ARMA-mallien rakentaminen

- ARMA(p,q)-prosessi määritellään yhtälöllä

  $$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}, \;\; \varepsilon_t \sim \mathrm{iid}(0, \sigma^2)$$

  jossa $1 - \phi_1 z - \cdots - \phi_p z^p \neq 0, \;\; |z| \leq 1$ (stationaarisuusehto) ja $1 + \theta_1 z - \cdots + \theta_q z^q \neq 0, \;\; |z| \leq 1$ (käännettävyysehto)

- Oletetaan, että havaittu aikasarja $y_1, ..., y_T$ voidaan (mm. kuvion perusteella) tulkita stationaarisen prosessin tuottamaksi ja sen mallintaminen jollain ARMA(p,q)-prosessilla tuntuu kohtuulliselta.

- Sationaarisuusoletus saattaa vaatia aikasarjan muuntamista.

- Tavallisimmat esimerkit muunnoksista ovat differensointi (mahdollisesti logaritmointiin yhdistettynä) ja deterministisen trendin eliminointi PNS-menetelmällä.

ARMA-mallien rakentaminen

- Sopivan ARMA(p,q)-prosessin tai, kuten mallintamisen yhteydessä tavallisesti sanotaan, ARMA(p,q)-mallin löytäminen on perinteisesti ajateltu koostuvan seuraavista toisiinsa liittyvistä vaiheista:

  1) Mallin asteiden p ja q valinta
  2) Parametrien alustava estimointi
  3) Parametrien tehokas estimointi
  4) Estimoidun mallin riittävyyden tutkiminen eli mallin diagnostiikka

- Jos estimoitu malli havaitaan vaiheessa 4 puutteelliseksi, on valittava uusi malli (uudet asteet), estimoitava valitun mallin parametrit ja tutkittava estimoidun mallin riittävyyttä.
- Tyypillinen puutteellisuus mallissa on virheiden ε_t autokorreloituneisuus, jolloin asteista p ja q (ainakin) toinen on valittu liian pieneksi.

ARMA-mallien rakentaminen

- Kannattaisiko ARMA(p,q)-mallin asteet valita varmuuden vuoksi niin suuriksi, että virhetermin autokorreloituneisuutta ei ole mitään syytä epäillä?
- Tähän liittyvä ongelma on identifioituvuus eli voidaan päätyä malliin

$$\delta(B)\,\phi(B)\,y_t = \delta(B)\,\theta(B)\,\varepsilon_t, \quad \delta(B) = 1 + \delta_1 B + \cdots + \delta_r B^r,$$

  jota ei voida erottaa mallista

$$\phi(B)\,y_t = \theta(B)\,\varepsilon_t, \quad \varepsilon_t \sim \text{iid}\left(0, \sigma^2\right).$$

- Tällöin ensin mainitun mallin parametreja ei voida identifioida, joten estimointi on epävakaata, mikä heijastuu epätarkkuutena ennustamisessa.
- Kannattaa suosia ns. *säästäväisyysperiaatetta* eli pyrkiä valitsemaan mahdollisimman yksinkertainen, mutta kuitenkin riittävä malli.

- Aikasarjan kuvan piirtämisen jälkeen on hyvä tutustua aineistoon tutkimalla sarjan autokorrelaatio-ominaisuuksia estimoidun autokorrelaatio- ja osittaisautokorrelaatiofunktion avulla.
- Tämä saattaa antaa joissakin tapauksissa vihjeitä sopivista mallin asteista. Erityisesti, jos havaitaan katkoksia.
- Katkos autokorrelaatiofunktiossa (osittaisautokorrelaatiofunktiossa) viittaa MA-prosessiin (AR-prosessiin). Jos katkoksia ei ole, voidaan epäillä ARMA-prosessia.
- Usein selvää johtopäätöstä on vaikea tehdä, mutta joitakin mallivaihtoehtoja voidaan ehkä sulkea pois.

Havaittu aikasarja $y_1, ..., y_T$. Mallina on nyt

$$y_t - \mu \quad = \quad \phi_1 \left(y_{t-1} - \mu\right) + \cdots + \phi_p \left(y_{t-p} - \mu\right) + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{iid}(0, \sigma^2).$$

