Fræðilegur bakgrunnur fyrir kafla 2 í Tsay

- Grundvallarskilningsatriði við spár með þekktum parametrum er:

$$\left[\begin{array}{c}\mathbf{X}_1\\ \mathbf{X}_2\end{array}\right] \sim N(\left[\begin{array}{c}\mu_1\\ \mu_2\end{array}\right], \left[\begin{array}{cc}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{12}' & \Sigma_{22}\end{array}\right]) \tag{1}$$

þá er:

$$E(\mathbf{X}_1|\mathbf{X}_2 = x_2) = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) \tag{2}$$

og

$$V(\mathbf{X}_1|\mathbf{X}_2 = x_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'$$

- Hér má hugsa sér að $\mathbf{X}_2 = (X_1, X_2, \ldots, X_T)'$ tákni fortíð, þ.e. T mælingar og $\mathbf{X}_1 = (X_{T+1}, X_{T+2}, \ldots, X_{T+h})'$ séu fyrstu h mælingarnar í framtíðinni.

- Ef allir parametrar, $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \Sigma_{11}, \Sigma_{12}, \Sigma_{22}$ eru þekktir er einungis tæknilegt atriði að reikna út bestum (least-squares) spá. (Þetta eru fylkjaframsetning af formúlunum á bls. 342 í Spanos)

- Þegar talað er um spálíkön er átt við í hvaða mæli fortíð nýtist til að spá framtíð. Covaríans-fylkin lýsa tengslum fortíðar og framtíðar. Þess vegna er oft gengið út frá að $\mu = 0$

ARMA líkön

- AR(p) líkan

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t$$
$$E(\varepsilon_t) = 0 \quad V(\varepsilon_t) = \sigma^2 \quad E(\varepsilon_t \varepsilon_s) = 0 \text{ ef } t \neq s$$
$$\varepsilon_t \quad \text{er white noise}$$

- MA(q) líkan

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots \varepsilon_{t-q}$$

- ARMA(p,q)

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots \varepsilon_{t-q}$$

- Oft er heppilegt að setja þetta þetta fram með aðstoð „backward"-virkjans B (tafa-virkinn L)

$$BX_t = X_{t-1} \quad LX_t = X_{t-1} \quad \text{þ.e. ARMA(p,q) er}$$
$$\phi(B)X_t = \theta(B)\varepsilon_t \quad \text{þar sem } \phi \text{ og } \theta \text{ eru margliður}$$
$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \text{ og } \theta(z) = 1 - \theta_1 z - \cdots \theta_q z^q$$

- Athugið að stundum er hægt að flytja svona margliður yfir jafnaðarmerkið:

$$X_t = \frac{\theta(B)}{\phi(B)} \varepsilon_t = \psi(B)\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

Stationarity/sístæðni

. Tímaraðagreining er athugun á tilraun þar sem útkoman er ferill. Til að geta alhæft eitthvað um eðli ferils út frá mælingum á takmörkuðum hluta hans er nauðsynlegt að ganga út frá einhvers konar stöðugleika ferilsins. Statinonarity er dæmi um stöðugleikahugtak. Röð er sögð veikt sístæð ef:

$E(X_t) = \mu$ ekki fall af t $\quad$, $V(X_t) = \sigma^2$

$E(X_t - \mu)(X_{t-k} - \mu) = \gamma_k$ einungis fall af k og ekki t.

$\gamma(k)$ er kallað auto-covariance function, ACF X_t

$\rho_k = \dfrac{\gamma_k}{\gamma_0}$ er kallað auto-correlation function, sjálffylgnifallX_t

- Veik sístæðni er það sem venjulega er unnið með. Einnig til sterk síðstæðni.
- Til að gagnasöfnun dugi til að meta parametra í líkani þarf X_t einnig að vera ergodic.
- Ergodicity gefur skilyrði sem duga t.d. til að hægt sé að meta meðal tal X_t. Þ.e. að

$$V(\frac{1}{T}\sum_{t=1}^{T} X_t) \to 0$$

- Þá er sagt að X_t sé ergodic in mean. Sjá nánar t.d. Spanos (1999) bls 424-426.
- Nægjanlegt skilyrði fyrir að X_t sé ergodic í variance og covariance er:

$$\sum_{-\infty}^{\infty} |\gamma_k| < \infty$$

Af hverju ARMA líkön

- *Wold's representation theorem*: Sístætt ferli X_t má skrifa sem:

  $$MA(\infty) + \text{ deterministískt ferli}$$

- Það er ekki praktískt að vinna með óendanlega marga parametra og því heppilegt að velja ARMA(p,q) í stað MA(∞).
- Mörg ferli eru greinilega ekki sístæð. Við segjum að ferli X_t sé „integrated" af gráðu d ef

  $$\Delta^d X_t \text{ er sístætt}$$

- Ferlið

  $$\Delta^d \phi(B) X_t = \theta(B) \varepsilon_t$$

  þar sem $\phi(z)$ og $\theta(z)$ eru margliður af gráðu p og q er kallað ARIMA(p,d,q)

