ARMA-mallien rakentaminen

Parametrien tehokas estimointi

- Tarkastellaan *normaalista* ARMA(p,q)-prosessia

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

  ja oletetaan stationaarisuus-, käännettävyys- ja identifioituvuusehdot.

- Tarkastellaan *normaalista* ARMA(p,q)-prosessia

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

  ja oletetaan stationaarisuus-, käännettävyys- ja identifioituvuusehdot.
- Merkitään $\mathbf{y} = (y_1, ..., y_T)$, $\boldsymbol{\phi} = (\phi_1, ..., \phi_p)$, $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)$ ja $\boldsymbol{\beta} = (\boldsymbol{\phi}, \boldsymbol{\theta})$.

- Tarkastellaan *normaalista* ARMA(p,q)-prosessia

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

ja oletetaan stationaarisuus-, käännettävyys- ja identifioituvuusehdot.

- Merkitään $\mathbf{y} = (y_1, ..., y_T)$, $\boldsymbol{\phi} = (\phi_1, ..., \phi_p)$, $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)$ ja $\boldsymbol{\beta} = (\boldsymbol{\phi}, \boldsymbol{\theta})$.
- Koska $\mathbf{y} \sim \text{N}(0, \sigma^2 \boldsymbol{\Sigma})$, jossa $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\beta})$, on sv:n $\mathbf{y}$ tiheysfunktio

$$f_{\mathbf{y}}(\mathbf{y}) = (2\pi)^{-T/2} \det\left(\sigma^2 \boldsymbol{\Sigma}\right)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2} \mathbf{y}' \boldsymbol{\Sigma}^{-1} \mathbf{y}\right\}.$$

- Tarkastellaan *normaalista* ARMA(p,q)-prosessia

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

ja oletetaan stationaarisuus-, käännettävyys- ja identifioituvuusehdot.

- Merkitään $\mathbf{y} = (y_1, ..., y_T)$, $\boldsymbol{\phi} = (\phi_1, ..., \phi_p)$, $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)$ ja $\boldsymbol{\beta} = (\boldsymbol{\phi}, \boldsymbol{\theta})$.
- Koska $\mathbf{y} \sim \mathrm{N}(0, \sigma^2 \boldsymbol{\Sigma})$, jossa $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\beta})$, on sv:n $\mathbf{y}$ tiheysfunktio

$$f_{\mathbf{y}}(\mathbf{y}) = (2\pi)^{-T/2} \det\left(\sigma^2 \boldsymbol{\Sigma}\right)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2} \mathbf{y}' \boldsymbol{\Sigma}^{-1} \mathbf{y}\right\}.$$

- Tunnetusti, $\det(\sigma^2 \boldsymbol{\Sigma}) = \sigma^{2T} \det(\boldsymbol{\Sigma})$, joten log-uskottavuusfunktio on

$$l\left(\boldsymbol{\beta}, \sigma^2\right) = -\frac{T}{2} \log \sigma^2 - \frac{1}{2} \log\det\left(\boldsymbol{\Sigma}(\boldsymbol{\beta})\right) - \frac{1}{2\sigma^2} \mathbf{y}' \boldsymbol{\Sigma}(\boldsymbol{\beta})^{-1} \mathbf{y}.$$

- Log-uskottavuus funktiossa

$$l\left(\boldsymbol{\beta},\sigma^2\right)=-\frac{T}{2}\log\sigma^2-\frac{1}{2}\log\det\left(\boldsymbol{\Sigma}\left(\boldsymbol{\beta}\right)\right)-\frac{1}{2\sigma^2}\mathbf{y}'\boldsymbol{\Sigma}\left(\boldsymbol{\beta}\right)^{-1}\mathbf{y}.$$

  $\boldsymbol{\Sigma}\left(\boldsymbol{\beta}\right)$ on varsin monimutkainen funktio parametrista $\boldsymbol{\beta}$, joten log-uskottavuusfunktiota voida esittää yksinkertaisessa muodossa.

- Käytännössä uskottavuusfunktio joudutaan maksimoimaan numeerisia menetelmiä käyttäen.