Jakso 3.1: Parametrien $\boldsymbol{\phi} = (\phi_1, ..., \phi_p)$ ja σ^2 Yule-Walker -estimaatit

$$\tilde{\boldsymbol{\phi}}_{\mathrm{YW}} \quad = \quad \begin{bmatrix} c_0 & c_1 & \cdots & c_{p-1} \\ c_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_1 \\ c_{p-1} & \cdots & c_1 & c_0 \end{bmatrix}^{-1} \begin{bmatrix} c_1 \\ \vdots \\ \vdots \\ c_p \end{bmatrix} = \mathbf{C}_p^{-1} \mathbf{c}_p,$$

$$\tilde{\sigma}^2_{\mathrm{YW}} \quad = \quad c_0 - \tilde{\boldsymbol{\phi}}_{\mathrm{YW,1}} c_1 - \cdots - \tilde{\boldsymbol{\phi}}_{\mathrm{YW,p}} c_p,$$

jossa $c_h = (T-h)^{-1} \sum_{t=1}^{T-h} (y_t - \bar{y}) \left(y_{t+h} - \bar{y}\right)$ on otosautokovarianssi.

- Voidaan osoittaa, että

$$\tilde{\boldsymbol{\phi}}_{\text{YW}} \underset{as}{\sim} \mathrm{N}\left(\boldsymbol{\phi}, T^{-1}\sigma^2\boldsymbol{\Gamma}_p^{-1}\right),$$

jossa $\boldsymbol{\Gamma}_p = \left[\gamma_{i-j}\right]_{i,j=1,\dots,p}$ ja $\tilde{\sigma}^2_{\text{YW}}\mathbf{C}_p^{-1} \xrightarrow{p} \sigma^2\boldsymbol{\Gamma}_p^{-1}$.

- Lisäksi, jos $\varepsilon_t \sim \mathrm{nid}\left(0, \sigma^2\right)$, niin yo asymptoottinen jakauma on sama kuin SU-estimaattorin asymptoottinen jakauma, joten Yule-Walker estimaattori $\tilde{\boldsymbol{\phi}}_{\text{YW}}$ on asymptoottisesti tehokas.

- Yo asymptoottista jakaumatulosta käyttäen voidaan muodostaa Waldin testejä ja luottamusvälejä tavanomaiseen tapaan.

- Kun $\mathrm{E}\left(y_t\right) = \mu \neq 0$ sallitaan, voidaan AR(p)-malli esittää myös käyttäen parametointia

$$y_t = \nu + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{iid}\left(0, \sigma^2\right),$$

jossa $\nu = \mu - \phi_1 \mu - \cdots - \phi_p \mu$.

- Tuntuu luontevalta käyttää parametrien $\boldsymbol{\phi}$ ja ν estimointiin PNS-menetelmää eli minimoidaan neliösummafunktio

$$\mathrm{S}\left(\nu, \boldsymbol{\phi}\right) = \sum_{t=p+1}^{T} \left(y_t - \nu - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p}\right)^2.$$

- Jos $\tilde{\nu}_{PNS}$ ja $\tilde{\boldsymbol{\phi}}_{PNS}$ ovat näin saatuja PNS-estimaattoreita, voidaan σ^2 estimaattoriksi valita $\mathrm{S}(\tilde{\nu}_{PNS}, \tilde{\boldsymbol{\phi}}_{PNS}) / \left(T - p\right)$ tai $\mathrm{S}(\tilde{\nu}_{PNS}, \tilde{\boldsymbol{\phi}}_{PNS}) / \left(T - 2p - 1\right)$.

- Voidaan osoittaa, että PNS-estimaattorilla $\tilde{\boldsymbol{\phi}}_{PNS}$ on sama asymptoottinen jakauma kuin Yule-Walker estimaattorilla, joten se on asymptoottisesti yhtäpitävä normaalijakaumaan perustuvan SU-estimaattorin kanssa ja siten asymptoottisesti tehokas.
- Lisäksi, kaikki tavanomaiset lineaarisen mallin testeihin ja luottamusväleihin liittyvät tulokset ja menetelmät pätevät *likimääräisesti, mutta eivät äärellisillä T:n arvoilla tarkasti*.
- PNS-estimaattori poikkeaa SU-estimaattorista vain siinä, että havainnot $y_1, \ldots, y_p$ tulkitaan ei-satunnaisiksi vakioiksi eli ehdollistetaan niiden suhteen.

- Oletetaan stationaarisuus ja käännettavyys sekä $\mathrm{E}(y_t) = 0$.

$$
\begin{aligned}
y_t &= \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}, \\
\varepsilon_t &\sim \mathrm{iid}\left(0, \sigma^2\right).
\end{aligned}
$$

- Jos viivästetyt innovaatiot $\varepsilon_{t-1}, ..., \varepsilon_{t-q}$ tunnettaisiin, voitaisiin menetellä kuten AR(p)-mallin tapauksessa ja estimoida parametrit $\phi_1, ..., \phi_p$ ja $\theta_1, ..., \theta_q$ yksinkertaisesti PNS-menetelmällä.

