

Idősorelemzés

Ferenci Tamás
tamas.ferenci@medstat.hu

Utoljára frissítve: 2023. május 12.

Tartalom

Tartalomjegyzék

1	Az idősorelemzés alapjai	5
1.1	Az idősor fogalma, jelentősége, idősorelemzés	5
1.1.1	Az idősor fogalma	5
1.1.2	Az idősorelemzés közgazdasági jelentősége, története	7
1.1.3	Idősorelemzési iskolák, a módszerek felosztása	8
1.2	Idősorok jellemzői a sokaságban	8
2	Determinisztikus idősorelemzés, dekompozíciós idősormodellek: trend, szezonaritás és ciklus	11
2.1	Determinisztikus idősorelemzés	11
2.1.1	Alapgondolat	11
2.1.2	Determinisztikus idősormodellezés regresszióval	11
2.1.3	Trend és szezonaritás	14
3	Idősorok szűrése	17
3.1	Idősorok szűrése	17
4	A stacionaritás és az ergodicitás fogalma	23
4.1	A stacionaritás fogalma, szükségessége	23
4.2	Idősor-jellemzők mintából becslése	25
5	A sztochasztikus idősormodellezési filozófia, és alapelemei	27
5.1	Matematikai emlékeztető	27
5.1.1	Valószínűségszámítás emlékeztető	27
5.2	A sztochasztikus idősorelemzési iskola	28
5.3	ARMA-modellek	31
5.3.1	WN-folyamat	31
5.3.2	MA-modellek	31
5.3.3	AR-modellek	33
5.3.4	ARMA-modellek	35
6	Késleltetési operátor és polinom	37
6.1	Matematikai emlékeztető	37
6.1.1	Algebra emlékeztető	37
6.2	Az ARMA-folyamatok mélyebb matematikája	38
6.2.1	A késleltetési operátor és a késleltetési polinom	38
6.2.2	ARMA-folyamatok reprezentációja késleltetési polinomokkal	40

7	A stacionaritás tesztelése	43
7.1	A stacionaritás tesztelése	43
7.1.1	A stacionaritás teszteléséről általában	43
7.1.2	Egységgyök	45
8	A nemstacionaritás kezelése	49
8.1	Stacionarizálás	49
9	Box-Jenkins eljárás, előrejelzés készítése	53
9.1	Box-Jenkins eljárás	53
9.2	Előrejelzés készítése	54
10	Idősorok regressziója	57
10.1	Exogén változós idősormodellek	57
10.1.1	Alapgondolatok, statikus regresszió	57
10.1.2	Dinamikus regressziók	58
10.1.3	Idősoros regressziók általános modellje	60
10.2	Idősoros regresszió becslése OLS-sel	61
10.2.1	Standard modellfeltevések	61
10.2.2	Az OLS véges mintás tulajdonságai idősorokra	64
10.2.3	Az OLS nagymintás tulajdonságai idősorokra	64

1 Az idősorelemzés alapjai

1.1. Az idősor fogalma, jelentősége, idősorelemzés

1.1.1. Az idősor fogalma

Idősor sokaságban és mintában I.

- Mi a sokaság és a minta? – ismétlés stat 2-ből
- („Sokaságban valszám kell, mintánál statisztika”)
- **Idősor** minta értelemben: időben rendezett megfigyelések

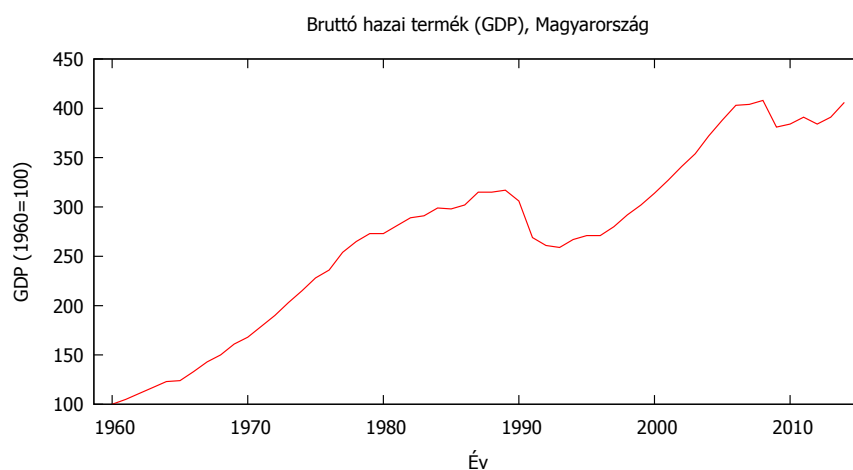
$$\{5492, 5640, \dots, 7317\},$$

általános jelöléssel

$$\{y_1, y_2, \dots, y_T\}$$

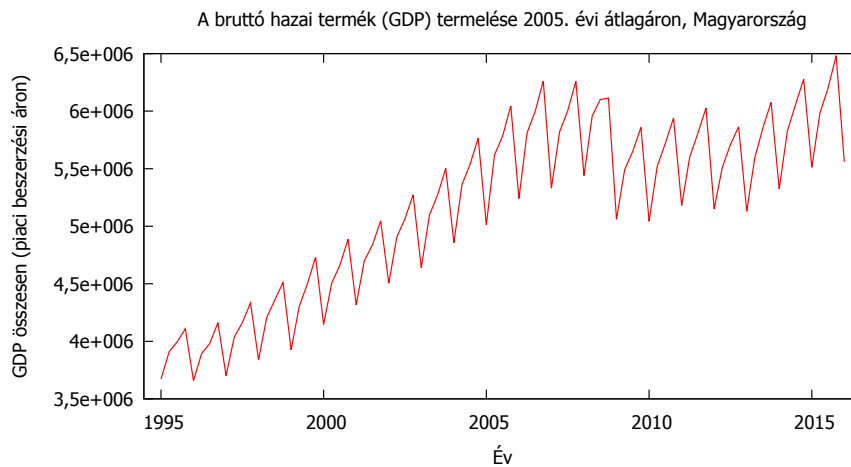
- Az egyetlen eltérés a keresztmetszethez képest: *sorrendjük* van (hatalmas jelentősége lesz majd)

Példa: magyar GDP (hosszú éves)

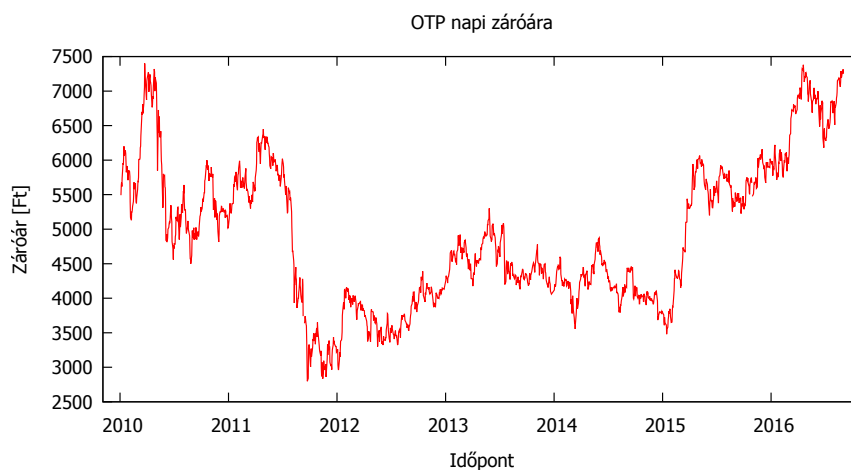


1 Az idősorelemzés alapjai

Példa: magyar GDP (negyedéves)



Példa: OTP napi záróár



Idősor sokaságban és mintában II.

- Sokasági definíció: valószínűségi változók indexelt halmaza (ahol az indexet „idő”-nek hívjuk):

$$\{Y_t, t \in N\}$$

- Az idősor egy időpontban tehát egy valószínűségi változó
- Itt az N tehát egy *rendezett* halmaz
- Valszámos szó: **sztochasztikus folyamat**

1.1 Az idősor fogalma, jelentősége, idősorelemzés

- Közgázban jellemzően N diszkrét, sőt véges: $N = \{1, 2, \dots, T\}$
- A minta tehát ennek egy realizációja (természetesen), itt a neve: **trajektória**

Mivel a valószínűségi változó pedig egy függvény ($X : \omega \rightarrow \mathbb{R}$), ezért a sztochasztikus folyamat igazából egy *kétváltozós* függvény: $Y : \omega \times N \rightarrow \mathbb{R}$. Ennek megfelelően két metszete van – úgy értve, hogy melyik paramétert rögzítjük – az egyik a trajektória, a másik egy adott időpillanathoz tartozó valószínűségi változó.

Egyetlen realizáció (reprodukálhatatlanság) problémája I.

- Mi a különbség 1406 lakás ára és az OTP 1406 napi záróárfolyama között?
- Mi van emögött?
- A függetlenség feltevés (tarthatósága)!

Egyetlen realizáció (reprodukálhatatlanság) problémája II.

- Elég, ha ismerjük az *egyes időpontok* eloszlásait? (F_{Y_1}, F_{Y_2} stb.)
- Ami teljes mértékben leírja az idősort: az *összes* időpont *együttes* eloszlása:

$$F_{Y_1, Y_2, \dots, Y_T}$$

- (Igazából keresztmetszetenél is ez volt, csak a függetlenségi feltevés miatt esett szét egyváltozósokra)

Egyetlen realizáció (reprodukálhatatlanság) problémája III.

- Azaz: 1 dimenziós adatra vett 100 megfigyelés vs. 100 dimenziósra vett 1 megfigyelés!
- (Ezért nem segít az sem, ha hosszabbítjuk az idősorunkat!)
- Hogyan lehet egyáltalán így bármi becsülni...?
- 1 megfigyelésből? Az érdekes lesz... → további feltevésekre lesz majd szükség!

1.1.2. Az idősorelemzés közgazdasági jelentősége, története

Az idősorelemzés közgazdasági jelentősége

- A legtöbb közgazdasági adat *igazából* idősorként érhető el!
- Számos feladatnál időbeli a fókusz, gondoljunk a – szó szoros értelmében vett – előrejelzési kérdésekre

Idősorelemzés a közgazdaságtanban

- Eleinte: egyszerű determinisztikus módszerek (pl. dekompozíciós modellek már a XIX. században)
- Később regresszió is, de még tekintet nélkül az idősoros jellegre
- Cochrane és Orcutt mutatott rá először 1949-ben, hogy ez nem jó ötlet
- Megindult a kutatás ennek figyelembevételére
- Box és Jenkins könyve 1970-ben fordulópont: sztochasztikus módszerek
- Korszerű eljárások és aktuális kérdések (nemstacionaritás, nemlinearitás, többváltozós módszerek, ARCH, ...)

1.1.3. Idősorelemzési iskolák, a módszerek felosztása

Az idősorelemzési módszerek csoportosítása

- Időtartomány vs. frekvenciatartomány (csak az előbbivel fogunk most foglalkozni)
- Determinisztikus vs. sztochasztikus (definíciós kavargások, most: a véletlennek van-e *folyamatépítő* szerepe; ld. később részletesen)
- Egyváltozós vs. többváltozós (+panel)

1.2. Idősorok jellemzői a sokaságban

Várhatóérték- és szórásnégyzet-függvény

Emlékezzünk rá, hogy egy adott időpontban az idősor egyszerűen egy valószínűségi változó, így definiálható várható értéke, szórásnégyzete, két ilyennek a kovarianciája; ez alapján:

- Várhatóérték-függvény ($\mu : N \rightarrow \mathbb{R}$):

$$\mu_t := \mathbb{E}Y_t$$

- Szórásnégyzet-függvény ($\sigma^2 : N \rightarrow \mathbb{R}_+$):

$$\sigma_t^2 := \mathbb{D}^2Y_t = \mathbb{E}(Y_t - \mathbb{E}Y_t)^2 = \mathbb{E}(Y_t - \mu_t)^2 = \mathbb{E}Y_t^2 - \mu_t^2$$

Autokovariancia- és autokorreláció-függvény

- Autokovariancia-függvény (ACVF, $\gamma : N \times N \rightarrow \mathbb{R}$):

$$\begin{aligned}\gamma_{t,s} &:= \text{cov}(Y_t, Y_s) = \mathbb{E}[(Y_t - \mathbb{E}Y_t)(Y_s - \mathbb{E}Y_s)] = \\ &\mathbb{E}[(Y_t - \mu_t)(Y_s - \mu_s)] = \mathbb{E}(Y_t Y_s) - \mu_t \mu_s\end{aligned}$$

- Nyilván $\gamma_{t,t} = \sigma_t^2$
- Autokorrelációs függvény (ACF, $\rho : N \times N \rightarrow [-1, 1]$):

$$\rho_{t,s} := \text{corr}(Y_t, Y_s) = \frac{\gamma_{t,s}}{\sigma_t \sigma_s}$$

Ne feledjük: mindezekben semmi sztochasztikus nincs, teljesen közönséges – determinisztikus – függvények!

Parciális autokorrelációs függvény (PACF)

- Úgy viszonyul az ACF-hez mint a sima (keresztmetszeti) parciális korreláció a korrelációhoz: bizonyos változókon keresztül terjedő hatásokat szűrjük (lineárisan)
- De melyikeket?
- Ami közbeesik: t és s közti korreláció ($s > t$), szűrve a $t+1, t+2, \dots, s-2, s-1$ időpontokon keresztül terjedő hatásokat
- Kiszámítható az ACF-ek ismeretében egyszerű mátrixműveletekkel
- (Avagy: a korreláció kijön egy olyan regresszióból, aminek egyetlen magyarázó változója van, a parciális korrelációhoz pedig hozzá kell adni a közbelső időpontokat is)

Korrelogram

- Korrelogram: ACF és PACF együtt ábrázolva
- (Egyelőre úgy tűnik, hogy ez egy kétdimenziós függvény, ezt majd később árnyalni fogjuk – és ezért nem is ábrázoljuk most még ténylegesen)
- Jelentősége: ha ismerjük nevezetes folyamatok – elméleti – korrelogramját, akkor egy minta mögötti, azt adó folyamatra következtethetünk az alapján, hogy a minta – empirikus – korrelogramja hogyan néz ki
- (Sajnos a gyakorlatban sokszor csak hozzávetőleges lehet)

2 Determinisztikus idősorelemzés, dekompozíciós idősormodellek: trend, szezonális és ciklus

2.1. Determinisztikus idősorelemzés

2.1.1. Alapgondolat

A determinisztikus idősorelemzés

- Az idősor alakulása *elvileg* függvényszerűen felírható bizonyos tényezők alapján
- Csak azért nem tudjuk tökéletesen megtenni, mert nem ismerjük e tényezőket, nem tudjuk milyen függvényformával hatnak, nem tudjuk pontosan mérni stb. ezért fogunk hibázni
- De pont: a hibának *csak* ennyi szerepe van...
- ...beállítja az aktuális időszakbeli értéket, és kész

Dekompozíciós idősormodellek

- Minderre a legtipikusabb – és egyben legklasszikusabb – példát a **dekompozíciós idősormodellek** jelentik
- A legismertebb additív modell:

$$Y_t = R_t + C_t + S_t + u_t,$$

ahol R_t , C_t és S_t a trend, a ciklus és a szezonális t -edik időszakbeli értéke rendre, u_t pedig a már említett eltérésváltozó

- Becslés?