- Tämä vaatii käänteismatriisin $\boldsymbol{\Sigma}\left(\boldsymbol{\beta}\right)^{-1}$ ja determinantin $\det\left(\boldsymbol{\Sigma}\left(\boldsymbol{\beta}\right)\right)$ laskemisen, kun parametrin $\boldsymbol{\beta}$ arvo on annettu. Annetulla $\boldsymbol{\beta}$:n ja σ^2:n arvoilla $\mathrm{Cov}\left(\mathbf{y}\right)=\sigma^2\boldsymbol{\Sigma}\left(\boldsymbol{\beta}\right)$ ja siten $\boldsymbol{\Sigma}\left(\boldsymbol{\beta}\right)$ voidaan laskea kuten jaksossa 3.3 kuvattiin.

- ARMA(p,q)-prosessin autokovartianssifunktio toteuttaa

$$\gamma_h = \begin{cases} \phi_1\gamma_{h-1} + \cdots + \phi_p\gamma_{h-p} + \sigma^2\sum_{j=h}^{q}\theta_j\psi_{j-h}, \\ \qquad\qquad 0 \le h < \max\{p, q+1\} \\ \phi_1\gamma_{h-1} + \cdots + \phi_p\gamma_{h-p}, \quad h \ge \max\{p, q+1\}, \end{cases}$$

jossa $\theta_0 = 1$ ja summaus $\sum_{j=h}^{q}$ tulkitaan nollaksi, kun $h > q$.

- Koska kertoimet ψ_j voidaan lausua parametrien $\phi_1, \ldots, \phi_p$ ja $\theta_1, \ldots, \theta_q$ funktiona, voidaan autokovarianssit γ_h, $h \ge 0$, ratkaista näistä yhtälöistä parametrien funktiona differenssiyhtälöiden ratkaisemisessa käytettävien menetelmien avulla.

- Miten lasketaan $\boldsymbol{\Sigma}(\boldsymbol{\beta})^{-1}$ ja $\det(\boldsymbol{\Sigma}(\boldsymbol{\beta}))$ log-uskottavuusfunktiossa

$$l\left(\boldsymbol{\beta}, \sigma^2\right) = -\frac{T}{2}\log\sigma^2 - \frac{1}{2}\log\det\left(\boldsymbol{\Sigma}\left(\boldsymbol{\beta}\right)\right) - \frac{1}{2\sigma^2}\mathbf{y}'\boldsymbol{\Sigma}\left(\boldsymbol{\beta}\right)^{-1}\mathbf{y} \text{ ?}$$

- Cholesky-hajotelma:

$$\boldsymbol{\Sigma} = \mathbf{C}\mathbf{C}',$$

jossa $\mathbf{C} = [c_{ij}]$ $(T \times T)$ on alakolmiomatriisi eli $c_{ij} = 0$, $i < j$, ja lisäksi $c_{ii} > 0$, $i = 1, ..., T$.

- Miten lasketaan $\mathbf{\Sigma}(\boldsymbol{\beta})^{-1}$ ja $\det(\mathbf{\Sigma}(\boldsymbol{\beta}))$ log-uskottavuusfunktiossa

$$l(\boldsymbol{\beta}, \sigma^2) = -\frac{T}{2}\log\sigma^2 - \frac{1}{2}\log\det(\mathbf{\Sigma}(\boldsymbol{\beta})) - \frac{1}{2\sigma^2}\mathbf{y}'\mathbf{\Sigma}(\boldsymbol{\beta})^{-1}\mathbf{y} \text{ ?}$$

- Cholesky-hajotelma:

$$\mathbf{\Sigma} = \mathbf{C}\mathbf{C}',$$

  jossa $\mathbf{C} = [c_{ij}]$ $(T \times T)$ on alakolmiomatriisi eli $c_{ij} = 0$, $i < j$, ja lisäksi $c_{ii} > 0$, $i = 1, ..., T$.

- Kun $\boldsymbol{\beta}$:n arvo on annettu, voidaan $\mathbf{C}$, $\det(\mathbf{C}) = \det(\mathbf{C}')$ ja $\mathbf{C}^{-1}$ ja muodostaa nykyisin helposti ($\mathbf{C}$:n alakolmiorakenne!).

