Nemstacionaritás kezelése

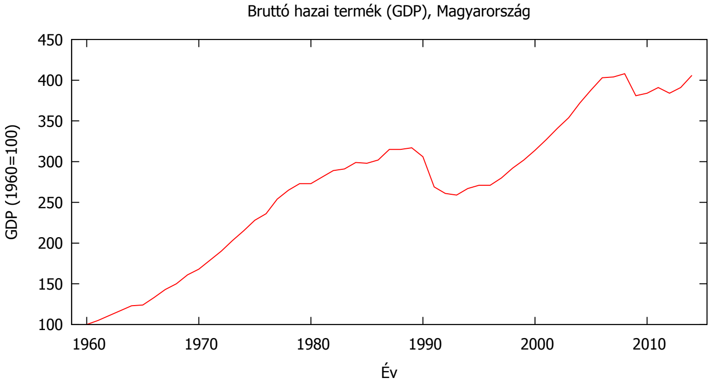
Ferenci Tamás
tamas.ferenci@medstat.hu

Tartalom

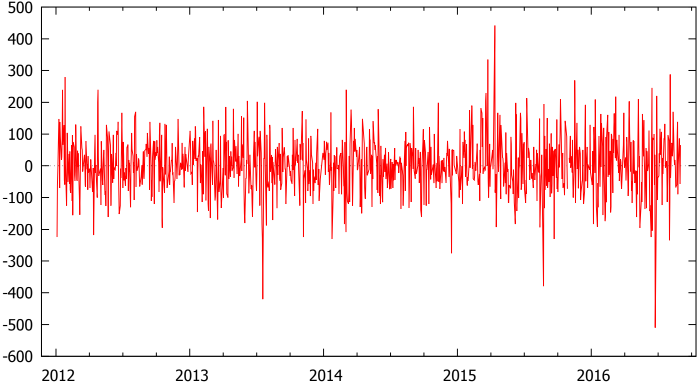
A tesztelés lehetőségei

- Egy módszerről már volt szó (grafikus eljárás), de ez igen szubjektív
- Most kiegészítjük két újjal, a második, a korrelogram szemrevételezése még mindig inkább csak heurisztikus...
- ...de a harmadik, a statisztikai próbák alkalmazása már objektív (noha ez nem azt jelenti, hogy tökéletes!)

Grafikus módszer



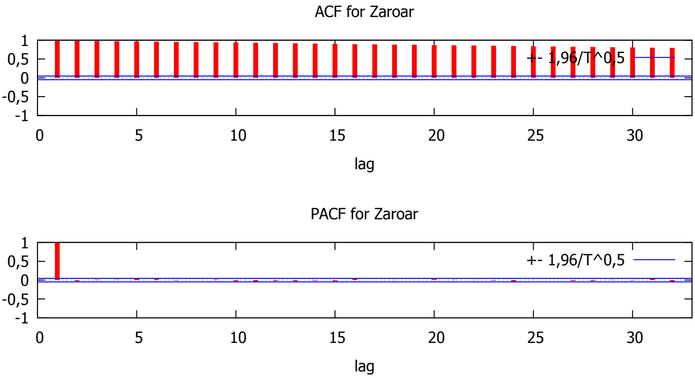
Grafikus módszer



A grafikus módszer határai

- A (gyenge) stacionaritás három feltételéből igazából csak kettő vizsgálható egyáltalán „ránézésre”
- Nem állandó várható érték, nem állandó szórás (akár csak átmenetileg is!)
- A mintavételi ingadozás figyelembevételére nincsen formális módszer, nehéz megítélni (különösen ha nem túl nagy a mintanagyság)
- Szubjektív

Korrelogram szemrevételezése



Korrelogram szemrevételezése – miért működik?