Berið saman við jöfnu (1).

$$\mathbf{X} = (\underbrace{X_{T+h}, \ldots, X_{T+1}}_{\mathbf{X}_1=\text{framtíð}}, \underbrace{X_T, X_{T-1}, \ldots, X_2, X_1}_{\mathbf{X}_2=\text{fortíð}})'$$

$$\mathbf{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 \cdots & \cdots & \cdots & \gamma_{T+h} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{T+h-1} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & & \\ \gamma_{T+h} & \gamma_{T+h-1} & \cdots & \gamma_1 & \gamma_0 \end{bmatrix}$$

- Athugið ef X_t fylgir ARMA(p,q) þá er $\mathbf{\Sigma} = \mathbf{\Sigma}(\phi, \theta, \sigma)$, þ.e. að þetta covaríans-fylki er (flókið) fall af parametrum ARMA líkansins.
- Ef fylkið $\mathbf{\Sigma}$ er þekkt þá má nota jöfnu (2) til að reikna spá. Það er tæknilega erfitt því að það þarf að andhverfa stórum fylkjum. Þess vegna má nota aðferðir eins og Durbin-Levinson og innovation-algorithm til að reikna spá og spámörk. Sjá t.d. Brockwell and Davis (1991).

Aðferðafræði Box-Jenkins

Hugmyndin gengur út á að nota úrtaksfræði fyrir ARMA ferli og reyna að finna sparlega (parsimonious) ARIMA(p,d,q) (þ.e. p+q lítið) framsetningu sem gæti lýst gögnum vel. Þetta er eins konar data-mining leið. Gangurinn samkvæmt þeirra leiðbeiningum er eftirfarandi:

1. Identificaton: Vel heppilegt form, t.d. tek logaritma, næ burt deterministískum þáttum, nota úrtaks-ACF, úrtaks-PACF til að ákvarða p og q. Finn það d sem lágmarkar $V(\Delta^d x_t)$ til að ákvarða d.

2. Estimation: Met ϕ, θ og σ.

3. Diagnostics: Metnir afgangsliðir $\hat{\varepsilon}_t = \hat{\varepsilon}(\hat{\phi}, \hat{\theta})$ skoðaðir. Ef vel hefur tekist til ættu metnir afgangsliðir að hafa svipaða eiginleika og ε_t, þ.e. virðast white-noise. Hér eru teiknuð gröf, reiknuð sjálffylgniföll, nota CUSUM og CUSUMSQ o.s.frv. Ef grunur leikur á misbresti er farið aftur í skref 1.

4. Forecasting: Hér er metið líkan notað til að reikna spár.

Úrtaks ACF fyrir stationary ferli

- Gögnin fást sem mælingar $(x_1, \ldots, x_T)'$ á $(X_1, \ldots, X_T)'$

$$\hat{\gamma}_l = \frac{1}{T} \sum_{j=l+1}^{T} (X_j - \hat{\mu})(X_{j-l} - \hat{\mu})$$
$$\hat{\rho}_l = \hat{\gamma}_l / \hat{\gamma}_0$$

- Dreifingu $\hat{\rho}_l$ má nálga með:

$$\hat{\rho}_l \sim N(\rho_l, \frac{1}{T} \sum_{i=-\infty}^{\infty} (\rho_l^2 + \rho_{i+l}\rho_{i-l} - 4\rho_l\rho_i\rho_{i-l} + 2\rho_l^2\rho_i^2))$$

  liðum í þessari formúlu fækkar fyrir $\hat{\rho}_l$ ef $\rho_l = 0$ fyrir $l > m$.

- Tilsvarandi fyrir PACF er einfaldar, þ.e. ef $\phi_{kk} = 0$ þá er

$$\hat{\phi}_{kk} \sim N(0, \frac{1}{T})$$

Grunnhugtök fyrir margvíðar raðir

- Sístæð (veikt) $\mathbf{X}_t = (X_{1t}, \ldots, X_{kt})'$ k-víð tímaröð

$$E(\mathbf{X}_t) = \mu \quad \text{ekki fall af } t$$

- Kross-covaríans fall:

$$\mathbf{\Gamma}_l = E(\mathbf{X}_t - \mu)(\mathbf{X}_{t-l} - \mu)' \quad kxk \text{ fylki einungis fall af } l$$
$$\mathbf{\Gamma}_{-l} = \mathbf{\Gamma}_l'$$

- Sjálffylgni-fall, ρ_l er skilgreint á hliðstæðan hátt, hornalína $\mathbf{\Gamma}_0$ sköluð í 1.

