sigma_{\mathbf{X}Y}$

(b) U ja $\mathbf{X}$ ovat korreloimattomia eli $\mathsf{Cov}(U, \mathbf{X}) = 0$

Vihje: Tehtävä voidaan ratkaista suoraan laskemalla käyttäen satunnaisvektorien odotusarvoa ja kovarianssimatriiseja koskevia tuloksia (ks. esim. Koistinen, P.: Todennäköisyyslaskenta, syksy 2013, luku 9.2 ja erityisesti tulokset (9.6) ja (9.7) sekä Lause 9.1).

2. Jatkoa edelliselle. Osoita, että millä tahansa ei-satunnaisella skalaarilla b ja $n \times 1$ vektorilla $\mathbf{c}$ pätee

$$\mathsf{E}\left(U^2\right) \le \mathsf{E}\left[\left(Y - b - \mathbf{c}'\mathbf{X}\right)^2\right].$$

Vihje: Muokkaa ensin lauseketta $Y - b - \mathbf{c}'\mathbf{X}$ lisäämällä ja vähentämällä sopivasti termejä, jotta saat lausuttua sen muodossa $Y - b - \mathbf{c}'\mathbf{X} = U + a + \left(\sigma_{Y\mathbf{X}}\boldsymbol{\Sigma}_{\mathbf{X}}^{-1} - \mathbf{c}'\right)(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})$, jossa skalaari a on ei-satunnainen. Suorita tämän jälkeen potenssiin korotus ja laske odotusarvo käyttäen edellisen tehtävän tuloksia.

Huom.: Tuloksen tulkinta on, että keskineliövirheen mielessä paras Y:n satunnaisvektoriin $\mathbf{X}$ perustuva *lineaarinen* ennuste on $\mu_Y + \sigma_{Y\mathbf{X}}\boldsymbol{\Sigma}_{\mathbf{X}}^{-1}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})$. Jos $\mathbf{Z}$ on multinormaalinen, on ennuste Y:n ehdollinen odotusarvo ehdolla $\mathbf{X}$ eli ennuste on paras kaikkien ennusteiden joukossa.

Jos y_t on stationaarinen ARMA(p,q)-prosessi odotusarvona nolla, niin valitsemalla $Y = y_{T+h}$ ja $\mathbf{X} = (y_T, ..., y_1)$ pätee $\boldsymbol{\mu}_{\mathbf{Z}} = \mathbf{0}$, $\sigma_{Y\mathbf{X}} = \left(\gamma_h, \gamma_{h+1}, ..., \gamma_{h+T-1}\right)$ ja $\boldsymbol{\Sigma}_{\mathbf{X}} = \left[\gamma_{i-j}\right]_{i,j=1,...,T}$, jossa kovarianssit voidaan laskea ARMA(p,q)-prosessin parametrien avulla kuten monisteen s. 31 on kuvattu. Tällöin edellä todetusta nähdään, miten ARMA(p,q)-prosessin optimaalinen lineaarinen ennuste muodostetaan, kun ennustamisessa käytetään prosessin äärellistä menneisyyttä (vrt. monisteen keskustelu s. 34 ja 35 taitteessa). Laskelmien helpottamiseksi on (mahdollisesti suuren) käänteismatriisin $\boldsymbol{\Sigma}_{\mathbf{X}}^{-1}$ muodostamiseksi kehitetty useita algoritmeja.

3. Olkoon y_t heikosti stationaarinen prosessi, jonka odotusarvo on nolla.

(a) Osoita edellisen tehtävän avulla, että y_{t+h}:n havaintoon y_t perustuva paras lineaarinen ennuste on $\rho_h y_t$, jossa ρ_h on prosessin h. autokorrelaatiokerroin.

(b) Oletetaan nyt, että y_t on heikosti stationaarinen AR(2)-prosessi

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim \mathsf{wn}(0, \sigma^2).$$

Mikä on y_{t+1}:n paras vektoriin (y_t, y_{t-1}) perustuva lineaarinen ennuste ja vastaava ennustevirheen varianssi?

(c) Oletetaan, että y_t on AR(2)-prosessi kuten kohdassa (b). Laske kohdasta (a) tapauksessa $h = 1$ saatavan y_{t+1}:n ennusteen ennustevirheen varianssi ja vertaa sitä kohdassa (b) saamaasi ennustevirheen varianssiin.

Huom.: Käytä kohdassa (c) (etenkin sen loppuosassa) tietoa, että AR(2)-prosessin varianssi γ_0 ja ensimmäinen autokorrelaatio ρ_1 voidaan lausua parametrien funktioina muodossa

$$\gamma_0 = \frac{(1-\phi_2)\,\sigma^2}{(1+\phi_2)\left[(1-\phi_2)^2 - \phi_1^2\right]} \quad \text{ja} \quad \rho_1 = \frac{\phi_1}{1-\phi_2}.$$

4. Tarkastellaan ARIMA(1,1,0)-prosessia $\Delta y_t = \phi \Delta y_{t-1} + \varepsilon_t, \;\; t = 1, 2, ..., \; \varepsilon_t \sim \mathsf{iid}\left(0, \sigma^2\right), \; |\phi| < 1$. Merkitään $\mathsf{E}_t(\cdot) = \mathsf{E}(\cdot\,|y_t, ..., y_0)$ $(t \geq 1, h \geq 1)$.

(a) Osoita, että $\mathsf{E}_t(y_{t+h}) - \mathsf{E}_t(y_{t+h-1}) = \phi^h \Delta y_t$

(b) Osoita, että $\mathsf{E}_t(y_{t+h}) = y_t + \dfrac{\phi\left(1-\phi^h\right)}{(1-\phi)} \Delta y_t$

(c) Oletetaan nyt, että $\phi = 0$. Laske ennustevirheen $\mathsf{E}_t(y_{t+h}) - y_t$ odotusarvo ja varianssi. Mitä tapahtuu ennustevirheen varianssille ja edelleen ennusteen oletukseen $\varepsilon_t \sim \mathsf{nid}\left(0, \sigma^2\right)$ perustuvalle luottamusvälille, kun $h \to \infty$?

Vihje: Menettele (a)-kohdassa kuten stationaarisen AR(1)-prosessin tapauksessa monisteen s. 35. (b)-kohdassa yksi mahdollisuus on ratkaista (a)-kohdassa saatu differenssiyhtälö.