- Miten lasketaan $\mathbf{\Sigma}\left(\boldsymbol{\beta}\right)^{-1}$ ja $\det\left(\mathbf{\Sigma}\left(\boldsymbol{\beta}\right)\right)$ log-uskottavuusfunktiossa

$$l\left(\boldsymbol{\beta},\sigma^2\right)=-\frac{T}{2}\log\sigma^2-\frac{1}{2}\log\det\left(\mathbf{\Sigma}\left(\boldsymbol{\beta}\right)\right)-\frac{1}{2\sigma^2}\mathbf{y}'\mathbf{\Sigma}\left(\boldsymbol{\beta}\right)^{-1}\mathbf{y}\ ?$$

- Cholesky-hajotelma:

$$\mathbf{\Sigma}=\mathbf{CC}',$$

jossa $\mathbf{C}=\left[c_{ij}\right]$ $\left(T\times T\right)$ on alakolmiomatriisi eli $c_{ij}=0$, $i<j$, ja lisäksi $c_{ii}>0$, $i=1,...,T$.

- Kun $\boldsymbol{\beta}$:n arvo on annettu, voidaan $\mathbf{C}$, $\det\left(\mathbf{C}\right)=\det\left(\mathbf{C}'\right)$ ja $\mathbf{C}^{-1}$ ja muodostaa nykyisin helposti ($\mathbf{C}$:n alakolmiorakenne!).
- Siis,

$$l\left(\boldsymbol{\beta},\sigma^2\right)=-\frac{T}{2}\log\sigma^2-\sum_{t=1}^{T}\log c_{tt}\left(\boldsymbol{\beta}\right)-\frac{1}{2\sigma^2}\mathbf{y}'\mathbf{C}\left(\boldsymbol{\beta}\right)'^{-1}\mathbf{C}\left(\boldsymbol{\beta}\right)^{-1}\mathbf{y}.$$

- Merkitään $\mathrm{S}(\boldsymbol{\beta}) = \mathbf{y}'\mathbf{C}(\boldsymbol{\beta})'^{-1}\mathbf{C}(\boldsymbol{\beta})^{-1}\mathbf{y}$, jolloin

$$l\left(\boldsymbol{\beta}, \sigma^2\right) = -\frac{T}{2}\log\sigma^2 - \sum_{t=1}^{T}\log c_{tt}(\boldsymbol{\beta}) - \frac{1}{2\sigma^2}\mathrm{S}(\boldsymbol{\beta}).$$

- $l\left(\boldsymbol{\beta}, \sigma^2\right)$ maksimoituu annetulla $\boldsymbol{\beta}$:n arvolla, kun $\sigma^2 = T^{-1}\mathrm{S}(\boldsymbol{\beta})$, joten SU-estimaateille $\hat{\boldsymbol{\beta}}$ ja $\hat{\sigma}^2$ pätee

$$\hat{\sigma}^2 = T^{-1}\mathrm{S}(\hat{\boldsymbol{\beta}}).$$

- Lisäksi, $\boldsymbol{\beta}$:n SUE $\hat{\boldsymbol{\beta}}$ saadaan maksimoimalla profiiliuskottavuusfunktio

$$l_{\mathrm{p}}(\boldsymbol{\beta}) = -\frac{T}{2}\log\mathrm{S}(\boldsymbol{\beta}) - \sum_{t=1}^{T}\log c_{tt}(\boldsymbol{\beta}).$$

- Profiiliuskottavuusfunktion $l_{\mathrm{p}}(\boldsymbol{\beta})$ maksimointi suoritetaan lähtemällä jostain alkuarvosta $\hat{\boldsymbol{\beta}}^{(0)}$ ja edeten iteratiivisesti arvoihin $\hat{\boldsymbol{\beta}}^{(1)}$, $\hat{\boldsymbol{\beta}}^{(2)}$, ..., jotka (toivottavasti) konvergoivat SU-estimaattiin $\hat{\boldsymbol{\beta}}$.

- Yleinen iteraatiokaava on tyyppiä

$$\hat{\boldsymbol{\beta}}^{(i+1)} = \hat{\boldsymbol{\beta}}^{(i)} + a^{(i)}\mathbf{d}^{(i)}, \quad i = 0, 1, 2, \ldots,$$

$\mathbf{d}^{(i)}$ on ns. suuntavektori ja $a^{(i)}$ (skalaari) on ns. askelpituus, jonka algoritmi valitsee siten, että ehto $l_{\mathrm{p}}(\hat{\boldsymbol{\beta}}^{(i+1)}) \geq l_{\mathrm{p}}(\hat{\boldsymbol{\beta}}^{(i)})$ toteutuu.