- Hannanin ja Rissasen menetelmän ideana on korvata viivästetyt innovaatiot empiirisillä vastineilla, jotka perustetaan y_t:n AR(∞)-esityksen AR(m)-approksimaatioon (m "suuri"). Tämä vaatii mallin käännettävyyden.

- Olkoon

$$\tilde{\varepsilon}_t = y_t + \tilde{\pi}_1 y_{t-1} + \cdots + \tilde{\pi}_m y_{t-m}, \quad t = m+1, ..., T,$$

y_t:n AR(m)-approksimaatioon perustuvat PNS- tai Yule-Walker -residuaalit.

- Asetetaan $n = \max\{m+p+1, m+q+1\}$ ja muodostetaan "apumalli" $(t = n, ..., T)$

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \tilde{\varepsilon}_{t-1} + \cdots + \theta_q \tilde{\varepsilon}_{t-q} + a_t,$$

josta parametrit $\boldsymbol{\phi} = (\phi_1, ..., \phi_p)$ ja $\boldsymbol{\theta} = (\theta_1, .., \theta_q)$ estimoidaan PNS-llä eli minimoimalla neliösummafunktio

$$\tilde{S}(\boldsymbol{\phi}, \boldsymbol{\theta}) = \sum_{t=n}^{T} \left(y_t - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p} - \theta_1 \tilde{\varepsilon}_{t-1} - \cdots - \theta_q \tilde{\varepsilon}_{t-q} \right)^2$$

Parametrille $\boldsymbol{\beta} = (\boldsymbol{\phi}, \boldsymbol{\theta})$, saadaan siten estimaattori

$$\tilde{\boldsymbol{\beta}}_{HR} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y},$$

jossa $\mathbf{y} = \begin{bmatrix} y_n & \cdots & y_T \end{bmatrix}'$ ja

$$\mathbf{X} = \begin{bmatrix} y_{n-1} & \cdots & y_{n-p} & \tilde{\varepsilon}_{n-1} & \cdots & \tilde{\varepsilon}_{n-q} \\ \vdots & & \vdots & \vdots & & \vdots \\ y_T & \cdots & y_{T-p} & \tilde{\varepsilon}_{T-1} & \cdots & \tilde{\varepsilon}_{T-q} \end{bmatrix}.$$

Innovaatiovarianssi σ^2 estimoidaan suureella $\tilde{\mathrm{S}}(\tilde{\boldsymbol{\beta}}_{HR}) / (T - n)$.

- Estimaattorin $\tilde{\beta}_{HR}$ toimivuus vaatii, että viipymä m valitaan "tarpeeksi suureksi", jotta AR(m)-approksimaatio tai $e_t \approx \varepsilon_t$ toimii "kohtuullisen hyvin".
- Toisaalta m ei saa olla "liian suuri" suhteessa havaintojen lukumäärään T, koska muutoin AR(m)-approksimaatiossa estimoidaan liian paljon parametreja suhteessa havaintojen lukumäärään ja $\tilde{\beta}_{HR}$ perustuu liian pieneen havaintomäärään.
- Käytännössä m:n arvo voidaan valita tutkimalla estimoitua osittaisautokorrelaatiofunktiota (vrt. jakso 3.1) tai ns. mallinvalintakriteerejä (tai molempia).
- **HUOM.**: *Estimaattori $\tilde{\beta}_{HR}$ ei ole (asymptoottisesti) tehokas eivätkä tavanomaiset lineaarisen mallin F- ja t-testit ole (edes asymptoottisesti) päteviä*. Sitä on tarkoitus käyttää (korkeintaan) edellä mainittujen mallinvalintakriteerien kanssa sekä tuottamaan alkuarvoja numeerisia menetelmiä vaativassa SU-estimoinnissa.

- Tavoitteena on valita (stationaarisen ja käännettävän) ARMA(p,q)-mallin

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}, \\ \varepsilon_t &\sim \text{iid}\left(0, \sigma^2\right). \end{aligned}$$

  asteet p ja q, kun "riittävän suuret" maksimiasteet p^* ja q^* on kiinnitetty.

- Lineaarisen mallin teorian perusteella voitaisiin ajatella menettelyä, jossa Hannanin ja Rissasen "apumallin" $(t = n, ..., T)$

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \tilde{\varepsilon}_{t-1} + \cdots + \theta_q \tilde{\varepsilon}_{t-q} + a_t$$

  virhetermin a_t estimoitua varianssia $\tilde{\sigma}^2_{p,q}$ $(0 \leq p \leq p^*, 0 \leq q \leq q^*)$ minimoitaisiin p:n ja q:n suhteen.

