

A nemstacionaritás kezelése: stacionarizálás, trend- és differenciastacioner idősorok, differenciázás, ARIMA-folyamat

Ferenci Tamás
tamas.ferenci@medstat.hu

Utoljára frissítve: 2023. május 12.

Tartalom

Tartalomjegyzék

1	Stacionarizálás	1
---	-----------------	---

1. Stacionarizálás

A stacionarizálás szükségessége

- Mi alapvetően stacioner idősorokat szeretnénk majd modellezni (például olyan ARMA-val akarunk idősort modellezni, ami stacioner)
- De: a legtöbb közgazdasági idősor *nem* stacioner!
- Mit csináljunk most?
- Olyan „visszacsinálható” (invertálható) transzformációt alkalmazunk, ami a nem-stacioner időorból stacionert csinál
- Azon elvégezzük a modellezést (és ha kell, a transzformáció inverzével visszatérünk az eredeti idősor nagyságrendjébe)
- Két módszert fogunk látni
- Ez nem univerzális: nem arról van szó, hogy valamelyiknek matematikai szükség-szerűség, hogy stacionarizálnia kell minden idősort, egész egyszerűen azért nézzük meg ezeket, mert a gyakorlatban sokszor beváltak

Determinisztikus trend szűrése

- Az első módszer a determinisztikus trend szűrése: az idősorra ráillesztünk egy analitikus trendet majd kivonjuk belőle, ezt jelenti a „szűrés”
- Például (lineáris trend szűrése):

$$Y_t = \alpha + \beta t + u_t$$

modell alapján megbecsüljük α és β értékét (OLS- vagy ML-elven), majd áttérünk a – reményeink szerint stacioner –

$$Y'_t = Y_t - (\hat{\alpha} + \hat{\beta}t)$$

idősorra

- Ilyen értelemben mondjuk, hogy kiszűrtünk belőle egy determinisztikus trendet
- (Lényegében az egyenes illesztése utáni reziduumokra tértünk át)
- Visszatérés: a trend hozzáadása

Determinisztikus trend szűrése

- Ha ezzel a transzformációval az idősor stacionarizálható (tehát Y'_t már stacioner), akkor azt mondjuk, hogy az eredeti Y_t idősor **trendstacioner folyamat** (TSP, trend stationary process) volt
- Természetesen nem muszáj egyszerű lineáris trendet szűrni, illeszthetünk kvadratikusan trendet ($\alpha + \beta t + \gamma t^2$), exponenciális trendet ($\alpha e^{\beta t}$), szezonaritást, bármit, a lényeg, hogy egy előre megadott determinisztikus függvényforma legyen
- (A dolog ugyanis azért fog működni, mert amit illesztünk, annak a paraméterei szigorúan exogének)
- Vegyük észre, hogy ez filozófiájában a korábban látott „determinisztikus trend” fogalmához illeszkedik: ha egy idősorban trend van, de az determinisztikus ($Y_t^{(D)} = \alpha t + u_t$), akkor épp ez a módszer fogja stacionarizálni

Differenciálás

- Ha az idősorban viszont sztochasztikus trend van ($Y_t^{(S)} = Y_{t-1}^{(S)} + \alpha + u_t$), akkor egy másik, de pofonegyszerű transzformációval stacionarizálhatjuk:

$$Y'_t = Y_t - Y_{t-1}$$

- Hiszen ha az idősor valóban az előbbi modell követi, akkor a fenti transzformáció eredménye

$$Y'_t = \alpha + u_t$$

lesz, ami feltevéseink szerint tényleg stacioner

- Ezt a transzformációt úgy hívjuk, hogy az idősor **differenciázása**, jele Δ , ez is egy *operátor*:

$$\Delta Y_t = Y_t - Y_{t-1} = Y_t - LY_t = (1 - L) Y_t,$$

tehát $\Delta = 1 - L$

Differenciázás

- Ha ezzel a transzformációval az idősor stacionarizálható (tehát Y'_t már stacioner), akkor azt mondjuk, hogy az eredeti Y_t idősor **differenciastacioner folyamat** (DSP, difference stationary process) volt
- Visszatérés: felkumulálás (kezdőértékre szükség lesz):

$$Y_t = Y_0 + \sum_{i=1}^t \Delta Y_i$$

- (Itt már ráismerhetünk, hogy a differenciázás igazából nem más, mint a *diszkrét deriválás*: ha diszkrét halmazon vagyunk, akkor a $\lim_{\Delta t \rightarrow 0}$ azt jelenti, hogy $\Delta t = 1$, azaz, hogy két egymást követő időpont különbségét nézzük, és ilyenkor persze le sem kell osztani Δt -vel)
- Emiatt azt is mondjuk, hogy a folyamat elsőrendben integrált, jelben $I(1)$

Differenciázás

- A differenciázás lineáris trendet tüntet el (ha az sztochasztikus értelmű)
- Mi van, ha kvadratikus trendet kell eltüntetnünk?
- Ugyanúgy, ahogy az $ax + b$ függvényt a deriválás teszi konstanssá, az $ax^2 + bx + c$ -t pedig a kétszeri deriválás, ilyenkor a *kétszeri differenciázás* (*másodrendű differenciázás*) lesz a megoldás:

$$\begin{aligned} \Delta(\Delta Y_t) &= \Delta(Y_t - Y_{t-1}) = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = \\ &= (Y_t - 2Y_{t-1} + Y_{t-2}) \end{aligned}$$

- A jele Δ^2
- Az előbbi eredmény nem meglepő, hiszen

$$\Delta^2 = (1 - L)^2 = 1 - 2L + L^2$$

- Ha egy folyamat így stacionarizálható, akkor azt mondjuk, hogy másodrendben integrált, jelben $I(2)$

Differenciálás

- Természetesen a dolog általánosítható: Δ^d a d -ed rendbeli (d -szeri) differenciálás, $\Delta^d = (1 - L)^d$
- Ha egy idősor nem stacioner, az első differenciázottja sem az, a második differenciázottja sem az, ..., de a d -szeri differenciázottja már igen (tehát d a legkisebb egész szám, hogy az annyiszoros differenciázott már stacioner), akkor azt mondjuk, hogy az idősor d -ed rendben integrált, jelben $I(d)$
- (Ennek megfelelően a stacioner idősor nulladrendben integrált, jelben $I(0)$)
- Ez a tipikusabb a közgazdasági gyakorlatban
- Olyannyira, hogy ennek ARMA-val való kombinációjára külön elnevezés van: ha egy d -ed rendben integrált idősor d -szeres differenciázottját modellezzük ARMA(p,q)-val, akkor azt is mondhatjuk, hogy az eredeti idősort **ARIMA**(p,d,q)-val modelleztük