- Profiiliuskottavuusfunktion $l_{\mathrm{p}}(\boldsymbol{\beta})$ maksimointi suoritetaan lähtemällä jostain alkuarvosta $\hat{\boldsymbol{\beta}}^{(0)}$ ja edeten iteratiivisesti arvoihin $\hat{\boldsymbol{\beta}}^{(1)}$, $\hat{\boldsymbol{\beta}}^{(2)}$, ..., jotka (toivottavasti) konvergoivat SU-estimaattiin $\hat{\boldsymbol{\beta}}$.
- Yleinen iteraatiokaava on tyyppiä

$$\hat{\boldsymbol{\beta}}^{(i+1)} = \hat{\boldsymbol{\beta}}^{(i)} + a^{(i)}\mathbf{d}^{(i)}, \quad i = 0, 1, 2, ...,$$

$\mathbf{d}^{(i)}$ on ns. suuntavektori ja $a^{(i)}$ (skalaari) on ns. askelpituus, jonka algoritmi valitsee siten, että ehto $l_{\mathrm{p}}(\hat{\boldsymbol{\beta}}^{(i+1)}) \geq l_{\mathrm{p}}(\hat{\boldsymbol{\beta}}^{(i)})$ toteutuu.
- Esimerkiksi tunnetussa Newton-Raphson menetelmässä

$$\mathbf{d}^{(i)} = -\left(\partial^2 l_{\mathrm{p}}(\hat{\boldsymbol{\beta}}^{(i)})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'\right)^{-1}\partial l_{\mathrm{p}}(\hat{\boldsymbol{\beta}}^{(i)})/\partial\boldsymbol{\beta}.$$

- Stationaarisuus-, käännettävyys- ja identifioituvuusehtojen voimassa ollessa SUE:t $\hat{\boldsymbol{\beta}}$ ja $\hat{\sigma}^2$ ovat tarkentuvia ja asymptoottisesti normaalisia.
- Erityisesti pätee

$$\hat{\boldsymbol{\beta}} \underset{as}{\sim} \mathrm{N}\left(\boldsymbol{\beta}, \mathbf{V}\left(\boldsymbol{\beta}\right)^{-1}\right), \qquad (*)$$

  jossa $\mathbf{V}\left(\boldsymbol{\beta}\right) = \mathrm{E}\left[-\,\partial^2 l(\boldsymbol{\beta}, \sigma^2)/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'\right]$ (σ^2 supistuu odotusarvon ottamisen myötä pois).
- Nämä tulokset pätevät myös ilman normaalisuusoletusta, joskaan estimaattorit eivät enää ole (asymptoottisesti) tehokkaita.
- Toisaalta stationaarisuus-, käännettävyys- ja identifioituvuusehtodot vaaditaan tulokseen $(*)$, joka voi tuottaa kehnon approksimaation myös silloin, kun nämä ehdot ovat "lähellä rikkoontua".

$$\hat{\beta} \underset{as}{\sim} \mathrm{N}\left(\beta, \mathbf{V}(\beta)^{-1}\right), \quad (*)$$

- AR(p): $\mathbf{V}(\beta) = \mathbf{V}(\phi) = T\sigma^{-2}\mathbf{\Gamma}_p$, jossa $\mathbf{\Gamma}_p = \left[\gamma_{i-j}\right]_{i,j=1,..,p}$ eli sama kuin YW - ja PNS-estimaattorilla.
- MA(q): $\mathbf{V}(\beta) = \mathbf{V}(\theta) = T\sigma^{-2}\mathbf{\Gamma}_q^*$, jossa $\mathbf{\Gamma}_q^* = \left[\gamma_{i-j}^*\right]_{i,j=1,...,q}$ ja γ_h^* on AR(q)-prosessin $\theta(B)\, y_t^* = \varepsilon_t$ autokovarianssifunktio ($\varepsilon_t \sim \mathrm{nid}\left(0, \sigma^2\right)$).
- ARMA(p,q): $\mathbf{V}(\beta) = T\sigma^{-2}\mathrm{Cov}(\mathbf{x})$, jossa

