

HARJOITUS 3

MALLIRATKAISUT

1. Tehtävässä tarkastellaan erään Lydia E. Pinkham -lääkeyhtiön valmistaman tuotteen myyntitulojenn ja sen mainontaan käytettyjen kulujen yhteyttä vuosina 1907 - 1960. Merkitään (kaksiulotteisen) aineiston vuoden t myyntituloja y_{1t} :llä ja mainoskuluja y_{2t} :llä.

(i) Selvitetään nyt monisteen sivujen 32–33 mallinvalintakriteerejä käyttäen sopiva VAR(p)-mallin aste p . Mallinvalintakriteerien tarkoituksena on etsiä ”sopivin” parametrien lukumäärä. Tilastotieteessä puhutaan yleisesti *säästäväisyysperiaatteesta*, jolla tarkoitetaan sitä, että liian suurta parametrien määrä halutaan välttää, koska se heikentää estimointitarkkuutta. Toisaalta on mahdollista, että pienellä määrällä parametreja aineistoa ei pystytä kuvaamaan kunnolla. Mallinvalintakriteerien avulla pyritään löytämään tasapaino näiden kahden tavoitteen välillä.

Tässä tapauksessa kaikki monisteessa mainitut kriteerit (AIC, BIC ja HQ) ehdottavat mallin asteeksi kolmea, jota käyttäen saadaan estimoitu VAR(3)-malli

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 215.87 \\ (100.48) \\ -64.42 \\ (80.59) \end{bmatrix} + \begin{bmatrix} 1.37 & -0.16 \\ (0.17) & (0.18) \\ 0.59 & 0.52 \\ (0.14) & (0.14) \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} -0.31 & -0.28 \\ (0.27) & (0.21) \\ -0.54 & -0.34 \\ (0.21) & (0.16) \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} \\ + \begin{bmatrix} 0.00 & 0.10 \\ (0.18) & (0.19) \\ 0.26 & 0.27 \\ (0.15) & (0.15) \end{bmatrix} \begin{bmatrix} y_{1,t-3} \\ y_{2,t-3} \end{bmatrix} + \begin{bmatrix} \hat{\varepsilon}_{1t} \\ \hat{\varepsilon}_{2t} \end{bmatrix}.$$

Estimoituja kertoimia ja keskihajontoja vertailemalla havaitaan, että jälkimmäisen yhtälön (mainoskulut) useimmat estimaatit ovat ”suuria” (likimääräisiin) keskivirheisiin verrattuna, kun taas edellisessä yhtälössä (myyntitulo) useimmat estimaatit ovat ”pieniä”. Keskivirheitä näin käytettäessä on kuitenkin syytä muistaa, että mallissa on olettu virheiden iid-ominaisuus, jonka realistisuutta tarkastellaan seuraavissa laskuharjoituksissa. Todetaan vielä, että residuaalien välinen korrelaation on 0.5 ja että estimoidusta VAR(3)-mallista lasketun matriisin \mathbf{A} itseisarvoltaan suurin juuri on 0.90, joka on riittävän kaukana ykkösestöstä, jotta stationaarisuusehtoa voidaan pitää perusteltuna.

(ii) Tutkitaan nyt VAR-mallin astetta käyttäen uskottavuusosamäärätestejä kuten monisteen s. 31. Koska edellä mallinvalintakriteerit päätyivät asteeseen $p = 3$, valitaan maksimiasteeksi tätä hieman suurempi $P = 5$. Testejä tehtäessä on tärkeää käyttää samaa havaintoaineistoa jokaisessa estimoinnissa. Menettelemällä kuten monisteen s. 31 saadaan seuraavassa taulukossa esitetyt tulokset.

Testisuure	LR(5)	LR(4)	LR(3)
p-arvo	0.45	0.34	0.001

Kukin taulukon LR-testisuure noudattaa asympotoottisesti χ_4^2 -jakaumaa, josta p-arvot on laskettu. Tulokset tukevat astetta $p = 3$, joten uuteen estimointiin ei ole tarvetta.

2. Ennustetaan seuraavaksi edellisessä tehtävässä tarkasteltua kahta aikasarjaa estimoidulla VAR(3)-malli vuosien 1907–1955 aineistolla ja laskemalla ennusteet vuosille 1956–1960 (ks. monisteen s. 13–15). Näin saadaan rekursiivinen ennuste aina viimeiseen vuoteen eli vuoteen 1960 asti.