2.1.2. Determinisztikus idősormodellezés regresszióval

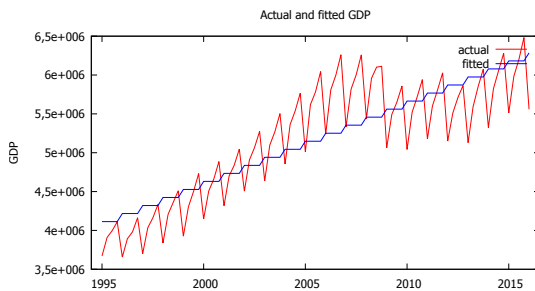
Regresszió alkalmazása

- Az előbbi modell teljesen természetesen becsülhető regresszióval, ha R_t , C_t és S_t helyébe beírjuk a feltételezett – paraméteres – függvényformákat

2 Determinisztikus időszorelemzés, dekompozíciós időszormodellek: trend, szezonális és ciklus

- (Most tehát mindvégig paraméteres regressziót fogunk használni)
- Legegyszerűbb eset: $R_t = \alpha + \beta t$, $C_t = 0$ és $S_t = 0$ (egyszerű lineáris trend)
- Az így kapott modell OLS-sel becsülhető

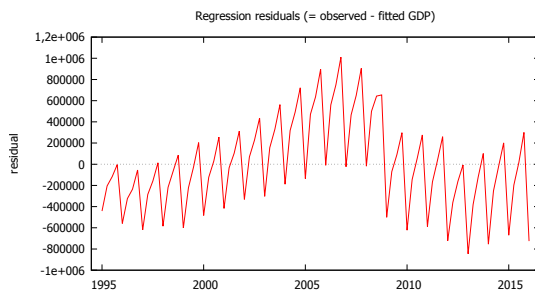
Negyedéves GDP (éves) lineáris trenddel I.



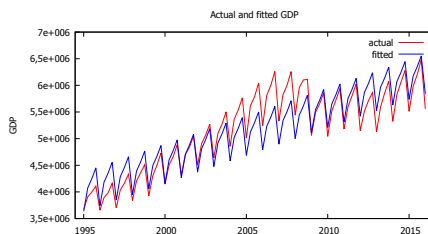
	Coefficient	Std. Error	t-ratio	p-value
const	-2,02165e+008	1,50629e+007	-13,4214	0,0000
EV	103398,	7512,17	13,7641	0,0000
Mean dependent var	5161052	S.D. dependent var	765270,3	
Sum squared resid	1,50e+13	S.E. of regression	424924,3	
R^2	0,695356	Adjusted R^2	0,691686	
$F(1, 83)$	189,4494	P-value(F)	3,94e-23	
Log-likelihood	-1221,169	Akaike criterion	2446,339	
Schwarz criterion	2451,224	Hannan-Quinn	2448,304	
$\hat{\rho}$	0,315141	Durbin-Watson	1,343686	

Negyedéves GDP (éves) lineáris trenddel II.

Mi ezzel a baj? Hibátag jól specifikált? Aligha!



Negyedéves GDP (éves) lineáris trenddel és szezonálisitással I.

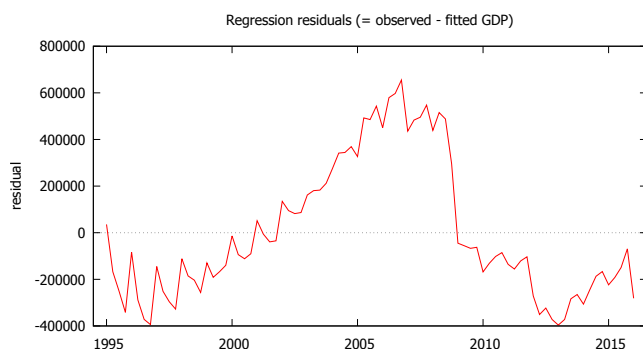


2.1 Determinisztikus idősorelemzés

	Coefficient	Std. Error	t-ratio	p-value
const	-2,04994e+008	1,06300e+007	-19,2845	0,0000
EV	104985,	5301,64	19,8024	0,0000
DNEGYEDEV_1	-815807,	91469,3	-8,9189	0,0000
DNEGYEDEV_2	-375072,	92487,9	-4,0554	0,0001
DNEGYEDEV_3	-203380,	92487,9	-2,1990	0,0308
Mean dependent var	5161052	S.D. dependent var	765270,3	
Sum squared resid	7,19e+12	S.E. of regression	299695,1	
R ²	0,853937	Adjusted R ²	0,846634	
F(4, 80)	116,9271	P-value(F)	1,34e-32	
Log-likelihood	-1189,928	Akaike criterion	2389,855	
Schwarz criterion	2402,068	Hannan-Quinn	2394,768	
$\hat{\rho}$	0,946516	Durbin-Watson	0,116617	

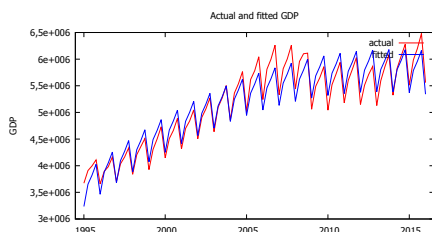
Negyedéves GDP (éves) lineáris trenddel és szezonalitással II.

A szezonális jónak tűnik, de az alaptrendet még mindig nem sikerült megragadni:



A szezonális azért tűnik jónak, mert nincs interakció az év és a szezon között, azaz minden évben hasonló a szezonális mintázata.

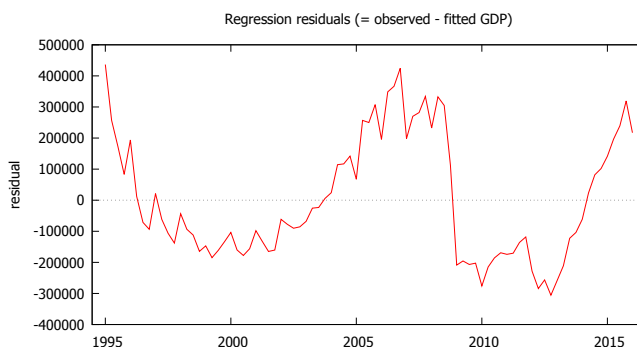
Negyedéves GDP (éves) kvadratikus trenddel és szezonalitással I.



	Coefficient	Std. Error	t-ratio	p-value
const	-2,60273e+010	2,61365e+009	-9,9582	0,0000
EV	2,58613e+007	2,60697e+006	9,9201	0,0000
DNEGYEDEV_1	-792077,	61608,7	-12,8566	0,0000
DNEGYEDEV_2	-375072,	62247,4	-6,0255	0,0000
DNEGYEDEV_3	-203380,	62247,4	-3,2673	0,0016
sq_EV	-6422,59	650,072	-9,8798	0,0000
Mean dependent var	5161052	S.D. dependent var	765270,3	
Sum squared resid	3,21e+12	S.E. of regression	201704,8	
R ²	0,934664	Adjusted R ²	0,930529	
F(5, 79)	226,0280	P-value(F)	2,80e-45	
Log-likelihood	-1155,736	Akaike criterion	2323,473	
Schwarz criterion	2338,128	Hannan-Quinn	2329,368	
$\hat{\rho}$	0,889365	Durbin-Watson	0,173391	

Negyedéves GDP (éves) kvadratikus trenddel és szezonálítással II.

Reziduumok kicsit jobbak:



Mindezek limitációi

- *Egyrészt* el kell találni a függvényformát
- Persze modelldiagnosztika (az előbb látott grafikus módszerek és tesztek is) ott van
- (Ez igazából már keresztmetszetnél is így volt)
- Pl. a kvadratikus nyilván csak erre az időszakra jó, az általánosítóképessége botrányos lenne
- *Másrészt* a hibatag diagnosztikája bonyolultabbá válik, egy új szempont is megjelenik (autokorreláció) → később még nagyon sokat fogunk róla beszélni

2.1.3. Trend és szezonális

A trend megadása

- **Trend:** „hosszú távú alapisirányzat”
- A mostani trend (determinisztikus trend) bármi lehet, amit paraméteres függvényformában megadunk; például:
 - Lineáris trend: $a + bt$
 - Kvadratikus trend: $a + b_1t + b_2t^2$
 - Polinomiális trend: $a + b_1t + b_2t^2 + \dots + b_kt^k$
 - Exponenciális trend: ae^{bt}
 - Aszimptotikus trend: $c + \frac{1}{a+bt}$
 - Logisztikus trend: $\frac{1}{c+e^{a+bt}}$
 - stb. stb.

- (Persze amelyik nem lineáris, ott vagy linearizálni kell vagy – ha ez nem lehetséges – akkor nem OLS-sel becsülni)
- Ezek mind paraméteres trendek voltak, elképzelhető nem-paraméteres trend is, a legismertebb a spline-ok használata (de ne feledjük, annak a becslése kevésbé hatásos, nem kapunk egyetlen vagy néhány számba sűrített – és jó esetben tárgyterületileg értelmezhető – eredményt, valamint az előrejelzés is problémásabb)

Szezonális megadása

- **Szezonális:** „éven belüli mintázat”, exogén módon rögzített hosszúságú, periodikus (vs. **ciklus:** „éven túli”, nem feltétlenül exogén módon adott, ismert hosszúságú)
- A szezonálisnál viszont tipikusabb a nem-paraméteres megadás: minden negyed-évnél (hónapnak, félévnek stb.) saját paramétere van
- (Dummy-kkal, ld. később, regressziós keretbe szintén szépen illeszkednek!)
- Persze itt is elképzelhető paraméteres megadás, a legismertebb a trigonometrikus (harmonikus) függvények használata

Dummy-kódolás szezonálitáshoz: referenciakódolás

- Az egyik szezon indikátorát elhagyjuk: **referenciakódolás**

	D_{Q1}	D_{Q2}	D_{Q3}
Q1	1	0	0
Q2	0	1	0
Q3	0	0	1
Q4	0	0	0

- Értelmezés: eltérés a referenciacsoporthoz képest (ami az elhagyott indikátorú csoport)

Dummy-kódolás szezonálitáshoz: kontrasztkódolás I.

- Egy másik népszerű megoldás a **kontrasztkódolás:** viszonyítsunk az *átlaghoz!*
- Ehhez hogyan kell kódolni...?

	C_{Q1}	C_{Q2}	C_{Q3}
Q1	1	0	0
Q2	0	1	0
Q3	0	0	1
Q4	-1	-1	-1

Dummy-kódolás szezonálitáshoz: kontrasztkódolás II.

Mert:

$$\alpha + \beta_{C_{Q_1}} + 0 + 0 = \bar{y}_{Q_1} \quad (2.1)$$

$$\alpha + 0 + \beta_{C_{Q_2}} + 0 = \bar{y}_{Q_2} \quad (2.2)$$

$$\alpha + 0 + 0 + \beta_{C_{Q_3}} = \bar{y}_{Q_3} \quad (2.3)$$

$$\alpha - \beta_{C_{Q_1}} - \beta_{C_{Q_2}} - \beta_{C_{Q_3}} = \bar{y}_{Q_4} \quad (2.4)$$

És így:

- (1)+(2)+(3)+(4) $\Rightarrow 4\alpha = \bar{y}_{Q_1} + \bar{y}_{Q_2} + \bar{y}_{Q_3} + \bar{y}_{Q_4} \Rightarrow \alpha$ tényleg a főátlag (mert azonosak voltak a csoportok elemszámai, különben ún. súlyozott kontraszt kellene)
- (2)+(3)+(4) $\Rightarrow 3\alpha - \beta_{C_{Q_1}} = \bar{y}_{Q_2} + \bar{y}_{Q_3} + \bar{y}_{Q_4} \Rightarrow \beta_{C_{Q_1}} = 3\alpha - (\bar{y}_{Q_2} + \bar{y}_{Q_3} + \bar{y}_{Q_4}) = 3\alpha - (4\alpha - \bar{y}_{Q_1}) \Rightarrow \beta_{C_{Q_1}} = \bar{y}_{Q_1} - \alpha \Rightarrow$ tényleg az átlagtól való eltérés (és hasonlóan a másik kettő)

Dummy-kódolás szezonálitáshoz: egyebek

- Az angol irodalomban az általunk kontrasztkódolásnak nevezett módszert nagyon gyakran „effect coding”-nak nevezik...
- ... a kontraszt pedig az, amikor a csoportok tetszőleges – általunk meghatározott – lineáris kombinációját teszteljük

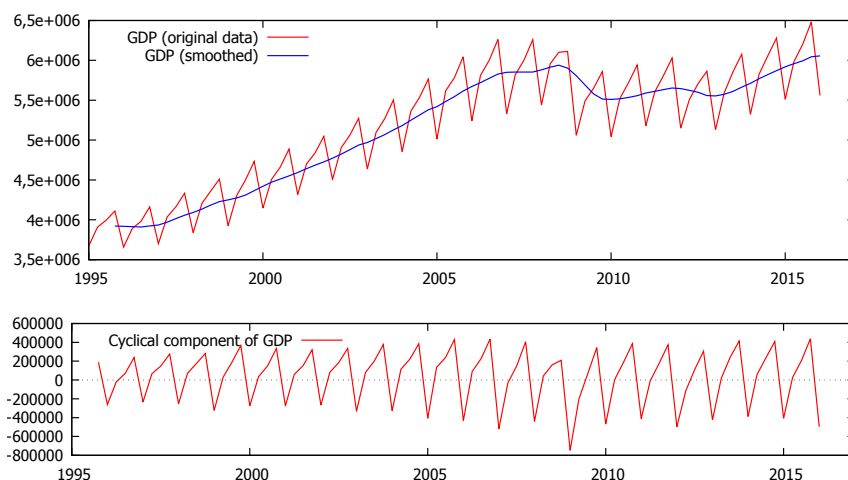
3 Idősorok szűrése

3.1. Idősorok szűrése

Célunk

- Szeretnénk elkülöníteni a trendet és a ciklikus komponenseket (szezón + ciklus); ez számos közgazdasági kérdésnél fontos feladat
- Egy lehetséges megoldás: „paraméteres szűrés”, azaz előírjuk a függvényformát és regresszióval megbecsüljük
- Igazából ezt megtettük az előbb is: megadtuk a trendet (lineáris vagy kvadratikus), megadtuk a szezonaritást (ezt nem-paraméteresen), és a látott reziduum a kiszűrt ciklus (+zaj) volt, ha van ilyen
- De ez függ a függvényforma-választás helyességén; nem lehetne ilyen feltevések nélkül is megtenni?
- Hogyne, például átlagoljunk ki évenként!
- Nem a legjobb, abrupt ugrások év végén, átlagoljunk inkább folyamatosan, mintegy csúszóablakot végigtolva (így is mindig négy különböző szezon lesz benne!)