- Azt kell nézni, hogy az ACF nagyon nem lecsengő-e
- (Figyelem: a „nagyon nem lecsengő” nem azt jelenti, hogy nem igaz az, hogy nullába tart, hanem azt, hogy az 1-től is alig szakad el!)
- Miért van ez így?
- Intuitív indoklás, gondoljunk arra, ha trendje van az idősornak
- (A végigtolt ablaknak mindkét vége vagy az átlag alatt, vagy az átlag felett lesz az esetek nagy részében; ez csak lassan oldódik az ablak szélességének növekedtével)

Statisztikai tesztek

- Ez az igazán objektív módszer, kapunk egy *p*-értéket, ebben nincsen szubjektív tényező
- A gyakorlatban legelterjedtebb módszerek valójában inkább azt ellenőrzik, hogy van-e ún. egységgyök a folyamat (nem általában azt, hogy „nem stacioner"), ld. mindjárt
- Például Dickey-Fuller teszt (DF), kiterjesztett Dickey-Fuller teszt (ADF, augmented DF)
- Lásd kicsit később

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 idősorból 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 szükségszerűség, 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 –

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

  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 kvadratikus trendet ($\alpha + \beta t + \gamma t^2$), exponenciális trendet ($\alpha e^{\beta t}$), szezonalitást, 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

Differenciázás

- 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 $\lim_{\Delta t \to 0}$ 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:

$$\Delta\left(\Delta Y_t\right) = \Delta\left(Y_t - Y_{t-1}\right) = \left(Y_t - Y_{t-1}\right) - \left(Y_{t-1} - Y_{t-2}\right) = \\ = \left(Y_t - 2Y_{t-1} + Y_{t-2}\right)$$

- 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 $I(2)$

Differenciázás

- Természetesen a dolog általánosítható: Δ^d a d-ed rendbeli (d-szeri) differenciázás, $\Delta^d = (1 - L)^d$
- 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

Az egységgyök fogalma

- Láttuk, hogy egy ARMA-folyamat akkor stacioner, ha az AR-rész polinomjának, $\phi(L)$-nek az összes gyöke az egységkörön kívül van
- Pontosítsuk ezt az állítást:
  1. Ha az összes gyök az egységkörön kívül van, akkor a folyamat stacioner (eltolásinvariánsak a momentumok, egyedi impulzusok hatása lecsengő)
  2. Ha akár csak egyetlen gyök is az egységkörön belül van, akkor a folyamat **explozív** (momentumok elmennek a végtelenbe – esetleg oszcillálva –, az egyedi impulzusok hatása felerősödő)
  3. Ha egységkörön *belül* nincs gyök, de az egységkör*ön* van – egy vagy több – akkor ugyan nem stacioner, de egy furcsa helyzet áll elő
- (Gondoljunk mindezeket végig az $Y_t = \alpha + \phi_1 Y_{t-1} + u_t$ példáján!)
- Az említett 3. helyzet határeset: az egyedi impulzusok hatása sem nem lecsengő, sem nem felerősödő, a momentumok felemásan viselkednek (pl. variancia elszáll, de várható érték állandó)
- Ezt hívjuk **egységgyök-folyamatnak** (azt is mondjuk, hogy a folyamatban egységgyök van)

Az egységgyök és a differenciastacionaritás kapcsolata

- Ha $\phi(x)$-nek egy darab 1 értékű gyöke van (a többi nagyobb), akkor úgy is írható mint $\phi(x) = \widetilde{\phi}(x)(1-x)$, ahol $\widetilde{\phi}(x)$-nek már minden gyöke 1-nél nagyobb
- Igen ám, de ezzel a $\phi(L) Y_t = \alpha + \theta(L) u_t$ úgy is írható, mint

$$\widetilde{\phi}(L)(1-L) Y_t = \alpha + \theta(L) u_t,$$

  azaz

$$\widetilde{\phi}(L) \Delta Y_t = \alpha + \theta(L) u_t$$

- Vagyis ilyenkor az idősor differenciázottjára adtunk egy ARMA-modellt!
- Egész pontosan ARMA(p-1,q)-t, hiszen az AR-polinomja ($\widetilde{\phi}(L)$) eggyel kisebb fokszámú

Az egységgyök és a differenciastacionaritás kapcsolata

- Azaz: ha – egyszeres – egységgyök van egy ARMA folyamatban, az épp azt jelenti, hogy differenciastacioner, mégpedig $I(1)$ lesz, mert a differenciázottja stacioner ARMA lesz
- Ez egy fontos magyarázat arra, hogy miért találjuk azt, hogy a differenciázás sokszor segít: épp az egységgyököt tünteti el!
- Hasonlóan, ha az 1 d-szeres gyök, akkor a d-szer differenciázott folyamat lesz stacioner, tehát az eredeti folyamat $I(d)$ volt