  $\mathbf{x} = \left(y_1^+, ..., y_p^+, y_1^*, ..., y_q^*\right)$ ja y_t^+ AR(p)-prosessi $\phi(B)\, y_t^+ = \varepsilon_t$.

$$\hat{\beta} \underset{as}{\sim} \mathrm{N}\left(\beta, \mathbf{V}(\beta)^{-1}\right) \qquad (*)$$

- Havaittuun informaatiomatriisiin $\hat{\mathbf{V}}(\hat{\beta}, \hat{\sigma}^2) = \left[-\partial^2 l(\hat{\beta}, \hat{\sigma}^2))/\partial\beta\partial\beta'\right]$ perustuvia Waldin testejä voidaan muodostaa tavanomaiseen tapaan (tässä $\hat{\sigma}^2$ ei supistu pois).
- Uskottavuusosamäärätestejä voidaan myös muodostaa tavanomaiseen tapaan log-uskottavuusfunktiota käyttäen.
- Nollahypoteesin voimassa ollessa identifioituvuusehdon täytyy kuitenkin päteä, joten esim. H_0: $\phi_p = \theta_q = 0$ ei ole mahdollinen.

- $\hat{\boldsymbol{\beta}} \underset{as}{\sim} \mathrm{N}\left(\boldsymbol{\beta}, \mathbf{V}(\boldsymbol{\beta})^{-1}\right), \quad \hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2) = \left[-\partial^2 l(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2))/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'\right]$

- Esimerkiksi, jos $\mathbf{V}(\boldsymbol{\beta})^{-1} = \left[v^{ij}\right]$ ja $\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2)^{-1} = \left[\hat{v}^{ij}\right]$, niin $\hat{\beta}_i \underset{as}{\sim} \mathrm{N}\left(\beta_i, v^{ii}\right)$ ja nollahypoteesia $\beta_i = 0$, hylätään 5%:n merkitsevyystasolla, jos

$$\left|\frac{\hat{\beta}_i}{\sqrt{\hat{v}^{ii}}}\right| \geq 1.96 \quad \Leftrightarrow \quad \left|\hat{\beta}_i\right| \geq 1.96\sqrt{\hat{v}^{ii}}$$

- Vastaavasti, approksimatiivinen 95%:n luottamusväli β_i:lle on

$$\left(\hat{\beta}_i - 1.96 \cdot \text{s.e.}\left(\hat{\beta}_i\right), \hat{\beta}_i + 1.96 \cdot \text{s.e.}\left(\hat{\beta}_i\right)\right), \quad \text{s.e.}\left(\hat{\beta}_i\right) = \sqrt{\hat{v}^{ii}}$$

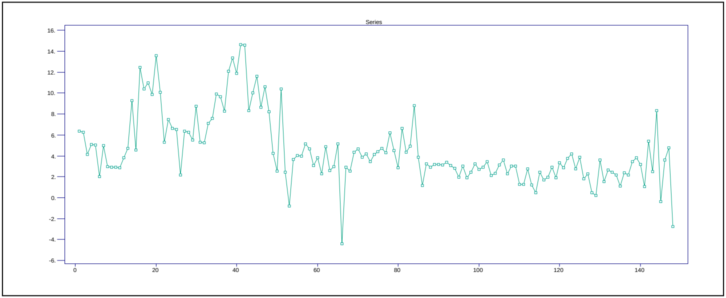
- Mallinvalintakriteerejä voidaan käyttää myös SU-estimaattorien kanssa valittaessa mallin asteita p ja q.
- Kriteerifunktio voidaan perustaa profiiliuskottavuusfunktioon

$$l_{\mathrm{p}}(\boldsymbol{\beta}) = -\frac{T}{2}\log \mathrm{S}(\boldsymbol{\beta}) - \sum_{t=1}^{T} \log c_{tt}(\boldsymbol{\beta}).$$