Motiváló példa



Mozgóátlagolás

- Ez volt az (egyszerű) **mozgóátlag**:

$$y'_t = \frac{y_t + y_{t-1} + y_{t-2} + \dots + y_{t-p}}{p+1}$$

- (Néha nem visszafelé átlagolnak, hanem az eredmény az átlagolt ablak közepén van („centrált” mozgóátlag), a dolognak nincs nagy jelentősége: ez talán kicsit jobban néz ki, viszont jövőbeli megfigyeléseket is igényel)
- Ez azt jelenti, hogy minden megfigyelés ugyanolyan súlyú; kézenfekvő gondolat, hogy a régebbiek kevésbé számítsanak
- Például:

$$y'_t = \frac{py_t + (p-1)y_{t-1} + (p-2)y_{t-2} + \dots + y_{t-(p-1)}}{p + (p-1) + (p-2) + \dots + 1}$$

- Ez az ún. súlyozott mozgóátlag

Exponenciális mozgóátlag

- Ökonometriában gyakoribb az exponenciális súlyozás:

$$y'_t = \alpha y_t + (1 - \alpha) y'_{t-1},$$

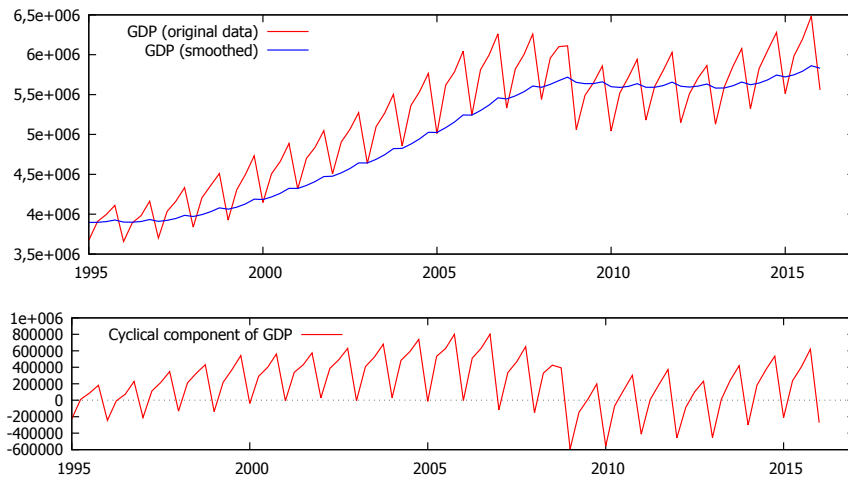
hiszen ez nyilván annak felel meg, hogy

$$y'_t = \alpha y_t + (1 - \alpha) \alpha y_{t-1} + (1 - \alpha)^2 \alpha y_{t-2} + \dots + (1 - \alpha)^{t-2} \alpha y_2 + (1 - \alpha)^{t-1} \alpha y_1 + (1 - \alpha)^t \alpha y_0$$

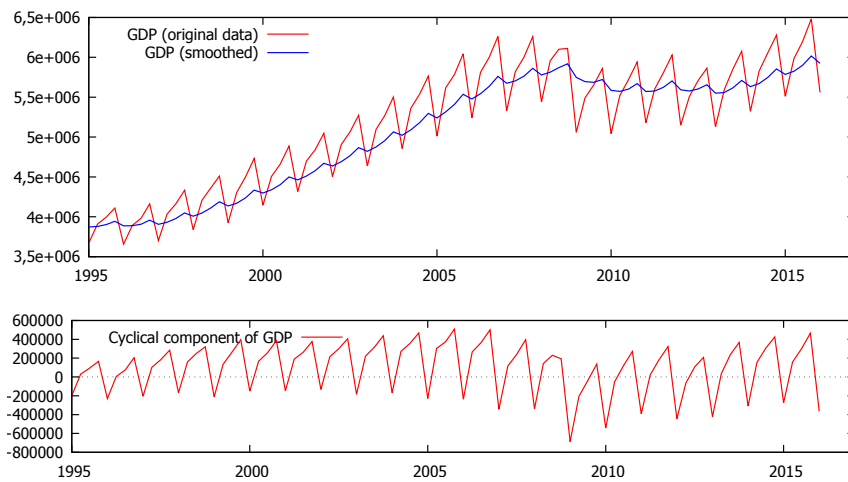
- (Az y_0 kezdőértéket nekünk kell megadni, a tipikus választások: y_1 , az első néhány megfigyelés átlaga, az egész idősor átlaga; a következő példákban az első 4 megfigyelés átlaga került alkalmazásra)
- Tehát az ablak egyre nyúlik (végig az egész tartomány), és a súlyok exponenciálisan csengenek le

Negyedéves GDP exponenciális mozgóátlagolása, $\alpha = 0,1$

3.1 Idősorok szűrése

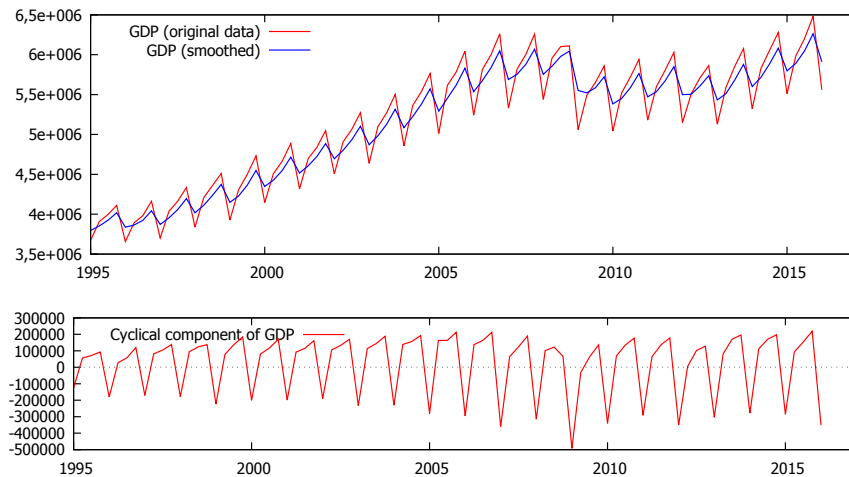


Negyedéves GDP exponenciális mozgóátlagolása, $\alpha = 0,2$

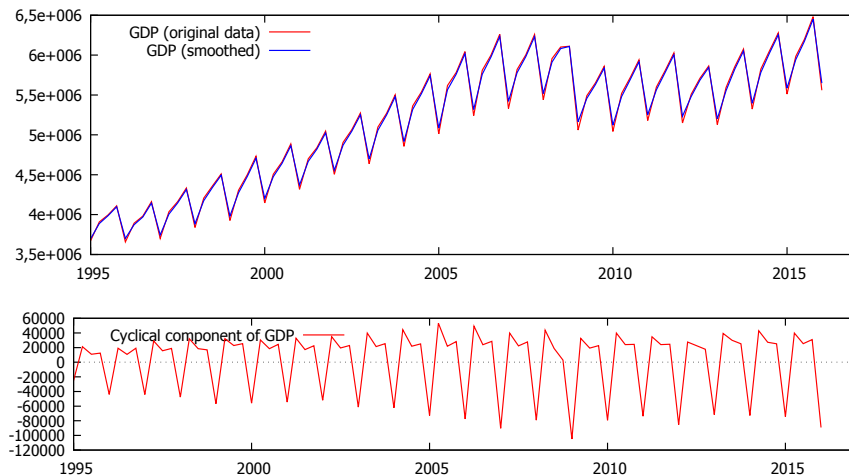


Negyedéves GDP exponenciális mozgóátlagolása, $\alpha = 0,5$

3 Idősorok szűrése



Negyedéves GDP exponenciális mozgóátlagolása, $\alpha = 0,9$



Lineáris szűrő

- Az egyszerű és a súlyozott mozgóátlag speciális esete annak, hogy

$$y'_t = a_0 y_t + a_1 y_{t-1} + \dots + a_p y_{t-p} = \sum_{i=0}^p a_i y_{t-i}$$

- Például egyszerű mozgóátlagra $a_i = 1/(p+1)$
- Lényegében diszkrét konvolúció
- Ezt hívjuk lineáris szűrőnek, a tulajdonságait értelemszerűen teljes mértékben meghatározzák az a_0, a_1, \dots, a_p szűrőegyütthatók

- Roppant érdekes kérdés, hogy a szűrt idősor hogyan néz ki az eredetihez képest a szűrőegyütthatók függvényében, lehet otthon kísérletezgetni...

Hodrick–Prescott-szűrő

- Különösen makroökonómiában népszerű
- Alapgondolat: a trend (g_t) követi az idősort (azaz az $y_t - g_t$ kicsi), de nem nagyon ugrándoza (azaz g_t sima)
- A megoldandó feladat:

$$\min_{\{g_t\}_{t=1}^T} \sum_{t=1}^T (y_t - g_t)^2 + \lambda \sum_{t=2}^{T-1} [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2$$

- Első tag: mennyire követi jól az idősort a trend, második tag: mennyire „rángatózik” a trend
- A λ együttható határozza meg a két szempont egymáshoz viszonyított fontosságát ($\lambda = 0$: a trend nem kell, hogy sima legyen \rightarrow pontosan az idősor lesz; $\lambda \rightarrow \infty$: trend legyen tökéletesen sima, nem érdekes, hogy mennyire követi az idősort \rightarrow egyenes lesz)
- Negyedéves adatokra a tipikus választás $\lambda = 1600$

A második tagra azt írtuk, hogy „mennyire rángatózik a trend”; később majd pontosabban is fogjuk látni, hogy ez micsoda, mert van egy nagyon konkrét tartalma.

A HP-szűrő matematikája I.

- Az érdekes az, hogy a fenti minimalizációs feladatnak van zárt alakú megoldása!
- Legyen \mathbf{y} az y_t -k, \mathbf{g} a g_t -k vektora és legyen

$$\mathbf{Q}_{(T-2) \times T} = \begin{pmatrix} 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -2 & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & -2 & 1 \end{pmatrix},$$

akkor a megoldandó feladat

$$\min_{\mathbf{g}} (\mathbf{y} - \mathbf{g})^T (\mathbf{y} - \mathbf{g}) + \lambda (\mathbf{Q}\mathbf{g})^T (\mathbf{Q}\mathbf{g})$$

A HP-szűrő matematikája II.

- Deriválva \mathbf{g} szerint:

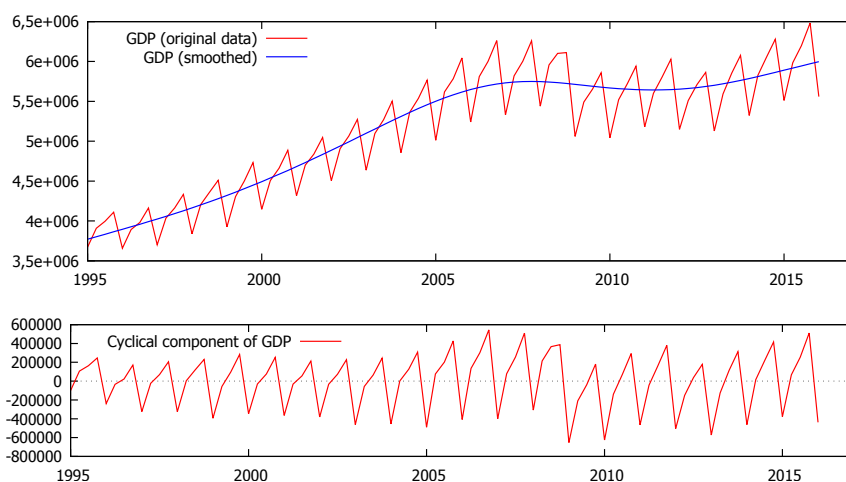
$$-2\mathbf{y} + 2\mathbf{g} + 2\lambda\mathbf{Q}^T\mathbf{Q}\mathbf{g}$$

- Egyenlővé téve nullával és megoldva (a másodrendű feltételek teljesülnek ahhoz, hogy ez tényleg minimum legyen):

$$-2\mathbf{y} + 2\mathbf{g} + 2\lambda\mathbf{Q}^T\mathbf{Q}\mathbf{g} = 0 \Rightarrow \hat{\mathbf{g}} = (\mathbf{I} + \lambda\mathbf{Q}^T\mathbf{Q})^{-1} \mathbf{y}$$

- (A megoldás mindegyik megfigyeléstől függ)

Negyedéves GDP HP-szűrése, $\lambda = 1600$



A HP-szűrő kritikája

Érdekes olvasmány: Hamilton, James D. „Why You Should Never Use the Hodrick-Prescott Filter.” (2016). <http://econweb.ucsd.edu/~jhamilto/hp.pdf>.