Reglur Bartlett's fyrir úrtaks-kross-sjálffylgni

- Ef X_{1t} og X_{2t} er óháð sístæð normal-ferli (+ nokkrar tæknilegar forsendur) þá:

$$[\hat{\rho}_{12}]_l \sim N(0, \frac{1}{T} \sum_{j=-\infty}^{\infty} [\rho_1]_j [\rho_2]_j)$$

- Þ.e. að dreifing úrtaksfylgnistuðuls er fall að sönnu sjálffylgnigildum.

- Það er einnig til hliðstæð niðurstaða þegar X_{1t} og X_{2t} eru háð normalferli

- Kjarni málsins er að ef annað hvort X_{1t} eða X_{2t} er white-noise þá gildir (eins og segir í byrjendakennslubókum)

$$TV([\hat{\rho}_{12}]_l) \xrightarrow[T\to\infty]{} 1 \quad \text{ef } [\rho_{12}]_l = 0 \tag{3}$$

Nálgun Box-Jenkins á tvívíðri greiningu

- Hugmyndin er að önnur röðin sé input.
  1. Meta ARIMA líkan fyrir input röðina
  2. Afgangsliðirnir eiga að líkjast white-noise. Maður lítur á afgangsliðina sem hvíta útgáfu af input röðinni. Fyrsta skrefið er kalla prewhitening.
  3. Reikna kross-fylgnistuðla milli afgangsliða input raðar og output raðar. Nota niðurstöðurnar til að finna form (identify) á líkani sem skýrir output röð með töfðum gildum af sjálfri er (AR), MA-liðum og afgangsliði úr input líkani á ýmsum töfum.
  4. Framkvæmi diagnostic. Ef vel hefur tekist eiga afgangsliðaraðir að líkjast óháðum white-noise röðum og sýna lítil tengsl við mældu raðirnar.
  5. Nota líkanið við spár og/eða stýringu.
- Í kafla 2 í Tsay er þessi leið farin til að leita tengsla milli skammtímavaxta og langtímavaxta.

Nútímaaðferðafræði

- Það er hagræn gagnrýni að fyrsta skrefið, þ.e. prewhitening, hefur tilhneigingu til að sigta út merkilegar upplýsingar.
- Það er tölfræðilegur vandi að þegar reiknuð er fylgni ósístæðra ferla þá eru úrtaksdreifing flóknar.
- ECM=Error-Correction-Model tengir saman breytingar og level í röðum.
- ECM er því túlkanlegt og hagmælingamenn hafa lengi leitað (data-mining) að slíku formi.
- Granger-representation theorem segir að tilvist ECM sé jafngild því að fyrir hendi séu co-integrating sambönd.
- Það er eitt að skilja fyrirbærið, co-integration og ECM, annað að álykta um slíkt út frá mælingum.

Verkefni

1. Ef X_t er AR(p), mælingar eru $(X_1, \ldots, X_T)$, reiknið bestu spá fyrir X_{T+1}
2. Ef X_t er MA(1) og mælingar eru (X_1, X_2), reiknið bestu spá fyrir X_3.
3. Setjið AR(1) líkan fram sem MA(∞).
4. Finnið bestu spá fyrir MA(1).
5. Skrifið niður likelihood fall (normal) fyrir MA(1) líkan þar sem tvær mælingar eru fyrir hendi.
6. Skrifið niður likelihood fall (normal) fyrir AR(1) líkan þar sem T mælingar eru fyrir hendi.
7. Hermið 25 mælingar úr tveim óháðum AR(1) þar sem $\phi_1 = \phi_2 = 0.95$. Eru ferlin sístæð? Reiknið dreifingu fylgnistuðuls milli raðanna (notið reglu Bartlett's). Berið staðalfrávikið í þeirri dreifingu saman við staðalfrávik hendingar sem er jafndreifð á bilinu (-1,1). Túlkið þann samanburð.
8. Hermið tvo óháða random walk. Reiknið prófstærð úr venjulegri aðhvarfsgreiningu og athugið hvort kenningunni um enga fylgni sé hafnað. Endurtakið nokkrum sinnum.

Heimildir

Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*. Springer-Verlag.

Spanos, A. (1999). *Probability Theory and Statistical Inference*. Cambridge University Press.