- Jos "apumallissa" $(t = n, ..., T)$

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \tilde{\varepsilon}_{t-1} + \cdots + \theta_q \tilde{\varepsilon}_{t-q} + a_t$$

  p ja/tai q kasvaa, niin (PNS:ään perustuva) residuaalineliösumma ja siten myös $\tilde{\sigma}^2_{p,q}$ pienenee. (Olettaen, että $n \geq \max(p^*, q^*)$ pidetään samana kaikilla kokeiltavilla p:n ja q:n arvoilla.)
- Siis, $\tilde{\sigma}^2_{p,q}$:n minimointi p:n ja q:n suhteen $(0 \leq p \leq p^*,\ 0 \leq q \leq q^*)$ johtaa väistämättä laajimpaan mahdolliseen malliin.
- Ratkaisu: Valitaan asteet minimoimalla funktio

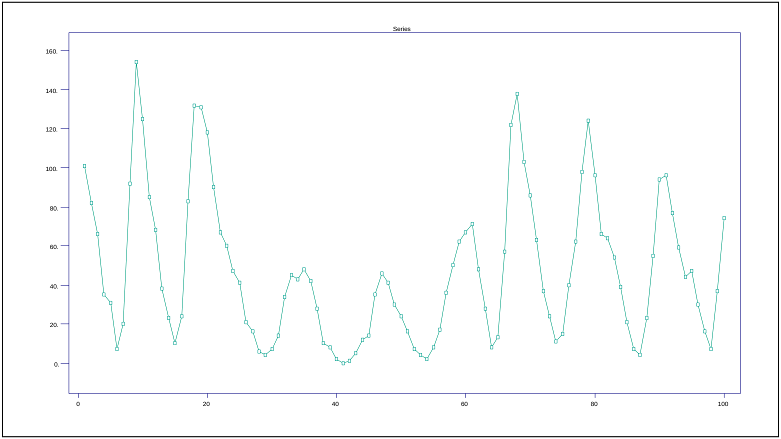
$$C_{HR}(p, q) = \log \tilde{\sigma}^2_{p,q} + (p+q)\, g(T) / T, \quad 0 \leq p \leq p^*, 0 \leq q \leq q^*,$$

  jossa ns. sakkofunktio $g(\cdot)$ on positiivinen ja $g(T)/T \to 0$, kun $T \to \infty$.

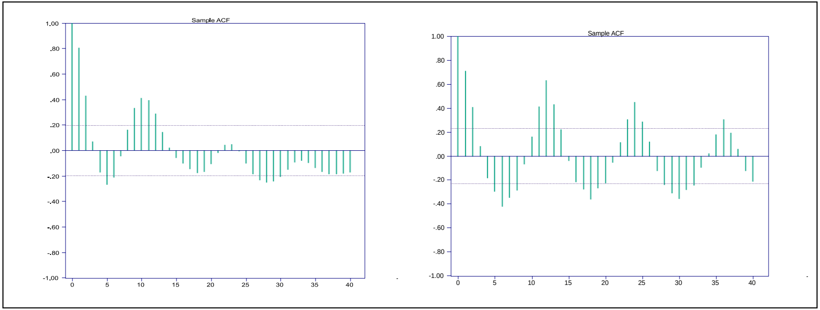
- Sakkofunktion idea on rankaista tarpeettoman laajan mallin käyttämisestä. Jos asteen p tai q kasvattaminen ei pienennä estimaattorin $\tilde{\sigma}^2_{p,q}$ arvoa tarpeeksi, ei laajempaa mallia valita.
- Tunnettuja sakkofunktioita ovat

$$
\begin{array}{lll}
\text{AIC:} & g(T) = 2 & \text{(Akaike, 1974)} \\
\text{HQ:} & g(T) = 2\log(\log T) & \text{(Hannan ja Quinn, 1979)} \\
\text{BIC:} & g(T) = \log T & \text{(Schwarz, 1978, Rissanen, 1978).}
\end{array}
$$

- Näistä ensimmäinen sakottaa vähiten (suosii laajempia malleja) ja viimeinen eniten (suosii suppeampia malleja). Kriteerifunktiosta HQ on myös versioita, joissa vakion 2 paikalla on joku muu vakio.
- Käytännössä ei ole suositeltavaa minimoida kriteerifunktiota $C_{HR}(p, q)$ mekaanisesti, vaan käyttää vain yhtenä mallin valinnan apuvälineenä.