Egyéb szűrők

- Több szűrő viselkedése alapvetően frekvencia-tartományon érthető meg
- Trend/ciklus szétválasztás: aluláteresztő szűrés (mert a trend az, ami lassan változik, a ciklus az, ami gyorsabban, persze kérdés, hogy hol a határ)
- Igazából már az egyszerű mozgóátlag is egy aluláteresztő szűrő volt!
- Vagy: LOESS-használata (STL-dekompozíció)

4 A stacionaritás és az ergodicitás fogalma

4.1. A stacionaritás fogalma, szükségessége

Az alapprobléma

- Egyetlen realizáció problémája
- Hogyan lehet egyetlen mintából bármit megbecsülni? Nyilván sehogy...
- És ezen ráadásul – ugyebár – a hosszabb megfigyelés sem segít
- A megoldás: valamilyen plusz-feltevés kell!

Jön a jótündér

- Mondjuk egy jótündér megsúgja nekünk, hogy minden időpontban (ami ugye egy valószínűségi változó) ugyanaz a várható érték:

$$\mu_t \equiv \mu$$

- (A \equiv természetesen azt jelenti, hogy $\forall t$ -re)
- Ez tehát egyfajta *eltolásinvariancia*; írhattuk volna, teljesen egyenértékűen azt is, hogy $\mu_t = \mu_s$, vagy azt is, hogy $\mu_t = \mu_{t+h}$ (most elég nyakatekert, de később ez lesz a jobb)
- Ekkor már ezt az – immár létező – *közös* várható értéket, igaz *csak ezt*, de ezt meg tudjuk becsülni mintából!
- Ehhez – de csak ehhez! – „összeönthetők” a különböző időponthoz tartozó megfigyelések (igazából ehhez még valami további dolog is kell, de erről majd később, most fogadjuk el, hogy megvalósítható)

- Hiszen:

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T y_t$$

- Menjünk tovább...

További jótünetek

- Ha $\sigma_t^2 \equiv \sigma^2$ akkor létezik közös szórásnégyzet, és becsülhető: $\widehat{\sigma^2} = \frac{1}{T} \sum_{t=1}^T (y_t - \widehat{\mu})^2$ (torzított, de aszimptotikusan torzítatlan, továbbá konzisztens)
- Menjünk tovább...
- Ha létezik az $F_{Y_i} \equiv F$ közös eloszlás, akkor \widehat{F} becsülhető (pl. hisztogrammal)
- Menjünk tovább...
- Két változónál is az eltolásinvarianciát akarjuk megtartani: $\gamma_{t,s}$ legyen ugyanaz mint $\gamma_{t+h,s+h}$
- De gondoljuk végig, ez magyarul azt jelenti, hogy $\gamma_{t,s}$ csak a $t - s$ -től függ!
- Ez esetben becsülhető: $\widehat{\gamma}_k = \frac{1}{T-k} \sum_{t=1}^{T-k} (y_t - \widehat{\mu})(y_{t+k} - \widehat{\mu})$
- Menjünk tovább...
- $F_{Y_t, Y_s} = F_{Y_{t+h}, Y_{s+h}}$, minden s, t, h -ra, akkor a kétváltozós eloszlás komplettül becsülhető

És végül...

- Az utolsó, immár tényleg mindent lefedő szint: $F_{Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}} = F_{Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_k+h}}$ minden értelmes k -ra, t_1, t_2, \dots, t_k -ra és h -ra
- Ennek a neve: **erős stacionaritás**
- Ez az a feltevés, ami a keresztmetszethez hasonló becsülhetőséget tesz lehetővé annak ellenére is, hogy idősorban vagyunk
- A keresztmetszet azért volt egyszerűbb, mert a függetlenséget feltettük, de itt most erről szó nincs: pont az a lényeg, hogy úgy teremtettük meg a keresztmetszethez hasonló becsülhetőséget, hogy semmilyen függetlenséget nem kellett feltételeznünk – szerencsére, mert annak ugye nem lenne értelme (persze a függetlenség implikálja az erős stacionaritást)
- Viszont: rengeteget követel, elméleti kezeléshez jó, de gyakorlatban nagyon nehéz ellenőrizni a teljesülését
- Éppen ezért gyengítsük kicsit...

A gyenge stacionaritás

- Mint az erős stacionaritás, de
 - csak $k = 1, 2$ -re
 - teljes eloszlás-egyezőség helyett csak első két momentumban egyezőséget követelünk meg
- Mi adódik ebből?
 1. $\mu_t \equiv \mu$
 2. $\sigma_t^2 \equiv \sigma^2$
 3. $\gamma_{t,s} \equiv \gamma_{t-s}$
- (Valójában a 2. feltétel redundáns, hiszen $\gamma_{t,t} = \sigma_t^2$)
- A továbbiakban, ha mást nem mondunk, stacionaritás alatt mindig ezt a **gyenge (vagy kovariancia-) stacionaritást** értjük

A stacionaritás tesztelése

- A fentiek már adnak egy – teljesen szubjektív – módszert a stacionaritás tesztelésére: nézzünk rá az idősorra, az 1. és 2. feltétel megítélhető
- Ez a „grafikus tesztelés” persze abszolút szubjektív
- Később látni fogunk objektív módszert (statisztikai próbát) is

4.2. Idősor-jellemzők mintából becslése**Egy gondolat a mintából történő becslésekről**

- A már látott mintából történő becsléseknél ($\hat{\mu}$, $\hat{\sigma}^2$, $\hat{\gamma}_k$ stb.) ne feledjük el, hogy mindezeket mind terheli a mintavételi ingadozás, az abból fakadó mintavételi hiba
 - ...konfidenciaintervallum szerkeszthető a valódi értékre
 - ...tesztelhető nevezetes értékre vonatkozó hipotézis
- Ez utóbbi tipikus példája az autokorrelálatlanság ($H_0 : \rho_k = 0$ vs. $H_1 : \rho_k \neq 0$)
- Nem túl kis mintaméretnél már jó a normális approximáció: $\hat{\rho}_k \sim \mathcal{N}(\rho_k, 1/T)$, így

$$\frac{\hat{\rho}_k}{1/\sqrt{T}} \stackrel{H_0}{\sim} \mathcal{N}(0, 1)$$

Autokorrelálatlanság tesztelése: Ljung–Box-teszt

- Nagyon sok esetben grafikusán is feltüntetik a korrelogramon az autokorrelációra vonatkozó $\pm z_{1-\alpha/2} \frac{1}{\sqrt{T}}$ kritikus értékeket
- De vigyázat: ez a *teljes* autokorrelálatlanság ($H_0 : \rho_1 = \rho_2 = \dots = \rho_M = 0$) tesztelésére nem alkalmas!
- (Többszörös összehasonlítások helyzete, α -infláció)
- Legnépszerűbb teszt erre: Ljung–Box-teszt:

$$Q = T(T+2) \sum_{k=1}^M \frac{\hat{\rho}_k^2}{T-k} \stackrel{H_0}{\sim} \chi_M^2$$

- M megválasztása kérdés (ha túl kicsi, elnézhetünk egy magasabbrendű autokorrelációt, ha túl nagy, eltérhetünk a χ^2 eloszlástól)!
- Tipikus alkalmazás majd: modelldiagnosztikában

Autokorrelálatlanság tesztelése: LM-tesztek

- A Ljung–Box-teszttel komoly elméleti agályok vannak (ld. Maddala, 13.5 vagy Hayashi 2.10)
- Ennek ellenére teljesen általánosan használják...
- Lehetséges alternatíva: LM-elvű tesztek, például a Breusch–Godfrey-teszt a már említett modelldiagnosztikai helyzetben

5 A sztochasztikus idősormodellezési filozófia, és alapelemei: a fehérzaj-, az AR-, az MA- és ARMA-folyamatok

5.1. Matematikai emlékeztető

5.1.1. Valószínűségi számítás emlékeztető

Várható érték

Ki fogjuk használni a következőket:

- A várható érték lineáris: $\mathbb{E}(\sum_i X_i) = \sum_i \mathbb{E}X_i$
- A várható érték lineáris: $\mathbb{E}(cX) = c\mathbb{E}X$
- Konstans várható értéke saját maga: $\mathbb{E}c = c$

Szórásnégyzet

Ki fogjuk használni a következőket:

- A szórásnégyzet nem lineáris: $\mathbb{D}^2(\sum_i X_i) = \sum_i \mathbb{D}^2 X_i$ ha X_i -k (páronként) korrelálatlanok (szemben a várható értékkel, ez *nem* mindig igaz!); ne feledjük, a függetlenség implikálja a korrelálatlanságot
- A szórásnégyzet nem lineáris: $\mathbb{D}^2(cX) = c^2 \mathbb{D}^2 X$
- Konstans szórásnégyzete nulla: $\mathbb{D}^2 c = 0$

Kovariancia és korreláció

Ki fogjuk használni a következőket:

- A kovariancia/korreláció bilineáris: $\text{cov}(\sum_i X_i, \sum_j Y_j) = \sum_i \sum_j \text{cov}(X_i, Y_j)$
- A kovariancia/korreláció bilineáris: $\text{cov}(aX, bY) = ab \text{cov}(X, Y)$
- Konstans mindennel korrelálatlan: $\text{cov}(c, X) = 0$
- Az önkovariancia a variancia: $\text{cov}(X, X) = \mathbb{D}^2 X$

5.2. A sztochasztikus idősorelemzési iskola

Filozófiai alapok

- Determinisztikus (például dekompozíciós idősormodellek) vs. sztochasztikus idősorelemzés
- A determinisztikus iskolában is van – természetesen – véletlen, csak a szerepe más: pusztán arra korlátozódik, hogy az *adott időszaki* értéket beállítsa
- A sztochasztikus iskolában ezzel szemben a véletlen az *egész későbbi* lefutást befolyásolja, a véletlennek „folyamatépítő szerepe” van
- Lássunk egy példát, hogy jobban megértsük mit jelentenek ezek a kissé homályos megfogalmazások!

Példa a két iskolára

- Az egyik idősorunk – sokasági modellel megadva – legyen

$$Y_t^{(D)} = \alpha t + u_t,$$

ahol α konstans, $u_t \sim \mathcal{N}(0, \sigma^2)$ függetlenül

- A másik modell legyen

$$Y_t^{(S)} = Y_{t-1}^{(S)} + \alpha + u_t,$$

ahol α és u_t mint előbb, $Y_0^{(S)}$ pedig legyen 0

- A további elemzésekhez hasznos lesz a következő átalakítás:

$$\begin{aligned} Y_t^{(S)} &= Y_{t-1}^{(S)} + \alpha + u_t = \left(Y_{t-2}^{(S)} + \alpha + u_{t-1} \right) + \alpha + u_t = \\ &= \left[\left(Y_{t-3}^{(S)} + \alpha + u_{t-2} \right) + \alpha + u_{t-1} \right] + \alpha + u_t = \dots = \alpha t + \sum_{i=1}^t u_i \end{aligned}$$

- Hasonlítanak is, meg nem is...

Hasonlóság

Számítsuk ki a μ_t várható érték függvényeket:

$$\mu_t^{(D)} = \mathbb{E}(\alpha t + u_t) = \mathbb{E}(\alpha t) + \mathbb{E}(u_t) = \alpha t + 0 = \alpha t$$

$$\mu_t^{(S)} = \mathbb{E} \left(\alpha t + \sum_{i=1}^t u_i \right) = \alpha t + \sum_{i=1}^t 0 = \alpha t$$

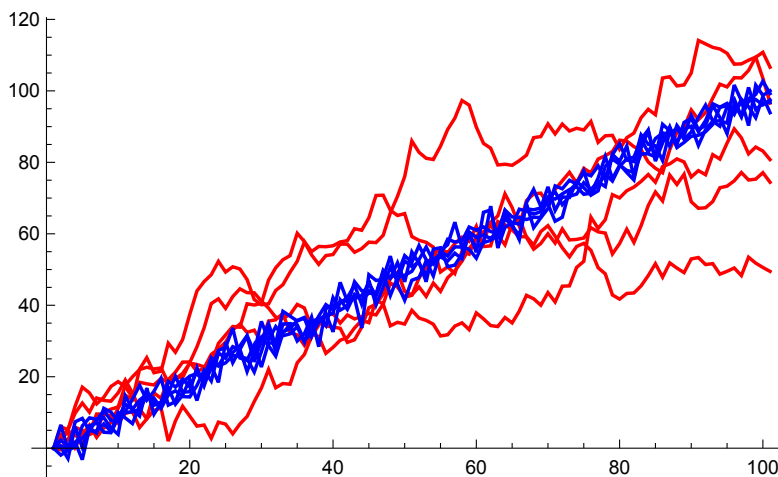
Különbség

Nézzük most meg a σ_t^2 szórásnégyzet függvényeket:

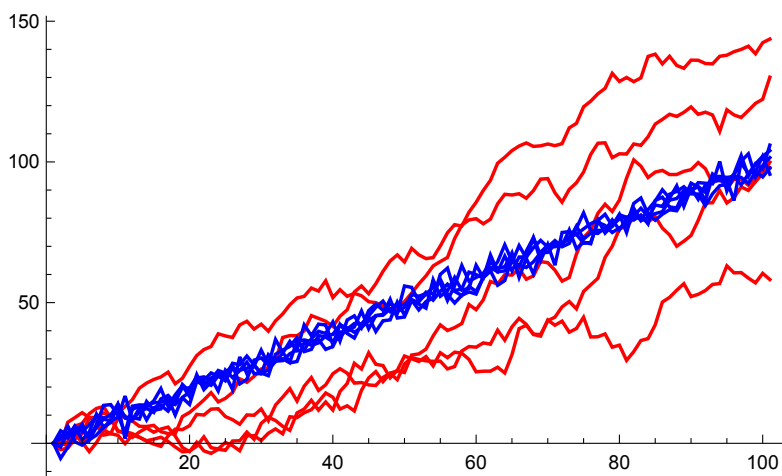
$$\begin{aligned}\sigma_t^{2(D)} &= \mathbb{D}^2(\alpha t + u_t) = \mathbb{D}^2(\alpha t) + \mathbb{D}^2(u_t) = 0 + \sigma^2 = \sigma^2 \\ \sigma_t^{2(S)} &= \mathbb{D}^2\left(\alpha t + \sum_{i=1}^t u_i\right) = \mathbb{D}^2(\alpha t) + \mathbb{D}^2\left(\sum_{i=1}^t u_i\right) = 0 + \sum_{i=1}^t \sigma^2 = \\ &= t\sigma^2\end{aligned}$$