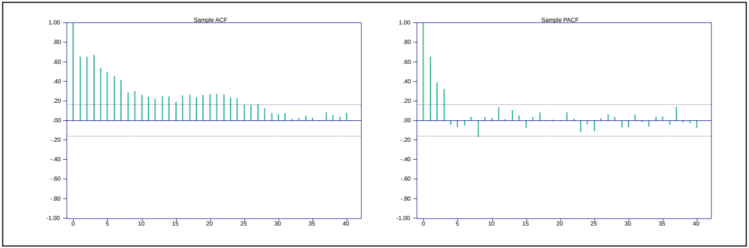
- Korvaamalla $\boldsymbol{\beta}$ SUE:lla $\hat{\boldsymbol{\beta}}$, kertomalla $-2/T$:llä ja lisäämällä sakkotermi $-\log T$ sekä asteet estimaattoreihin saadaan kriteerifunktio

$$C_{SU}(p,q) = \log \hat{\sigma}^2_{p,q} + \frac{2}{T}\sum_{t=1}^{T} \log c_{tt}(\hat{\boldsymbol{\beta}}_{p,q}) + (p+q)\, g(T)/T,$$

jossa $0 \le p \le p^*, 0 \le q \le q^*$ ja sakkofunktio $g(\cdot)$ on kuten Hannanin ja Rissasen menetelmässä. Pienet kriteerifunktion arvot liittyvät jälleen "hyviin" malleihin.



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



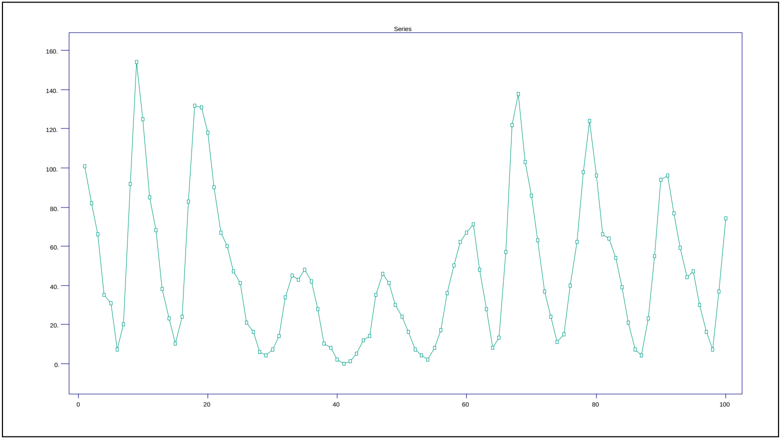
Kuvio 4.2. Kuvion 4.1 inflaatiosarjan otosautokorrelaatiofunktio (vas.) ja otososittaisautokorrelaatiofunktio (oik.) viipymillä $h = 0, ..., 40$.

- 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 $p^* = q^* = 5, ..., 9$).

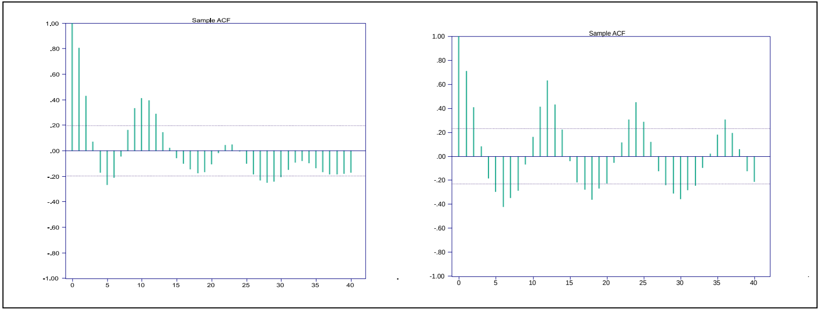
- Hannanin ja Rissasen menetelmä $\Rightarrow$ AR(3)-malli Kuvion 4.1 inflaatiosarjalle. SU-estimointi $(\bar{y}_t = y_t - 4.56)$:

$$\bar{y}_t \quad = \quad \underset{(0.080)}{0.275}\,\bar{y}_{t-1} + \underset{(0.081)}{0.263}\,\bar{y}_{t-2} + \underset{(0.080)}{0.321}\,\bar{y}_{t-3} + \hat{\varepsilon}_t, \quad \hat{\sigma}^2 = 4.637,$$