Kuvasta nähdään aikasarjojen reaalisatiot mustalla, ennuste sinisellä ja 95 % luottamusväli punaisella katkoviivalla. Vuosien 1956–1960 havainnot osuvat ennusteen luottamusvälin sisälle. Vertaamalla ennusteita toteutuneisiin arvoihin voidaan havaita, että ennusteet eivät kykene havaitsemaan sarjojen jatkuvaa laskua. Toisaalta luottamusvälit ovat suuria, mikä ei ole yllättävää mm. siitä syystä, että myyntitulon yhtälössä on usieta pieniä estimaatteja.

3. Tutkitaan seuraavaksi saksalaisten yritysten liiketoiminnan odotuksia kuvaavasta indeksistä (y_{1t}) ja Saksan tavaroiden tuotantoindeksin kuukausittaisista muutoksista (y_{2t}) koostuvaa kaksiulotteista aikasarjaa ajanjaksolta 1991I–2007XII.

(i) Valitaan VAR-mallin aste käyttäen hyväksi mallinvalintakriteereitä. Sekä AIC että HQ suosittelevat astetta $p = 3$, kun taas BIC suosittelee pienempää astetta $p = 2$. Estimoidaan VAR(3)-malli ja tutkitaan asteen valintaa uudestaan seuraavassa kohdassa. Estimoitu malliyhtälö on

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 10.91 \\ (2.36) \\ -8.43 \\ (1.87) \end{bmatrix} + \begin{bmatrix} 1.20 & 0.08 \\ (0.07) & (0.09) \\ 0.15 & -0.48 \\ (0.05) & (0.07) \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} -0.07 & 0.20 \\ (0.11) & (0.10) \\ -0.03 & 0.26 \\ (0.09) & (0.08) \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} \\ + \begin{bmatrix} -0.25 & 0.20 \\ (0.07) & (0.09) \\ -0.03 & 0.06 \\ (0.05) & (0.07) \end{bmatrix} \begin{bmatrix} y_{1,t-3} \\ y_{2,t-3} \end{bmatrix} + \begin{bmatrix} \hat{\varepsilon}_{1t} \\ \hat{\varepsilon}_{2t} \end{bmatrix}.$$

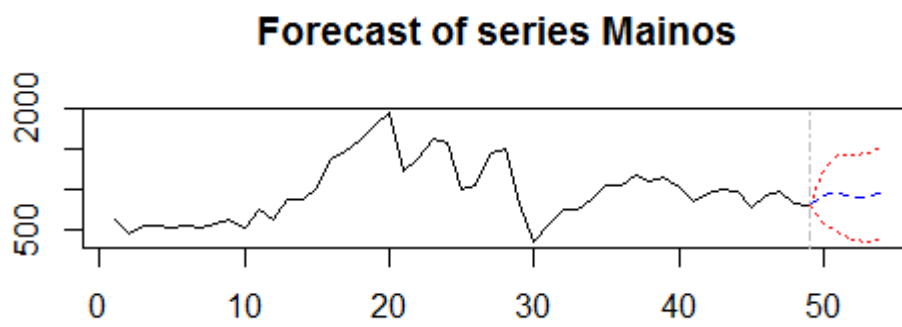
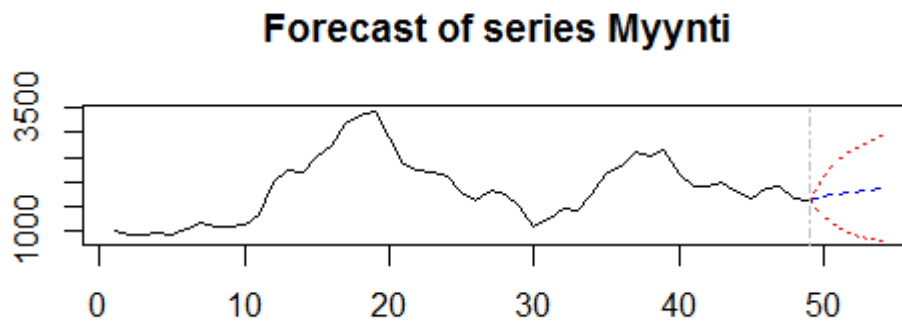
Molemmissa yhtälöissä on keskivirheisiin verrattuna joitakin pieniä estimaatteja ja jälkimmäisessä (tuotantoindeksin muutokset) niitä on parametrin A_3 estimaatissa, mikä selittää BIC:n päätymistä asteeseen $p = 2$. Estimointitulokset tukevat stationaarisuusoletusta, sillä estimoidusta VAR(3)-mallista lasketun matriisin \mathbf{A} itseisarvoltaan suurin juuri on (ykköistä pienempi) 0.81. Residuaalien välinen korrelaation on heikkoa, vain 0.10.