Az igazi eltérés a viselkedésben

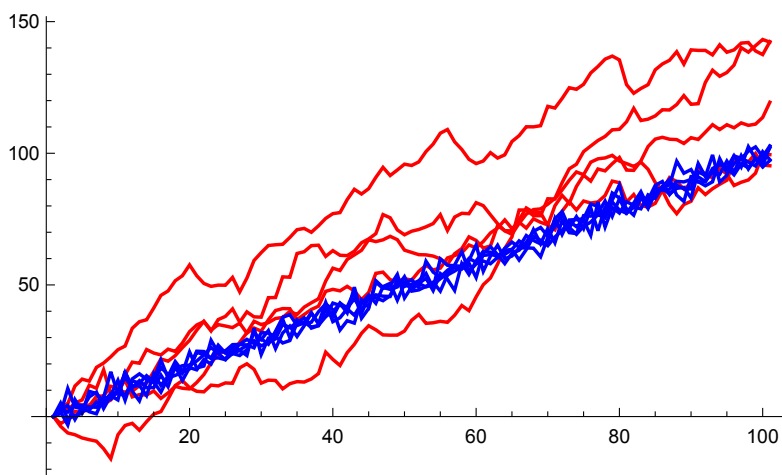
- Ennek sokkal mélyebb magyarázatát kapjuk, ha arra gondolunk, hogy a viselkedésük miben más
- Segítség: kidobunk egy nagyon deviáns u_t -t (pl $\sigma^2 = 1$ mellett +5-öt vagy -5-öt), miben tér el a két idősor későbbi viselkedése?
- Ez a két extrém véglet:
 - Az $Y_t^{(D)}$ -nél már a *rögtön következő* időpontban sincsen *semmilyen* hatása ennek
 - $Y_t^{(S)}$ -nél viszont az idősor *egész későbbi lefutását* befolyásolja, *csorbítatlanul*
- Bizonyos értelemben *ugyanazt* a trendet jelentik – gondoljunk a várható érték függvényre – de mégis teljesen eltérő *viselkedéssel*
- Megtestesítik a két iskolát: $Y_t^{(D)}$ a lineáris trend a determinisztikus szemléletben (véletlen szerepe: csak az adott időszakra korlátozódik), az $Y_t^{(S)}$ a lineáris trend sztochasztikus értelemben (véletlen szerepe: folyamatépítő)

Hogy néznek ki?

Hogy néznek ki?



Hogy néznek ki?



Más elnevezések

- Főleg sztochasztikus folyamatos kontextusban az $Y_t^{(S)}$ -et $\alpha = 0$ esetén **véletlen bolyongásnak** (random walk, RW) is szokás nevezni
- Rárákok egy bábut az origóra a számegeyenesen, dobok egy véletlen számot (u_t) és annyival odébb rakom, majd ezt ismétlem \rightarrow bolyongani fog a számegeyenesen
- (Folytonos határa a Wiener-folyamat)
- Az $\alpha \neq 0$ esetben pedig **eltolásos véletlen bolyongásról** (random walk with drift, RWD) szoktak beszélni

- Amennyiben az RW-t log-skálán vesszük, tehát $\log Y_t = \log Y_{t-1} + u_t'$ a modellünk, $Y_0 \neq 0$ mellett (eredeti skálára visszavetítve: $Y_t = Y_{t-1} \cdot u_t = Y_0 \cdot \prod_{i=1}^t u_i$; nem a növekmények, hanem a hányadosok adott fae változók, nagyobb értékeknél nagyobb ingadozás) akkor **geometriai véletlen bolyongásról** szokás beszélni (pénzügyes szóhasználatban!)

5.3. ARMA-modellek

5.3.1. WN-folyamat

A fehérzaj (WN) folyamat

- A folyamat

$$u_t,$$

melyre $\mathbb{E}(u_t) = 0$, $\mathbb{D}^2(u_t) = \sigma^2$ és $\text{cov}(u_t, u_s) = 0$ ($t \neq s$)

- Jele: $\mathcal{WN}(0, \sigma_u^2)$
- Az eloszlásról nem mondtunk semmit
- Néha feltesszük, hogy nem csak korrelálatlan, de független is (általában nem ez az alapértelmezés, külön kell mondani); egyedül normális eloszlás feltevése esetén mindegy
- Zaj: logikus, ez valamilyen teljesen modellezhetetlen, struktúra nélküli folyamat
- De mitől fehér? ...optikai analógia!

A fény is felfogható egy idősornak, hiszen egy elektromágneses rezgés. Itt a spektrális elemzés nagyon jól érthető tartalommal bír, hiszen a fény frekvenciája egész egyszerűen a színe! (Ilyen értelemben a prizma, mely a ráejtett fényt „szétbontja” összetevő színeire, igazából nem más, mint egy spektrális elemzés fizikai implementációja!) Ismert, hogy a fehér fény az, amely azonosan tartalmaz minden színt; a fenti értelemben tehát a fehér fény spektruma egy vízszintes egyenes. A két dolog ott kapcsolódik össze, hogy belátható, hogy a fenti definiált idősor spektruma szintén egy vízszintes egyenes lesz – ilyen értelemben tehát a színek közül a fehér idősoros megfelelője.

5.3.2. MA-modellek

A mozgóátlagú (MA) modell

A q -ad rendű mozgóátlagú modell (modell, mivel most a sokaságban specifikáljuk):

$$Y_t = \alpha + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q},$$

ahol u_t hiba(folyamat), szokták itt úgy is hívni, hogy innováció, fehérzaj-folyamatnak tételezzük fel: $u_t \sim \mathcal{WN}(0, \sigma_u^2)$; α és θ_i -k valós, σ_u^2 pozitív valós konstans paraméterek

MA(1)-folyamat: várhatóérték-függvény

Közvetlenül a definíció alapján (a várhatóérték-képzést „ráeresztve” a definícióra):

$$\mu_t = \alpha + 0 + \theta_1 \cdot 0 = \alpha,$$

tehát μ_t időfüggetlen

MA(1)-folyamat: szórásnégyzet-függvény

Közvetlenül a definíció alapján (a szórásnégyzet-képzést „ráeresztve” a definícióra):

$$\sigma_t^2 = 0 + \sigma_u^2 + \theta_1^2 \sigma_u^2 = \sigma_u^2 (1 + \theta_1^2)$$

tehát σ_t^2 időfüggetlen (ez lesz γ_0)

MA(1)-folyamat: autokovariancia-függvény

Közvetlenül a definíció alapján (a kovariancia-képzést „ráeresztve” a definícióra):

$$\text{cov}(Y_t, Y_{t-1}) = \text{cov}(\alpha + u_t + \theta_1 u_{t-1}, \alpha + u_{t-1} + \theta_1 u_{t-2}) = \dots$$

összesen 9 tag, ebből azonban csak 1 nem-nulla (a többiben vagy konstans van, vagy különböző időpontokhoz tartozó u -k érintkeznek):

$$\dots = \text{cov}(\theta_1 u_{t-1}, u_{t-1}) = \theta_1 \sigma_u^2,$$

tehát ez időfüggetlen, jogos a γ_1 jelölés

MA(1)-folyamat: autokovariancia-függvény

Közvetlenül a definíció alapján (a kovariancia-képzést „ráeresztve” a definícióra):

$$\text{cov}(Y_t, Y_{t-k}) = \text{cov}(\alpha + u_t + \theta_1 u_{t-1}, \alpha + u_{t-k} + \theta_1 u_{t-k-1}),$$

amiben immár – az előbbi logikát követve – mindegyik tag nulla ha $k > 1$.

Összefoglalva:

$$\gamma_k = \begin{cases} \sigma_u^2 (1 + \theta_1^2) & \text{ha } k=0 \\ \theta_1 \sigma_u^2 & \text{ha } k=1 \\ 0 & \text{ha } k>1 \end{cases}$$

MA(1)-folyamat: stacionaritás

Az előbbieket összerakva (μ_t időfüggetlen, σ_t^2 időfüggetlen, γ_k csak késleltetéstől függ) tehát kapjuk, hogy az MA(1) folyamat stacioner.

Mégpedig *mindig* az (értsd: paraméter-választástól függetlenül).

MA(1)-folyamat: korrelogram

- ACF: $\rho_k = \frac{\gamma_k}{\gamma_0}$; eltűnik 1 késleltetés után
- PACF: belátható, hogy lecsengő (azaz $\lim_{k \rightarrow \infty} \text{PACF}(k) = 0$)

MA(q)-folyamatok

- $\mu_t = \alpha$ (ugyanazért)
- $\sigma_t^2 = \sigma_u^2 (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)$ (ugyanazért)
- ACF q késleltetés után eltűnő (ugyanazért)
- PACF lecsengő (ugyanúgy kiszámolható lenne), adott esetben bonyolultabb min-tázat szerint
- Mindig stacioner (paraméter-választástól függetlenül)!

5.3.3. AR-modellek**Az autoregresszív (AR) modell**

A p -ed rendű autoregresszív modell (modell, mivel most a sokaságban specifikáljuk):

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + u_t,$$

ahol u_t hiba(folyamat), szokták itt úgy is hívni, hogy innováció, fehérzaj-folyamatnak tételezzük fel: $u_t \sim \mathcal{WN}(0, \sigma_u^2)$; α és ϕ_i -k valós, σ_u^2 pozitív valós konstans paraméterek

Megjegyzések

- Speciális esetek: ha $p = 1$ és $\phi_1 = 1$, akkor RWD (ha ráadásul $\alpha = 0$ akkor RW)
- Stacionaritás: ez – szemben az MA-folyamatokkal – nyilván nem lehet *mindig* stacioner, hiszen az RW sem az, de *néha* lehet az is (pl. $p = 1$ és $\phi_1 = 0$), a stacionaritásnak tehát itt valamilyen – paraméterekre vonatkozó – feltétele kell legyen
- Ennek vizsgálatát későbbre halasztjuk, és most azt mondjuk, hogy teljesültek ezek a feltételek

AR(1) folyamat: várhatóérték-függvény

Vegyük mindkét oldal várhatóértékét (feltettük a stacionaritást, $\mathbb{E}(Y_t) \equiv \mu$)

$$\mu = \alpha + \phi_1 \mu + 0,$$

mivel a várhatóérték lineáris, innen

$$\mu = \frac{\alpha}{1 - \phi_1}$$

($\phi_1 \neq 1$ kell legyen: látni fogjuk, hogy ez tényleg fennáll, ha stacioner)

AR(1) folyamat: szórásnégyzet-függvény

Vegyük mindkét oldal szórásnégyzetét (feltettük a stacionaritást, $\mathbb{D}^2(Y_t) \equiv \sigma^2$)

$$\sigma^2 = \phi_1^2 \sigma^2 + \sigma_u^2,$$

kihasználva, hogy a három tag korrelálatlan, innen

$$\sigma^2 = \frac{\sigma_u^2}{1 - \phi_1^2}$$

($|\phi_1| < 1$ kell legyen: látni fogjuk, hogy ez tényleg fennáll, ha stacioner)

AR(1) folyamat: autokovariancia-függvény

Kezdjük az 1 késleltetéssel (természetesen a stacionaritást most is feltételezzük):

$$\begin{aligned} \text{cov}(Y_t, Y_{t-1}) &= \text{cov}(\alpha + \phi_1 Y_{t-1} + u_t, Y_{t-1}) = \\ &= 0 + \phi_1 \text{cov}(Y_{t-1}, Y_{t-1}) + 0 = \phi_1 \sigma^2 = \phi_1 \frac{\sigma_u^2}{1 - \phi_1^2}, \end{aligned}$$

időfüggetlen; innen rekurzívan mehetünk tovább:

$$\text{cov}(Y_t, Y_{t-k}) = \text{cov}(\alpha + \phi_1 Y_{t-1} + u_t, Y_{t-k}) = \phi_1 \gamma_{k-1},$$

szintén időfüggetlen, ezekből tehát indukcióval kapjuk, hogy

$$\gamma_k = \phi_1^k \sigma^2 = \frac{\phi_1^k \sigma_u^2}{1 - \phi_1^2}$$

($|\phi_1| < 1$ kell legyen: látni fogjuk, hogy ez tényleg fennáll, ha stacioner)

AR(1) folyamat: autokorreláció és parciális autokorreláció-függvény

Definíció alapján az autokovariancia-függvényből (mivel stacioner):

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k,$$

tehát az ACF *geometriailag lecsengő*

Külön kellene igazolni, de a mechanika alapján is elég nyilvánvaló, hogy

$$\text{PACF}(k) = \begin{cases} \rho_1 & \text{ha } k=1 \\ 0 & \text{ha } k>1 \end{cases}$$

Épp az MA(1) „fordítva”: a kettő korrelogramja egymás *duálisa*

AR(1) folyamatok RWD-nél látott rekurzív visszafejtése

$$\begin{aligned} Y_t &= \alpha + \phi_1 Y_{t-1} + u_t = \alpha + \phi_1 (\alpha + \phi_1 Y_{t-2} + u_{t-1}) + u_t = \\ &= \alpha + \phi_1 \alpha + \phi_1^2 (\alpha + \phi_1 Y_{t-3} + u_{t-2}) + \phi_1 u_{t-1} + u_t = \dots \end{aligned}$$

Ha feltételezzük, hogy „végtelenből jön” a folyamat (ekkor a kezdőérték mindegy lesz), akkor ez

$$\dots = \alpha \sum_{i=0}^{\infty} \phi_1^i + \sum_{i=0}^{\infty} \phi_1^i u_{t-i} = \frac{\alpha}{1 - \phi_1} + \sum_{i=0}^{\infty} \phi_1^i u_{t-i}$$

($|\phi_1| < 1$ kell legyen: látni fogjuk, hogy ez tényleg fennáll, ha stacioner)

Mint egy MA-modell: ez az $AR(1)$ modell $MA(\infty)$ -reprezentációja

AR(p) folyamatok

- Stacionaritást egyelőre itt is feltételezzük
- Várhatóérték-függvény: $\mu = \frac{\alpha}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$
- Szórásnégyzet-függvény bonyolultabb (az autokovarianciák is megjelennek benne)
- Az ACF lecsengő (végtelenben 0-ba tart), de már bonyolultabb mintázat szerint is lehet
- PACF-ből az első p nem-nulla, a többi viszont már nulla
- Tehát az azonos rendű AR és MA folyamatok korrelogramja általánosságban is egymás duálisa
- Van $MA(\infty)$ -reprezentációja

5.3.4. ARMA-modellek**Az autoregresszív-mozgóátlagú (ARMA) modell**

A p, q rendű autoregresszív-mozgóátlagú modell (modell, mivel most a sokaságban specifikáljuk):

$$\begin{aligned} Y_t &= \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \\ &+ u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q}, \end{aligned}$$

ahol u_t hiba(folyamat), szokták itt úgy is hívni, hogy innováció, fehérzaj-folyamatnak tételezzük fel: $u_t \sim \mathcal{WN}(0, \sigma_u^2)$; α , ϕ_i -k és θ_i -k valós, σ_u^2 pozitív valós konstans paraméterek.