Egységgyök-tesztelés: DF-teszt

- Tekintsünk először egy AR(1)-modellt: $Y_t = \alpha + \phi_1 Y_{t-1} + u_t$ a szokásos feltevésekkel
- Az egyértelmű, hogy $H_0 : \phi_1 = 1$, klasszikusan legtöbbször a $H_1 : \phi_1 < 1$ alternatívával szemben vizsgálódunk
- (Mert: az explozív idősorokat teljesen kizárjuk a vizsgálódásunk köréből)
- Rögtön érthetővé válik, amit arról mondtunk, hogy ez nem „stacionaritási teszt", hanem egységgyök teszt (bár ebben az esetben a kettő *majdnem* ugyanaz, az egyetlen különbség az explozivitás kizárása)
- A teszteléshez térjünk át a differenciákra:

$$\Delta Y_t = \alpha + \delta_1 Y_{t-1} + u_t,$$

  ahol $\delta_1 = \phi_1 - 1$
- Ennek megfelelően itt a tesztünk: $H_0 : \delta_1 = 0$ vs $H_1 : \delta_1 < 0$

Egységgyök-tesztelés: DF-teszt

- Egyszerűen eresszünk rá egy t-próbát?
- Nem jó ötlet, mert δ_1 t-hányadosának nem t-eloszlása lesz
    - Klasszikusan azzal indokoljuk a t-eloszlást, hogy ha nagy a mintánk, akkor ez (aszimptotikusan) eloszlási feltevésektől függetlenül teljesül, a centrális határeloszlás tétel miatt
    - Csakhogy itt a CLT nem fog érvényesülni, mert Y_{t-1} integrált idősor (gondoljunk bele, a varianciája minden határon túl nőni fog, ha a mintanagyság egyre nagyobb!)
- David Dickey és Wayne Fuller 1979-ben nagy számú szimulációval tisztázta, hogy – legalábbis aszimptotikusan – milyen eloszlása van akkor ennek, ha nem t, ezt hívjuk DF-eloszlásnak
- Ez alapján (vagy legalábbis a kitáblázott kritikus értékek alapján) már végezhető teszt: DF-teszt

Egységgyök-tesztelés: DF-teszt

- A gyakorlatban három módon szoktuk alkalmazni (más a DF-eloszlás mindegyikhez):
  1. $\alpha = 0$ (konstans és trend nélkül): sztochasztikusan sem lehet benne trend (nulla körül kell ingadozzon a differenciázott)
  2. α-ra nincs megkötés (konstanssal, de trend nélkül): sztochasztikusan lehet benne trend (ez a tipikusabb)
  3. Determinisztikus – lineáris – trend kiszűrése után az előbbi (konstanssal és trenddel): a trendszűrt idősort teszteljük, azaz itt a trend-stacionaritást, és nem a stacionaritást tudjuk vizsgálni (azzal ekvivalens, hogy az $Y_t = \alpha_0 + \alpha_1 t + \phi_1 Y_{t-1} + u_t$-ből indulunk ki)
- (Esetleg másféle trend, vagy szezonális dummy-k is használhatóak)
- Hátrányok: sajnos kicsi lehet ez ereje ha a ϕ_1 kisebb mint 1, de csak kevéssel
- Hátrányok: csak akkor valid, ha az eredeti idősorra *tényleg* igaz volt az AR(1)-specifikáció (dinamikailag helyesen specifikált a modell, tehát tényleg ilyen alakú, és tényleg elég 1 késleltetés)

Egységgyök-tesztelés: ADF-teszt

- Próbáljuk kijavítani az előbbi hátrányt!
- Belátható, hogy ez úgy érhető el, ha áttérünk a

$$\Delta Y_t = \alpha + \delta_1 Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-2} + \ldots + \gamma_p \Delta Y_{t-p} + u_t$$

  modellre, ami akkor is működni fog, ha az eredeti folyamat magasabb – de p-nél nem nagyobb – rendű AR-folyamatot követ
- A rend megválasztása külön kérdés; általában információs kritériummal, vagy γ-k tesztelésével végzik (azoknak szerencsére szokásos, azaz t és – együttesen – F eloszlásaik vannak, legalábbis aszimptotikusan)