(ii) Tutkitaan nyt mallin asteen valintaa uskottavuusosamäärätestejä kuten tehtävässä 2(ii) ja valitaan jälleen maksimiasteeksi $P = 5$. Tulokseksi saadaan

Testisuure	LR(5)	LR(4)	LR(3)
p-arvo	0.25	0.54	0.003

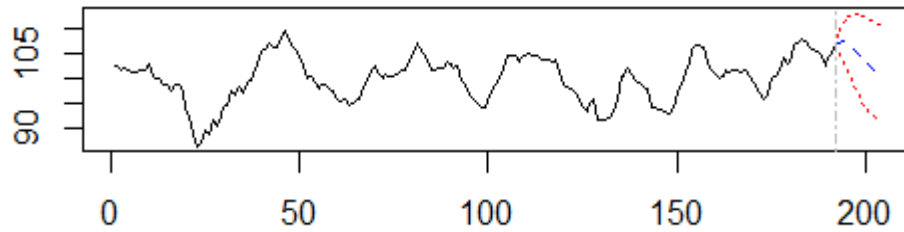
Tämä tukee edellisen kohdan valintaa $p = 3$, joten uuteen estimointiin ei ole tarvetta.

4. Ennustetaan seuraavaksi edellisessä tehtävässä tarkasteltua kahta aikasarjaa estimoimalla VAR(3)-malli käyttäen havaintoja vuoden 2006 joulukuuhun saakka. Mallin estimoinnin jälkeen muodostetaan ennusteet vuoden 2007 jokaiselle kuukaudelle. Kuvaan 2 on piirretty nämä aikasarjat ja niiden ennusteet (siniset katkoviivat) ja lisäksi ennusteiden 95 %:n luottamusrajat (punaiset katkoviivat). Havaintojen voidaan jälleen nähdä sisältyvän luottamusvälin sisään. Vertaamalla ennusteita toteutuneisiin arvoihin voidaan havaita, että odotusindeksin kohdalla ennuste seuraa toteutunutta indeksiä varsin hyvin. Tuotannon muutoksissa ennuste ei ota huomioon aikasarjan voimakasta vaihtelua, vaikka seuraakin toteutunutta laskevaa trendiä melko hyvin.

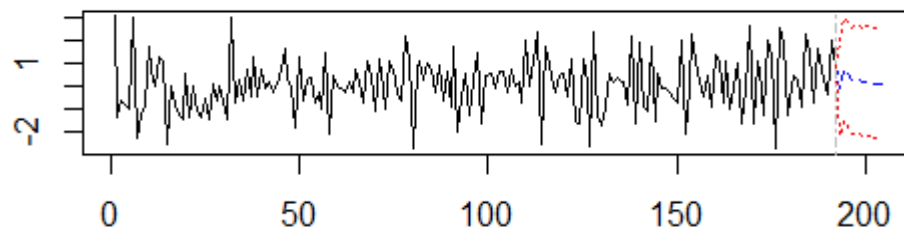


Kuva 1: Kuviin on piirretty myyntituotot sekä mainosmenot. Sininen katkoviiva kuvaa VAR(3)-mallin tuottamia ennusteita vuosille 1956-1960 ja punaiset katkoviivat ennusteiden 95 %:n luottamusväliä.

Forecast of series Odotukset



Forecast of series Tuotanto



Kuva 2: Kuviin on piirretty yritysten odotuksia kuvaava indeksi sekä tuotantoindeksin muutos. Sininen katkoviiva kuvaa VAR(3)-mallin tuottamia ennusteita vuodelle 2007 ja punaiset katkoviivat ennusteiden 95 %:n luottamusväliä.