Tulajdonságok

- Stacionaritásnak feltétele van (ami csak az AR együtthatóktól függ)
- Ha fennáll, akkor mind az ACF, mind a PACF lecsengő (nem eltűnő), lehet, hogy bonyolultabb mintázat szerint
- Van $MA(\infty)$ -reprezentációja

6 Késleltetési operátor és polinom, ARMA-folyamatok felírása késleltetési polinommal, az ARMA-folyamatok stacionaritása

6.1. Matematikai emlékeztető

6.1.1. Algebra emlékeztető

Változó, hatvány, polinom, polinom gyöke és inverze

- Legyen x egy változó, x^k egy hatványa, ekkor $\omega_0 + \omega_1 x + \omega_2 x^2 + \dots + \omega_k x^k = \omega(x)$ egy k -ad fokú, egyváltozós – polinom, $\omega_0, \omega_1, \omega_2, \dots, \omega_k$ együtthatókkal
- (Az együtthatók és a változó értéke legegyszerűbb esetben valós számok, de ez nem szükségszerű)
- Megengedjük, hogy a fokszám végtelen is lehessen: $\omega(x) = \sum_{i=0}^{\infty} \omega_i x^i$
- Polinom inverze: $\omega^{-1}(x)$ olyan, hogy $\omega^{-1}(x)\omega(x) = 1$
- Polinom gyöke: az $\omega(x) = 0$ egyenlet megoldása
- Az algebra alaptétele: egy k -ad fokú valós polinomnak k darab – nem feltétlenül különböző – gyöke van, melyek vagy valósak, vagy ha komplexek, akkor konjugált párokban jönnek
- Az előbbi miatt egy polinom mindig felírható úgy – gyöktényezős alak – mint $\omega(x) = \left(1 - \frac{1}{r_1}x\right) \left(1 - \frac{1}{r_2}x\right) \dots \left(1 - \frac{1}{r_k}x\right)$, ahol r_i az i -edik gyök

Polinom invertálása

- Például $1 - ax$ inverze $1 + ax + (ax)^2 + (ax)^3 + \dots$ (egyesével egyeztetve az együtthatókat)
- Ez egy hatványsor, konvergál, ha $|ax| < 1$

- Általános esethez induljunk ki a gyöktényezőes alakból:

$$\begin{aligned}\omega^{-1}(x) &= \left[\left(1 - \frac{1}{r_1}x\right) \left(1 - \frac{1}{r_2}x\right) \cdots \left(1 - \frac{1}{r_k}x\right) \right]^{-1} = \\ &= \left(1 - \frac{1}{r_1}x\right)^{-1} \left(1 - \frac{1}{r_2}x\right)^{-1} \cdots \left(1 - \frac{1}{r_k}x\right)^{-1} = \\ &= \prod_{i=1}^k \left[1 + \frac{1}{r_i}x + \left(\frac{1}{r_i}x\right)^2 + \left(\frac{1}{r_i}x\right)^3 + \dots \right]\end{aligned}$$

- Ami konvergál, ha minden i -re $\left|\frac{1}{r_i}x\right| < 1$
 - Ha $|x| = 1$, akkor a feltétel, hogy $|r_i| > 1$, azaz, hogy mindegyik gyök 1-nél nagyobb abszolútértékű legyen, más szóval, hogy a komplex egységkörön kívül legyen (ugye a gyökök komplexek is lehetnek)

6.2. Az ARMA-folyamatok mélyebb matematikája

6.2.1. A késleltetési operátor és a késleltetési polinom

A késleltetési operátor

- Legyen L valami, ami idősorból egy másik idősort csinál (ha y az eredeti idősor, akkor Ly jelöli az újat)
- ...mégpedig úgy, hogy $(Ly)_t = y_{t-1}$
- Az egyszerűség kedvéért most fókuszáljunk a minta (realizálódott) idősorra, ne a sokasági szemléletre
- Fogjuk fel úgy, mint egy függvényt, ami az időkhöz értékeket rendel: $y : \{1, 2, \dots, T\} \rightarrow \mathbb{R}$ és $y : t \mapsto y_t$
- Az L tehát függvényből egy másik függvényt csinál: *operátor*

A késleltetési operátor (precízebben)

- A „függvény” itt igazából egy vektor ($\mathbf{y} = (y_1 \ y_2 \ \cdots \ y_T)^T$)
- (Ez rendben is van: egy n dimenziós valós vektor felfogható egy $\{1, 2, \dots, n\} \rightarrow \mathbb{R}$ függvényként!)
- Ezek a függvények egy vektorteret alkotnak (függvény: a vektortér eleme, skálárral szorzás: pontonként szorzás, összeadás: pontonkénti összeadás), ezt szokás függvénytérnek nevezni

- A fenti esetben ez megfelel az n -dimenziós valós vektorokkal végzett szokásos műveleteknek
- Az operátor – úgy általában – igazából két vektortér közti leképezés
- A függvényteres értelmezés miatt mondhattuk azt, hogy az „operátor az, ami függvényből másik függvényt csinál”!

A késleltetési operátor (precízebben)

- Ha a vektoros felfogást, és azon belül is az n -dimenziós valós vektoroknak való megfeleltetést vesszük, akkor minden operátor reprezentálható mátrixszal (hiszen a mátrix az, ami vektorból vektort csinál!)
- Ez alól a késleltetési operátor sem kivétel, például:

$$\underbrace{\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\mathbf{L}} \cdot \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}}_{\mathbf{y}} = \begin{pmatrix} 0 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix},$$

- (Azért, hogy ne változzon az idősor hossza, az új első eleme legyen fixen 0)

A késleltetési operátor hatványai

- Micsoda L^2 ?
- Könnyen értelmezhető: $L(Ly_t) = Ly_{t-1} = y_{t-2}$
- Röviden: $L^2y_t = y_{t-2}$
- Megfeleltethető a mátrixoknak? Igen! Az \mathbf{L} mátrix négyzete épp a kettővel késleltetést valósítja meg, azaz $\mathbf{L}^2 = L^2$
- Szorozzuk össze, és ellenőrizzük le, hogy ez csakugyan teljesül!
- Hasonlóan $L^k y_t = y_{t-k}$, tehát ez a k -val késleltető operátor lesz
- (Ideértve azt is, hogy például $L^{-1}y_t = y_{t+1}$, „siettető operátor”)

A késleltetési polinom

- A késleltetett idősorokat kombinálhatjuk is, például $2y_t + 3y_{t-1} - 4y_{t-2} = 2y_t + 3Ly_t - 4L^2y_t = \dots$
- Most jön az érdekes rész: ez átírható mint

$$\dots = (2 + 3L - 4L^2) y_t$$

6 Késleltetési operátor és polinom

- Ami fontos, hogy ez nem „szintaktikai manipuláció”, az előbbi mátrixok nagyon is mutatják ennek a realitását: $2\mathbf{I} + 3\mathbf{L} - 4\mathbf{L}^2$ épp az a mátrix, amivel rászorozva az idősorra pont $2y_t + 3y_{t-1} - 4y_{t-2}$ -t kapjuk!
- Ennek általánosítása a késleltetési polinom:

$$\omega(L) = \omega_0 + \omega_1 L + \omega_2 L^2 + \dots + \omega_k L^k,$$

azaz az operátorokból is ugyanúgy gyárthatunk polinomot – az előbb definiált hatványaik segítségével – mint mondjuk valós ismeretlenekből

- Ezzel $\omega(L) y_t = \omega_0 y_t + \omega_1 y_{t-1} + \omega_2 y_{t-2} + \dots + \omega_k y_{t-k}$
- Természetesen $\omega(L)$ maga is egy operátor

A késleltetési polinom használatának előnye

Számos – egyébként bonyolult – művelet elvégezhető, mint (jól ismert) manipuláció polinomokkal: összeszorozhatóak, invertálhatóak stb.!

6.2.2. ARMA-folyamatok reprezentációja késleltetési polinomokkal

ARMA-folyamatok felírása késleltetési polinomokkal

- Emlékeztetőül:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q}$$

- Kicsit átrendezve:

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \dots - \phi_p Y_{t-p} = \alpha + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q}$$

ARMA-folyamatok felírása késleltetési polinomokkal

- Az előbbiek alapján ez átírható mint

$$Y_t - \phi_1 L Y_t - \phi_2 L^2 Y_t - \dots - \phi_p L^p Y_t = \alpha + u_t + \theta_1 L u_t + \theta_2 L^2 u_t + \dots + \theta_q L^q u_t$$

- Azaz:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) Y_t = \alpha + (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) u_t$$

- Vezessünk be két késleltetési polinomot: $\phi(x) = 1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p$ és $\theta(x) = 1 + \theta_1 x + \theta_2 x^2 + \dots + \theta_q x^q$
- Ezekkel az előbbi egész egyszerűen

$$\phi(L) Y_t = \alpha + \theta(L) u_t$$

ARMA-folyamatok vizsgálata polinomiális reprezentációval: stacionaritás

- Az előbbi egyenletet átalakítva:

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

- Legalábbis, ha $\phi(L)$ invertálható!
- Ehhez az kell, hogy a gyökei a polinomnak az egységkörön kívül legyenek
 - Mert az L úgy viselkedik, mint az 1 abszolútértékű szám (1 az operátornormája)
- Lényegében azt jelenti, hogy létezik $MA(\infty)$ -reprezentáció
- És most jön a lényeg: ez épp a *stacionaritás* feltétele!
- (Persze ez bizonyítást igényel)

ARMA-folyamatok vizsgálata polinomiális reprezentációval: invertálhatóság

- Ha viszont $\theta(L)$ gyökei vannak az egységkörön kívül, akkor az egész $AR(\infty)$ -folyamatként reprezentálható
- Ilyenkor azt mondjuk, hogy a folyamat *invertálható*
- Lényegében azt jelenti, hogy az u_t is felírható Y aktuális és múltbeli értékeivel (nem csak fordítva)

7 A stacionaritás tesztelése

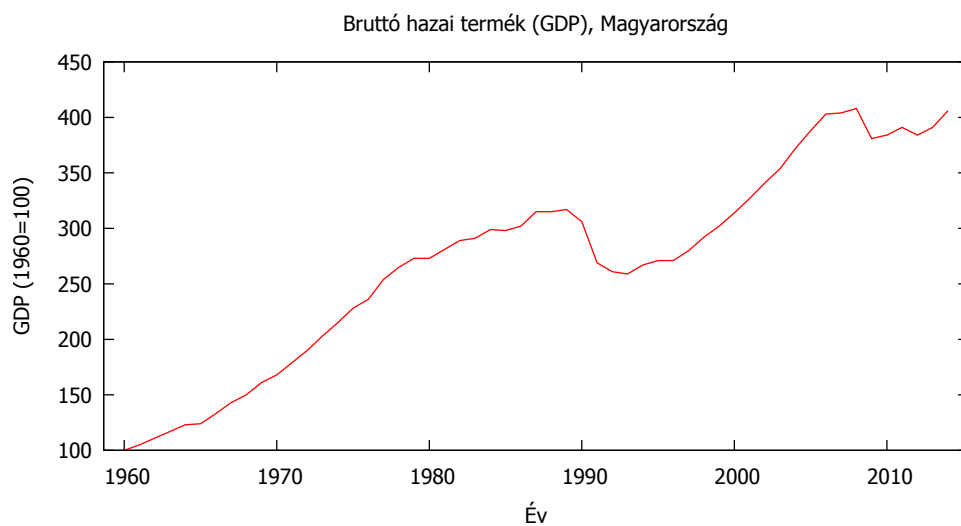
7.1. A stacionaritás tesztelése

7.1.1. A stacionaritás teszteléséről általában

A tesztelés lehetőségei

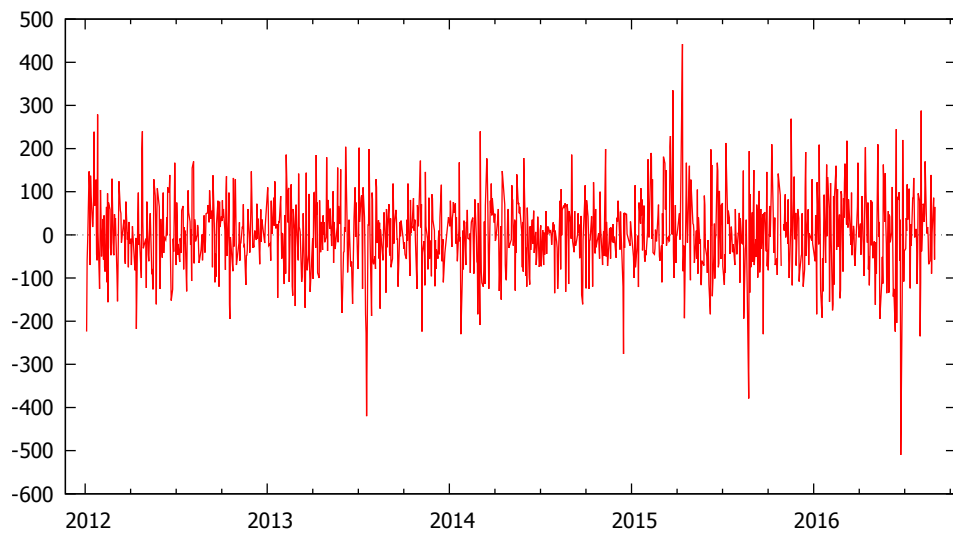
- Egy módszerről már volt szó (grafikus eljárás), de ez igen szubjektív
- Most kiegészítjük két újjal, a második, a korrelogram szemrevételezése még mindig inkább csak heurisztikus...
- ...de a harmadik, a statisztikai próbák alkalmazása már objektív (noha ez nem azt jelenti, hogy tökéletes!)