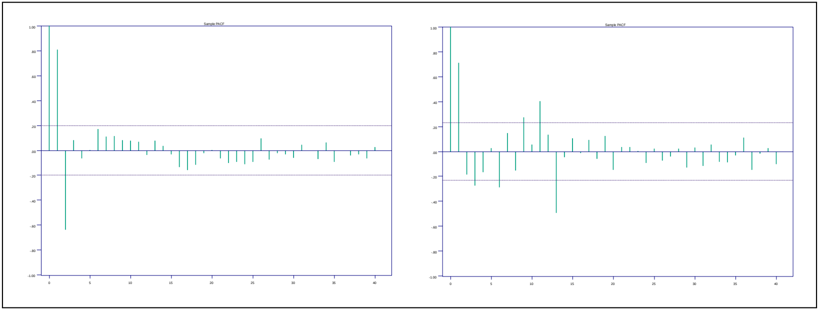
- Kaikki estimaatit ovat selvästi yli kaksi kertaa yli likimääräisen keskivirheensä, mikä viittaa siihen, ettei mallia voida supistaa.
- Lisäksi, ARMA(4,0), ARMA(3,1) ja ARMA(4,1) "huonompia" kuin AR(3) kriteerifunktion $C_{SU}(p,q)$ perusteella kaikilla sakkofunktiolla AIC, HQ ja BIC.



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 $h = 0, ..., 40$.



Kuvio 3.2. Kuvion 1.4 auringonpilkkusarjan otosositaisautokorrelaatio-funktio (vas.) ja Kuvion 1.2 onnettomuussarjan otosositaisautokorrelaatio-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.

ARMA-mallien rakentaminen

- Otosautokorrelaatio- ja otosositaisautokorrelaatiofunktiot $\Rightarrow$ AR(2)-malli.
- Hannanin ja Rissasen menetelmä $\Rightarrow$ ARMA(2,1)-malliksi, jonka estimointi tuotti

$$\bar{y}_t = \underset{(0.112)}{1.225}\,\bar{y}_{t-1} - \underset{(0.108)}{0.561}\,\bar{y}_{t-2} + \hat{\varepsilon}_t + \underset{(0.132)}{0.385}\hat{\varepsilon}_{t-1}, \quad \hat{\sigma}^2 = 213.96,$$

  Tulos osoittaa, että MA-parametrin estimaatti on selvästi "merkitsevä".
- Yleisemmin tämä esimerkki osoittaa, että autokorrelaatio- ja osittaisautokorrelaatiofunktioiden avulla ei aina ole helppo valita ARMA-mallin asteita.

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

- Rakennettaessa ARMA-mallia tavoitteena on, että mallin (tavalla tai toisella) valitut asteet p ja q on valittu siinä mielessä oikein, että virhetermi ε_t toteuttaa oletuksen $\varepsilon_t \sim \mathrm{iid}\left(0, \sigma^2\right)$ tai mielellään jopa $\varepsilon_t \sim \mathrm{nid}\left(0, \sigma^2\right)$, mutta p ja q eivät ole ainakaan samanaikaisesti suurempia kuin on tarpeen (säästäväisyysperiaate). Myös SU-estimointi olettaa tämän.
- Vaikka asteet olisikin valittu oikein, voivat virheet olla vain autokorreloimattomia, mutta eivät riippumattomia eivätkä nomaalisti jakautuneita, mitä on hyvä tutkia.

ARMA-mallien rakentaminen

Mallin riittävyyden tarkistaminen

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

- Valitun ARMA(p,q)-mallin riittävyyttä tutkittaessa kannattaa kiinnittää huomiota erityisesti "naapurimalleihin" eli ARMA(p+1,q)- ja ARMA(p,q+1)-malliin, joiden parametrit voidaan estimoida SU-menetelmällä ja tutkia poikkeaako lisätty AR-parametri tai MA-parametri nollasta.
- Laajempia vaihtoehtojakin voi tietysti kokeilla.
- Lisäksi on luontevaa tutkia ovatko virheiden ε_t empiiriset vastineet oletuksen $\varepsilon_t \sim \text{iid}(0, \sigma^2)$ tai $\varepsilon_t \sim \text{nid}(0, \sigma^2)$ mukaisia.