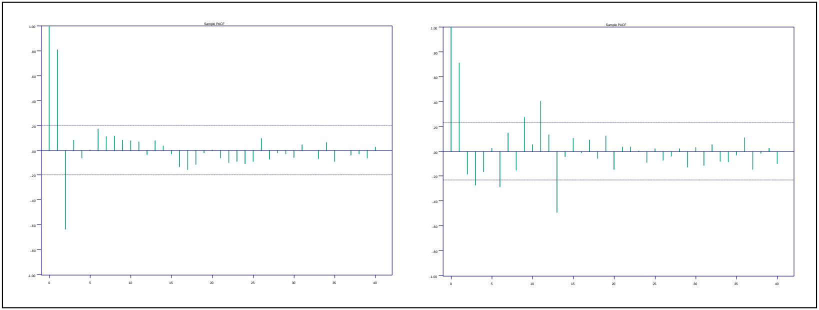
Kuvio 1.4. Vuotuisten auringonpilkkujen lukumäärä vuosilta 1770 - 1870.



Kuvio 2.1. Kuvion 1.4 auringonpilkkusarjan otosautokorrelaatiofunktio (vas.) ja Kuvion 1.2 onnettomuussarjan otosautokorrelaatiofunktio (oik.) viipymillä $h = 0, ..., 40$.

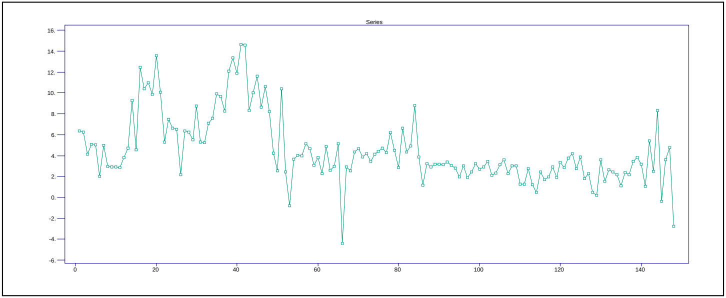
ARMA-mallien rakentaminen

Hannanin ja Rissasen mallinvalintamenetelmä



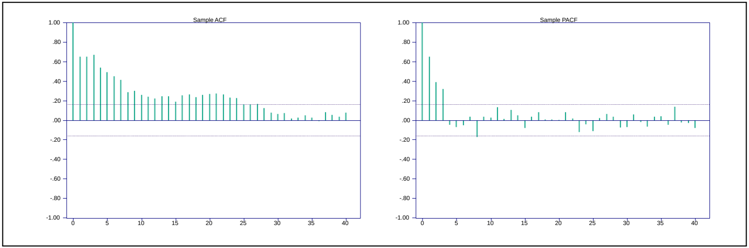
Kuvio 3.2. Kuvion 1.4 auringonpilkkusarjan otososittaisautokorrelaatio-funktio (vas.) ja Kuvion 1.2 onnettomuussarjan otososittaisautokorrelaatio-funktio (oik.) viipymillä $h = 0, \ldots, 40$.

- Estimoidun osittaisautokorrelaatiofunktion perusteella AR(2)-malli vaikuttaa sopivalta auringonpilkkusarjalle.
- Hannanin ja Rissasen menetelmä valinnoilla $m = 10$ ja $p^* = q^* = 4$ johtaa kaikilla kolmella sakkofunktiolla (AIC, HQ, BIC) kuitenkin laajempaan ARMA(2,1)-malliin.



Kuvio 4.1. USAn neljännesvuosittainen inflaatiosarja ajanjaksolta 1970I - 2006IV ($T = 148$). Alkuperäisen kausipuhdistetun kuluttajien hintaindeksin logaritmiset differenssit kerrottuna neljälläsadalla.

Sarjan alkupuoli näyttää hieman erilaiselta kuin loppupuoli, mutta stationaarisuus tuntuu kuitenkin melko kohtuulliselta oletukselta.



Kuvio 4.2. Kuvion 4.1 inflaatiosarjan otosautokorrelaatiofunktio (vas.) ja otosositaisautokorrelaatiofunktio (oik.) viipymillä $h = 0, \dots, 40$.

- Estimoitu autokorrelaatiofunktio vaimenee viipymän pituuden kasvaessa. Estimoidussa osittaisautokorrelaatiofunktiossa on varsin selvä katkos viipymällä 3. Siten AR(3)-malli voisi olla sopiva.
- Hannanin ja Rissasen menetelmä valinnoilla $m = 10$ ja $p^* = q^* = 4$ johtaa myös AR(3)-malliin kaikilla kolmella sakkofunktiolla (samoin, kuin valinnoilla $m = 10$ ja $p^* = q^* = 5, \dots, 9$).