Grafikus módszer



Grafikus módszer

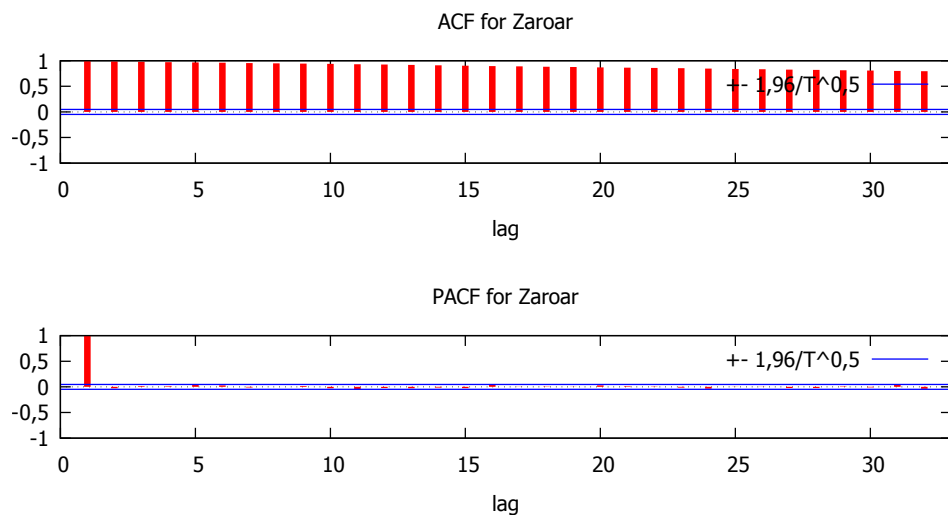
7 A stacionaritás tesztelése



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 növekedésével)

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

7.1.2. Egységgyök

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 egységkörön 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!)

7 A stacionaritás tesztelése

- 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 $\phi(x) = \tilde{\phi}(x)(1-x)$, ahol $\tilde{\phi}(x)$ -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

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

azaz

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

- Vagyis ilyenkor az idősor differenciázottjára adtunk egy ARMA-modellt!
- Egész pontosan ARMA(p-1,q)-t, hiszen az AR-polinomja ($\tilde{\phi}(L)$) 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 expozív idősorokat teljesen kizárjuk a vizsgáldásunk 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 expozitivitás 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 mindegyikéhez):
 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

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

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)

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

8.1. 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 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 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 kvadratikust $(\alpha + \beta t + \gamma t^2)$, 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 paramétereit 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 **differenciálása**, jele Δ , ez is egy *operátor*:

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

tehát $\Delta = 1 - L$

Differenciálás

- 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 differenciálás 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 $I(1)$

Differenciálás

- A differenciálás 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 differenciálás* (*másodrendű differenciálás*) 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, $\Delta^d = (1 - L)^d$

8 A nemstacionaritás kezelése

- 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

9 Box-Jenkins eljárás, előrejelzés készítése

9.1. Box-Jenkins eljárás

A Box-Jenkins eljárás lényege

- Az alapgondolat: az idősorokat – stationer – ARIMA(p,d,q)-modellel írjuk le...
- ...a paramétereiket úgy megválasztva, hogy a modellfeltevések teljesüljenek
- A nevét két fő proponentéről – George Box és Gwilym Jenkins – kapta, akiknek az 1970-es könyve (Time Series Analysis: Forecasting and Control) nagyon sokat tett a módszer széles körben történő megismertetéséért és elterjesztéséért

A Box-Jenkins eljárás lépései

1. A d meghatározása: már láttuk a módszereit (lényegében stacionarizálás/stacionaritás tesztelése)
2. A p és q rendek behatárolása: azért nem „meghatározása”, mert jellemzően nem egyértelmű, többféle lehetőséggel is próbálkozni kell (de általában igyekszünk kicsin tartani ezeket), egyedül a korrelogram segíthet, ha szemrevételezzük és összevetjük azzal, hogy az elméleti korrelogramok hogyan néznek ki (de ez általában csak tiszta AR vagy MA modelleknél működőképes)
3. Modell becslése: technikai lépés, most nem foglalkozunk vele
4. Modelldiagnosztika: reziduumok vizsgálata, minimum autokorrelálatlanságra (korrelogram, Ljung-Box teszt, Breusch-Godfrey teszt), esetleg normalitásra
5. Modellminősítés: jellemzően információs kritériumokat (AIC, BIC (SBC), HQC) használunk

A Box-Jenkins eljárás lépései

- A p és q behatárolásához tehát lényegében egy kétlépcsős megoldást alkalmazunk:
 - Szűrés: ami diagnosztikailag nem megfelelő, azok a modellek szóba sem jöhetnek, kidobjuk őket a jelöltek listájáról (ez tehát a modelldiagnosztika alapján megy)

9 Box-Jenkins eljárás, előrejelzés készítése

- Sorbarakás: ha nem egyetlen modell marad fenn, akkor azokat sorbarakjuk, és a – valamely metrika szerinti – legjobbat választjuk (ez tehát a modellminőség alapján megy)
- Az így kapott modellt pedig felhasználjuk
- Itt jellemzően a felhasználás nem elemzést, hanem előrejelzést jelent

9.2. Előrejelzés készítése

Az előrejelzés alapelve

- Természetesen itt is feltételes várható értékkel predikálunk, azaz az előrejelzéshez behelyettesítünk minden ismert változót (ARMA-modellnél ez a folyamat múltbeli értékeit, és a múltbeli hibákat jelenti), és a tárgyidőszaki hibatagot nullának vesszük
- Ilyen módon ARMA-modellben csak egyetlen időszakra tudunk előrejelezni; ennek neve **statikus előrejelzés**
- Statikus előrejelzésben csak realizálódott értékre támaszkodunk (a tárgyidőszaki hibától eltekintve, természetesen)
- Ha több időszakra kell előrejeleznünk, akkor
 - a későbbi hibákat mind nullának kell vennünk (nem csak a tárgyidőszakit)
 - a múltbeli értékek sem lesznek mind realizálódottak – ilyenkor a korábbi előrejelzésre támaszkodunk
- Ezt hívjuk **dinamikus előrejelzésnek**

Előrejelzések készítése

- Mindez összefoglalva azt jelenti, hogy
 - a hibatag helyébe a reziduumot írjuk, ha mintán belül vagyunk, 0-t, ha azon kívül
 - a múltbeli érték helyébe a realizálódott értékét írjuk, ha mintán belül vagyunk, a becsült értéket, ha azon kívül
- Az ARMA-folyamat tulajdonságaiból adódik, hogy nagyon messzire előremenne az előrejelzéssel a folyamat várható értékéhez fogunk konvergálni
- ARIMA-modellezésnél utolsó lépésben még vissza kell csinálni a differenciázást (kumulálni kell)

Az előrejelzés pontosságának a mérése

A két legtipikusabb mutató:

- Átlagos négyzetes hiba: $MSE = \frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2$
- Átlagos abszolút relatív hiba: $MAPE = \frac{1}{T} \sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{y_t}$

10 Idősorok regressziója

10.1. Exogén változós idősormodellek

10.1.1. Alap gondolatok, statikus regresszió

Idősorok regressziójának alap gondolata

- Az idősorunkat *más* idősor(ok)kal akarjuk magyarázni
- Lényegében tehát *ki akarjuk regresszálni* az idősorunkat (mint eredményváltozót), más idősorokkal (mint magyarázó változókkal)
- Bizonyos értelemben az eddigi AR-modellek is ilyenek voltak, csak a „más idősor” ugyanannak a késleltetettjeit jelentette
- Most viszont megengedjük, hogy tényleg eltérő idősorok (vagy azok késleltetettjei!) is belépjenek magyarázó változóként
- A fő kérdésünk az lesz, hogy e modelleknek milyen feltételeket kell teljesíteniük, hogy jó tulajdonsággal becsülhetőek legyenek a paramétereik
- És persze szokásosan az ökonometriai modellek két felhasználása: elemzés és előrejelzés

Statikus regresszió

- A legegyszerűbb idősoros regressziós modell:

$$y_t = \beta_0 + \beta_1 z_t + u_t,$$

ahol y_t és z_t tehát két idősor *ugyanazon* időpontbeli megfigyelései

- Például: statikus Phillips-görbe (infláció vs. munkanélküliség)
- Az u_t hibatag és tulajdonságai lesznek majd vizsgáldásunk fókuszában, ami a modellfeltevéseket illeti
- Legegyszerűbb eset, ha z_t valami „teljesen más”, u_t -től külső információ (ezt majd pontosítjuk), úgy fogjuk mondani, hogy **exogén változó**
- Természetesen lehet több exogén változó is:

$$y_t = \alpha + \beta_1 z_{t1} + \beta_2 z_{t2} + \dots + \beta_k z_{tk} + u_t$$

Statikus regresszió

- β_i jelentése: ha az i -edik idősor értéke egy adott időszakban egy egységgel megnő (minden mászt változatlanul tartva), akkor modellünk szerint várhatóan hány egységgel lesz nagyobb az eredményváltozó *ugyanazon* időszakban
- Azért hívjuk statikusnak, mert ugyanazon időszaki változásokat kapcsol össze, tehát nincs időszakok közötti hatás
- (A dinamika szó általában is időben kiterjedten lezajló dolgokra utal)

10.1.2. Dinamikus regressziók

A dinamika szükségessége

- Rengeteg helyzetben a statikusság irreális
- Nem várható, hogy egy beruházás rögtön *ugyanabban* az időpontban befolyásolja a kibocsátást, hogy egy felvilágosító kampány *azonnal* lecsökkenti a megbetegedések számát, hogy egy szociálpolitikai intézkedés *rögtön* megváltoztatja a jövedelmi viszonyokat stb.
- A legtöbb társadalmi-gazdasági jelenség csak időben elnyújtva, késleltetéssel hat, több időszakon keresztül fejt ki a hatását
- Ezért szükséges a dinamika beépítése is a regressziós modelljeinkbe

Osztott késleltetésű modellek

- A legegyszerűbb dinamikus modell: legyen most csak egyetlen magyarázó változónk, csak épp

$$y_t = \beta_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + \dots + \delta_p z_{t-p} + u_t$$

- Az idősor tárgyidőszaki értékére a korábbi z -k is hatást gyakorolnak...
- ... avagy – fordítva elmondva ugyanazt – a z mostani változásai a jövőben fognak kihatni y -ra
- β_i : ha z most megváltozik, akkor i időszakkal később ez hogyan hat y -ra
- Amennyiben véges sok korábbi z hat y -ra (később fogjuk látni hogyan hathat végtelen sok korábbi), akkor **véges osztott késleltetésű** modelltől (FDL, finite distributed lag) beszélünk

Az FDL-modell

- Ha z konstans, majd egy időszakra felugrik eggyel nagyobbra, majd után visszaáll a konstans szintre, akkor a tárgyidőszaki y β_0 -al lesz nagyobb mint az állandósult szintje, az eggyel később y β_1 -gyel, ..., a p -vel későbbi időszakban β_p -val, és a p utáni időszakokra már nem hat ez a módosulás
- Ezeket hívjuk **rövid távú hatásmultiplikátornak**, értéke tehát az i -edik időszakra épp β_i
- Éppen ezért szokás kiplotolni β_i -t a i -vel szemben
- A másik tipikus értelmezési keret, hogy z egy adott időszakban felugrik eggyel és *úgy is marad*, kérdés, hogy hosszú távon mi történik y -nal
- Ugyanabban az időszakban β_0 -al nő meg, a következőben $\beta_0 + \beta_1$ -gyel és így tovább
- Ezt hívjuk **hosszú távú hatásmultiplikátornak**, értéke tehát $\sum_{i=0}^p \beta_i$

FDL-modell strukturálatlan becslése

- Ha a fenti módon egyszerűen megbecsüljük β -kat, akkor **strukturálatlan becslésről** beszélünk (mert semmit nem tettünk fel a β -k értékeiről)
- A probléma ezzel, hogy nagyon sok esetben egy idősor egymást követő értékei nagyon korreláltak \rightarrow a fenti modellben rendkívül erős multikollinearitás lesz, a β -kat csak nagyon bizonytalanul (hatalmas CI-vel) tudjuk csak becsülni
- (*Együttesen* vizsgálhatóak, például F -teszttel, vagy a hosszú távú hatásmultiplikátort is jól meg tudjuk becsülni, csak külön-külön nem)

FDL-modell struktrált becslése, Almon-lag

- Éppen ezért gyakori, hogy nem teljesen szabadon becsüljük β_i -ket, hanem feltételezünk valamilyen struktúrát
- Lényegében: átcseréljük az eredeti paramétereket kisebb számú, kevésbé multikollineáris paraméterekre
- (Persze ennek az az ára, hogy a struktúrát el kell találnunk, az ugyanis nem az adatokból jön, hanem mi mondjuk meg kívülről)
- Az egyik népszerű választás az **Almon késleltetési struktúra**, amikor azt feltételezzük fel, hogy a β_i -k az i -ben polinomiálisak:

$$\beta_i = \sum_{j=0}^n w_j i^j,$$

ahol n tipikusan kicsi (pl. 2-3)

10 Idősorok regressziója

- Akármennyi is p , nekünk csak n darab – általában már nem túl multikollineáris – paramétert kell becsülnünk
- De még egyszer: fontos, hogy β_i tényleg kvadratikus (/kübös/stb.) legyen i -ben

FDL-modell strukturált becslése, Koyck-lag (GDL)

- Egy másik népszerű választás, hogy β_i geometriailag lecsengő i -ben:

$$\beta_i = \beta_0 \rho^i,$$

ahol természetesen $|\rho| < 1$

- (Azt is mondhattuk volna, hogy $\beta_i = \rho \beta_{i-1}$)
- Ezt hívják **Koyck késleltetési struktúrának**
- Ami nagyon érdekes, hogy ehhez igazából az sem kell, hogy csak véges sok késleltetés lépjen be a modellbe!
- Nyugodtan lehet az a modellünk, hogy

$$y_t = \beta_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + \dots + u_t,$$

nem lesz végtelen sok becsülendő paraméterünk, hiszen a DL részhez tartozó paraméterek száma *mindenképp* 2 (β_0 és ρ)

- Tehát értelmesen megbecsülhető a fenti specifikáció is, mintegy végtelen osztott késleltetésű modellként, a neve **geometriai osztott késleltetű** modell (GDL, geometric distributed lag)

A GDL-modell értelmezése

- A rövid távú hatásmultiplikátor tehát $\beta_i = \beta_0 \rho^i$
- A hosszú távú hatásmultiplikátor izgalmasabb:

$$\sum_{i=0}^{\infty} \beta_i = \sum_{i=0}^{\infty} \beta_0 \rho^i = \beta_0 \sum_{i=0}^{\infty} \rho^i = \frac{\beta_0}{1 - \rho}$$

10.1.3. Idősoros regressziók általános modellje

Az eddigiek kombinációja

- Az eddigiek természetesen kombinálhatóak is: lehet benne *több* z is, akár késleltetve
- Természetesen lehet vegyesen is (bizonyosak késleltetés nélkül, mások késleltetéssel, a rend sem kell, hogy azonos legyen)
- A jobb oldalra berakhatjuk az eredményváltozó késleltettjeit is (itt nyilván egyidejű tagot nem rakhatunk be...), ezzel AR-hatást is létrehozhatunk

Egy általános modell felé

- Láttuk tehát, hogy a magyarázó változó lehet:
 - Exogén z (mint a statikus regresszióban)
 - Exogén z késleltetettje (mint a DL-ben)
 - Az eredményváltozó késleltetettje (mint az AR-ben)
- Külön-külön mindegyiket néztük már – lásd a zárójeles megjegyzéseket – de semmi akadálya, hogy többet (vagy akár az összeset egyszerre) berakjuk egy modellbe!