- Virheiden ε_t empiiriset vastineet eli *residuaalit* $\hat{\varepsilon}_t$ $(t = 1, ..., T)$ määritellään seuraavasti.
- Koska oikean mallin tapauksessa

$$\mathrm{E}(\mathbf{y}) = 0 \quad \text{ja} \quad \mathrm{Cov}(\mathbf{y}) = \sigma^2 \mathbf{\Sigma} = \sigma^2 \mathbf{C}\mathbf{C}', \ \text{niin}$$

$$\begin{aligned} \mathrm{Cov}\left(\mathbf{C}^{-1}\mathbf{y}\right) &= \mathbf{C}^{-1}\mathrm{Cov}(\mathbf{y})\,\mathbf{C}^{-1'} \\ &= \sigma^2 \mathbf{C}^{-1}\mathbf{C}\mathbf{C}'\mathbf{C}'^{-1} \\ &= \sigma^2 I_T. \end{aligned}$$

- Siis, luonteva määritelmä on (mahdollisesti $y_t \to y_t - \bar{y}$)

$$\hat{\boldsymbol{\varepsilon}} = [\hat{\varepsilon}_1 \ \cdots \ \hat{\varepsilon}_T]' = \mathbf{C}(\hat{\boldsymbol{\beta}})^{-1}\mathbf{y}.$$

- Käytännössä voidaan käyttää myös skaalattuja residuaaleja $\hat{\varepsilon}_t / \hat{\sigma}$, joiden tulisi oikein spesifioidun mallin tapauksessa olla likimain riippumattomia odotusarvona nolla ja varianssina yksi.

- Residuaalien ominaisuuksien tutkiminen on syytä aloittaa piirtämällä residuaalien aikasarja.
- Oikean mallin tapauksessa, residuaaleissa ei saisi näkyä systemaattista vaihtelua "tasossa" eikä varianssissa (heteroskedastisuutta) eikä poikkeavia havaintoja, ajanjaksoja tms.
- Seuraavaksi voidaan tutkia residuaalista laskettuja autokorrelaatioita $\hat{r}_h$.
- Kriittiset rajat $\pm 1.96/\sqrt{T}$ eivät oikeita edes asymptoottisesti, mutta voivat silti toimia "karkena mittakaavana".
- Aikaisempi Ljungin ja Boxin testisuure vaatii modifikaation:

$$Q_{LB} = T\left(T+2\right)\sum_{h=1}^{H}\hat{r}_h^2/\left(T-h\right)\underset{as}{\sim}\chi^2_{H-p-q},$$

kun H on "suuri" (käytännössä ehkä 10-30 riippuen T:n koosta).

- Virheissä mahdollisesti esiintyvien epälineaaristen riippuvuuksien (ja ei-normaalisuuden) paljastamisessa voidaan käyttää (rajoitetussa mielessä) residuaalien neliöiden autokorrelaatioita ja McLeodin ja Lin testisuuretta.
- Jos teoreettisille virheille pätee $\varepsilon_t \sim \mathrm{nid}(0, \sigma^2)$, ovat neliöityjen residuaalien $\hat{\varepsilon}_t^2$ $(t = 1, \ldots, T)$ otosautokorrelaatiot $\mathrm{r}_h(\hat{\varepsilon}_t^2)$ likimain riippumattomia ja $\mathrm{N}\left(0, 1/T\right)$–jakautuneita.
- Siis, kriittiset rajat $\pm 1.96/\sqrt{T}$ toimivat ja samoin McLeodin ja Lin testisuure

$$T\left(T+2\right)\sum_{h=1}^{H}\mathrm{r}_h^2(\hat{\varepsilon}_t^2)/\left(T-h\right) \underset{as}{\sim} \chi_H^2$$

  ilman, että viipymän H tarvitsee olla "suuri".
- McLeod ja Li ovat osoittaneet nämä tulokset päteviksi myös oletuksilla $\varepsilon_t \sim \mathrm{iid}(0, \sigma^2)$ ja $\mathrm{E}(\varepsilon_t^8) < \infty$.

- Jos virheet ovat vain autokorreloimattomia, mutta eivät riippumattomia, ei tulos

$$\hat{\beta} \underset{as}{\sim} \mathrm{N}\left(\beta, \mathbf{V}(\beta)^{-1}\right) \qquad (*)$$

  (yleensä) päde.

- $\hat{\beta}$:n tarkentuvuus ja asymptoottinen normaalisuus voidaan kuitenkin osoittaa yleisin oletuksin, mutta kovarianssimatriisi ei ole (yleensä) kuten $(*)$:ssä (ks. jakso 6.3).

- Tällöin tulokseen $(*)$ perustuvat keskivirheet ja testit eivät myöskään ole (yleensä) päteviä.