Az általános modell

- Mindent összetéve:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + u_t = \beta_0 + \boldsymbol{\beta}^T \mathbf{x}_t + u_t,$$

ahol $x_{t,i}$ egyaránt *lehet* exogén változó, késleltett exogén változó vagy késleltetett eredményváltozó

- Formailag teljesen olyan, mint a regresszió keresztmetszetben, van eredményváltozó és vannak magyarázóváltozók
- Egyszerűen behúzzuk őket egy modellbe – történetesen nem keresztmetszeti adatok, hanem idősorok, de hát az OLS-nek mindegy, számok vannak így is, úgy is – és simán megbecsüljük OLS-sel... jó ötlet ez?
- A következőkben ezzel fogunk foglalkozni: mi történik akkor, ha a β -kat egyszerűen megbecsüljük OLS-sel, milyen feltételek mellett milyen tulajdonságúak lesznek az így kapott becslések?

10.2. Idősoros regresszió becslése OLS-sel

10.2.1. Standard modellfeltevések

A modellfeltevések és szerepük

- A helyzet, és a kérdés teljesen analóg a keresztmetszetenél látottakkal: milyen modellfeltevések mellett garantálhatóak, hogy az OLS szolgáltatja becsléseknek jó tulajdonságaik legyenek?
- Úgy fogjuk végignézni, hogy mindenhol a keresztmetszettel rakjuk párhuzamba
- Ugyanúgy 5 (+1) modellfeltevés lesz

Linearitás

- Keresztmetszetnél ez volt:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + u$$

és ez igaz mindegyik megfigyelési egységre, és így az egész mintára is:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + u_i$$

- Idősornál *pontosan ugyanez* a feltétel (legfeljebb i helyett t -t szokás írni)

Nincs egzakt multikollinearitás

- Keresztmetszetnél ez volt:

$$\mathbb{P}(\text{rank } \underline{X} = k) = 1$$

- Idősornál *pontosan ugyanez* a feltétel

Szigorú exogenitás

- Keresztmetszetnél, ha trehányak voltunk, ez volt:

$$\mathbb{E}(u_i | \underline{X}_i) = 0,$$

ha precízek, akkor ez:

$$\mathbb{E}(u_i | \underline{X}) = 0$$

- Idősornál, ha precízek voltunk, *pontosan ugyanez* a feltétel
- Tehát: *minden* időszaki hiba várható érték független *minden* (akár más időszaki!) magyarázó változótól

Szigorú és egyidejű exogenitás

- A trehányság azért volt megengedhető, mert ha fae a mintavétel – ami keresztmetszetnél egy elfogadható feltevés lehet – akkor a precíz tényleg a trehányra egyszerűsödik
- Ez teljesen logikus: ha a különböző mintaelemek függetlenek, akkor egy adott időszaki hiba az összes többi időszaki magyarázó változótól nyilván várható érték független lesz, tehát csak az ugyanazon időszakiktól való függetlenséget kell megkövetelni
- Idősornál, mivel a fae mintavétel itt már nem elfogadható általánosságban, ez a trehányság nem lesz megengedhető
- A továbbiakban a $\mathbb{E}(u_i | \underline{X}_i) = 0$ feltételt **egyidejű exogenitásnak**, az – erősebb – $\mathbb{E}(u_i | \underline{X}) = 0$ feltételt **szigorú (vagy erős) exogenitásnak** nevezzük
- (Most válik érthetővé, hogy a szigorú exogenitás elnevezésben mit jelent a szigorú!)
- A standard modellfeltevésben tehát a szigorú exogenitás szerepel

A szigorú exogenitás sérülései

- Természetesen minden, amit keresztmetszetenél is láttunk (pl. kihagyott változó, mérési hiba)
- Itt azonban más okok is lehetnek a háttérben:
 - Rosszul megragadott dinamika: például statikus regressziót becslünk, miközben FDL lenne a helyes
 - Az u -beli változás nem befolyásolhatja a későbbi x -et (ez meg hogy lehetne? úgy, ha y értékei visszahatnak a későbbi x -kre, például bűnözés regresszálása a rendőri erők létszámával)

Homoszkedaszticitás

- Keresztmetszetenél ez volt:

$$\sigma_i^2 := \mathbb{D}^2(u_i | \underline{X}) = \sigma^2$$

- Idősornál *pontosan ugyanez* a feltétel

Autokorrelálatlanság

- Keresztmetszetenél ez volt: ha fae a mintavétel, akkor automatikusan teljesül, különben

$$\text{cov}(u_i, u_j | \underline{X}) = 0$$

minden $i, j = 1, 2, \dots, n, i \neq j$

- Idősornál *pontosan ugyanez* (az utóbbi) a feltétel

Hibanormalitás

- Keresztmetszetenél ez volt:

$$\underline{u} | \underline{X} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$$

- Idősornál *pontosan ugyanez* a feltétel

Összefoglalva

- Ha precízen fogalmaztunk, akkor igazából a keresztmetszetenél látott feltételek egy-az-egyben ugyanazok, mint amire itt is szükség van
- Most már elárulható, hogy ez nem véletlen: a precíz fogalmazás *épp* azért kellett, hogy az ott látott dolgok valójában *univerzálisak* legyenek, tehát ne csak keresztmetszetre vonatkozzanak, hanem ugyanúgy idősorra is
- ...ami tulajdonképpen jól érthető is: a tiszta elmélet egységes kell legyen, hiszen a változóknak „mindegy”, hogy ők most idősorok, vagy keresztmetszeti adatok, vagy micsodák

10.2.2. Az OLS véges mintás tulajdonságai idősorokra

Az OLS véges mintás tulajdonságai idősorokra

- Az első három feltétel teljesülése esetén az OLS szolgáltatta becslések torzítatlanok
- Ha mind az öt feltétel teljesül, akkor az OLS szolgáltatta becslések ezen felül hatásosak is (azaz BLUE-k is)
- Ha mind az öt feltétel teljesül, akkor a σ^2 és a hibák kovarianciamátrixának OLS szolgáltatta becslése torzítatlan
- Ha még a hibanormalitás is teljesül, akkor az OLS szolgáltatta becslések eloszlása normális, a t (F) statisztikák nulleloszlásai tényleg t -k (F -ek), a szokásos tesztek és a konfidenciaintervallumok validak

A keresztmetszeti esethez való viszony

- Mindez lényegében azt jelenti, hogy ezen feltevések teljesülése esetén az idősoros adatokkal *pontosan ugyanúgy* hajthatunk végre regressziót, mintha keresztmetszetiek lennének!
- Persze látni kell, hogy ezek rettentő erős feltevések voltak, a gyakorlatban ritkán teljesülnek

Véges (vagy kis-) mintás tulajdonság mivolt

- A tulajdonságoknál nem mondtuk semmit a mintanagyságról: ez azt jelenti, hogy mintanagyságtól függetlenül – azaz minden mintanagyságra – igazak
- Ilyenkor azt szokták mondani, hogy ezek „kis” mintás (véges mintás) tulajdonságok voltak
- (A kismintás elég szerencsétlen elnevezés, hiszen természetesen nagy mintára is igazak, gyakorlati szempontból persze érthető a kifejezés oka)

10.2.3. Az OLS nagymintás tulajdonságai idősorokra

A nagymintás tulajdonság értelme és szükségessége

- Nagymintás: nem minden n -re igaz, hanem csak $\lim_{n \rightarrow \infty}$ értelemben (szokás még aszimptotikus tulajdonságnak is nevezni)
- Fontos, mert a gyakorlatban a véges mintás tulajdonságokhoz tartozó feltételek sokszor nem teljesülnek, de nagy mintát néha van módunk venni, így nagyon lényeges annak vizsgálata, hogy ezzel mit tudunk „kiváltani”
- Igazából már keresztmetszetnél is láttunk egy nagymintás tulajdonságot: amikor azt mondtuk, hogy az első három tulajdonság fennállása esetén az OLS szolgáltatta becslések konzisztensek

Kitérő: idősorok ergodicitása

- Egy idősort **ergodikusk** nevezünk, ha az időben távoli tagjai – bármely időpontból indulva – függetlenbe tartanak az időbeni távolságuk növekedtével (aszimptotikusan függetlenek)
- Egy ergodikus idősorra, ha még stationer is (és így létezik μ) teljesül, hogy

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T Y_i \xrightarrow{\text{m.b.}} \mu$$

- Néha ezzel definiálják az ergodicitást (pontosabban szólva a várható értékben ergodicitást) – ilyenkor a stationaritást meg kell követelni, vagy legalábbis óvatosan eljárni
- (Természetesen mindig definiálhatjuk az $I_{\{Y_i \in A\}}$ idősort, ilyenkor a várható érték valószínűség lesz)

Az ergodicitás – egy lehetséges! – pontos definíciója: minden korlátos $f : \mathbb{R}^{k+1} \rightarrow \mathbb{R}$ és $g : \mathbb{R}^{l+1} \rightarrow \mathbb{R}$ függvényre és minden i -re igaz, hogy

$$\begin{aligned} \lim_{n \rightarrow \infty} |\mathbb{E}[f(y_i, y_{i+1}, \dots, y_{i+k}) g(y_{i+n}, y_{i+n+1}, \dots, y_{i+n+l})]| &= \\ = |\mathbb{E}[f(y_i, y_{i+1}, \dots, y_{i+k})]| |\mathbb{E}[g(y_{i+n}, y_{i+n+1}, \dots, y_{i+n+l})]|. \end{aligned}$$

Az ergodicitás tartalma

- Lényegében azt mondja ki, hogy időátlag tart a sokasági – összességi – átlaghoz:
 - Azért fontos, mert azt mondja, hogy elég sok elemet megfigyelve (az időben – ugye mi csak ezt tudjuk megtenni!) *tényleg* tudunk következtetni a várható értékekre/valószínűségekre (ami igazából érdekel minket!)
 - A nagy számok törvényének megfelelője, illetve általánosítása (nem kellett a teljes függetlenséget feltenni)
- Néha szokás ezt gyenge függőségnek is nevezni

Ergodicitás és az autokovarianciák

- Érezhető, hogy ha egyszer az ergodicitás olyasmit követel meg, hogy az egyre távolabbi értékek egyre függetlenebbek legyenek (a teljes függetlenséghez tartva), akkor összefügg az autokovarianciákkal – hiszen azok is valami függetlenséggel kapcsolatban lévő dolgot mérnek
- Csakugyan, belátható, hogy egy idősor ergodik (a várható értékre), ha a kovarianciái nullába tartanak, mégpedig olyan gyorsan, hogy abszolút összegezhetőek is:

$$\sum_{i=1}^{\infty} |\gamma_i| < \infty$$

Stacionaritás és ergodicitás

- Egy stacioner idősor nem feltétlenül ergodik: $Y_t = X$ (ahol X egy valószínűségi változó), azaz az idősor konstans
- Egy ergodik idősor nem feltétlenül stacioner: $Y_t = \alpha t + u_t$, ahol $\alpha \neq 0$ és $u_t \sim \mathcal{N}(0, \sigma_u^2)$ függetlenül
- Nagyon sok esetben azonban a kettő ugyanaz (néhol keveredés is van emiatt a szóhasználatban)

Az új modellfeltevések

- Pluszban megköveteljük a linearitásnál, hogy az idősorok legyenek stacionerek és ergodikusak is
- Cserében viszont
 - Szigorú (erős) exogenitás helyett elég lesz az egyidejű exogenitás: $\mathbb{E}(u_i | \underline{X}_i) = 0$
 - Szigorú (erős) homoszkedaszticitás helyett elég lesz az egyidejű homoszkedaszticitás: $\mathbb{D}^2(u_i | \underline{X}_i) = \sigma^2$
 - Szigorú (erős) autokorrelálatlanság helyett elég lesz az egyidejű autokorrelálatlanság: $\text{cov}(u_i, u_j | \underline{X}_i, \underline{X}_j) = 0$
- (A hibanormalitásról nem tettünk fel semmit: nem kellett, mert úgyis aszimptotikus eredményeink lesznek, ahol a centrális határeloszlás tétel kihasználható – ugyanis a fenti feltételek mellett az is működni fog, nem csak a nagy számok törvénye)

Az OLS nagymintás tulajdonságai

- Az előbb vázolt modellfeltevések közül az első három teljesülése esetén az OLS szolgáltatott becslések konzisztensek
- Ha mind az öt teljesül, akkor az OLS szolgáltatott becslések aszimptotikusan normálisak, a t (F) statisztikák nulleloszlásai tényleg t -k (F -ek) aszimptotikusan, a szokásos tesztek és a konfidenciaintervallumok aszimptotikusan validak