

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

Lomoschitz Lilla

**TÖBBDIMENZIÓS ARMA FOLYAMATOK
KÜLÖNBÖZŐ FELÍRÁSAI ÉS ILLESZTÉSEI**

Szakdolgozat
Alkalmazott matematikus Msc

Témavezető:
Pröhle Tamás
Valószínűségelméleti és Statisztika Tanszék



Eötvös Loránd Tudományegyetem
Természettudományi Kar

2019

Köszönetnyilvánítás

Szeretném megköszönni Pröhle Tamásnak a sok segítségét, amit a szakdolgozat elkészülése során kaptam tőle. Köszönöm a segítséget az irodalom felkutatásában, az észrevételeket és javaslatokat, valamint a rugalmasságot a konzultációs időpontok megválasztásánál.

Hálás vagyok a családomnak, és a barátaimnak a támogatásukért, valamint a türelemért, amit a szakdolgozat megírása alatt tanúsítottak.

Tartalomjegyzék

Bevezetés	4
1. VARMA Modellek	5
1.1. Ekvivalens reprezentáció, egyértelműség	9
1.2. Végső forma	12
1.3. Echelon forma	13
1.4. Skalár komponens modell	15
2. Hiba javító modellek	17
2.1. Integrált folyamatok	17
2.2. Kointegrált folyamatok	17
2.3. Hiba javító modellek	19
2.4. Hibajavító modellek különböző becslései	19
2.5. Dickey-Fuller teszt	22
3. Kriptoaluták kointegráltságának vizsgálata	23
3.1. Stacionaritás vizsgálata	27
3.2. ECM modell illesztés	28

Bevezetés

Különböző idősor-modelleket széleskörűen alkalmaznak a matematikán kívüli tudományterületek. A makroökonómia a gazdaságra vonatkozó aggregált mérőszámok, változók, például a GDP, a munkanélküliség és az infláció vizsgálatára, előrejelzése használja. Pénzügyben különböző valuták, értékpapírok, és származtatott termékek árváltozásának elemzésére szolgál. A meteorológiában a hőmérséklet mellett a légnyomás, páratartalom, szélsebesség, csapadékmennyiség is idősor jellegű változók, és mivel több helyen mérünk, így még több dimenziós lesz az adatsorunk.

Az adatsorok független vizsgálata történik abban az esetben, ha azokat külön-külön, egy dimenziós idősorokként vizsgáljuk. Többdimenziós idősorként a különböző változók időben eltolt egymásra való hatását is vizsgáljuk. Ilyen lineáris modellek például a többdimenziós VAR, és VARMA modell. Ezeknél jóval bonyolultabb, nem lineáris modellek is vannak. Ez a szakdolgozat többdimenziós lineáris modellekkel foglalkozik: VARMA folyamatokkal, melyek két részből részből tevődnek össze: az egyik az ismert változókból álló autoregresszív részre, és a másik egy ismeretlen/véletlent tartalmazó változóra vonatkozó mozgó átlag részből.

A szakdolgozat első része a VARMA folyamatok különböző felírásaival foglalkozik. Először azt mutatja be, hogy hogyan tevődik össze egy autoregresszív és egy mozgó átlag rész egy, illetve több dimenzióban. Tobábbiakban arra tér ki, hogy miért nem egyértelmű egy általános VARMA felírás, és hogyan lehet azt különböző módokon egyértelművé tenni. A második rész kointegrált folyamatokról, hibajavító modellekről és ezek kapcsolatáról szól. A harmadik részben ezt ültetem át a gyakorlatba, kriptovaluták árfolyamanak modellezésére alkalmazok hibajavító modellt. A Bitcoin, a Litecoin és az Ethereum kapcsolatát vizsgáltam, hogy a megelőző időszakbeli árak, és megelőző időszakok árai hogyan hatnak az aktuális időszak árváltozására.

1. fejezet

VARMA Modellek

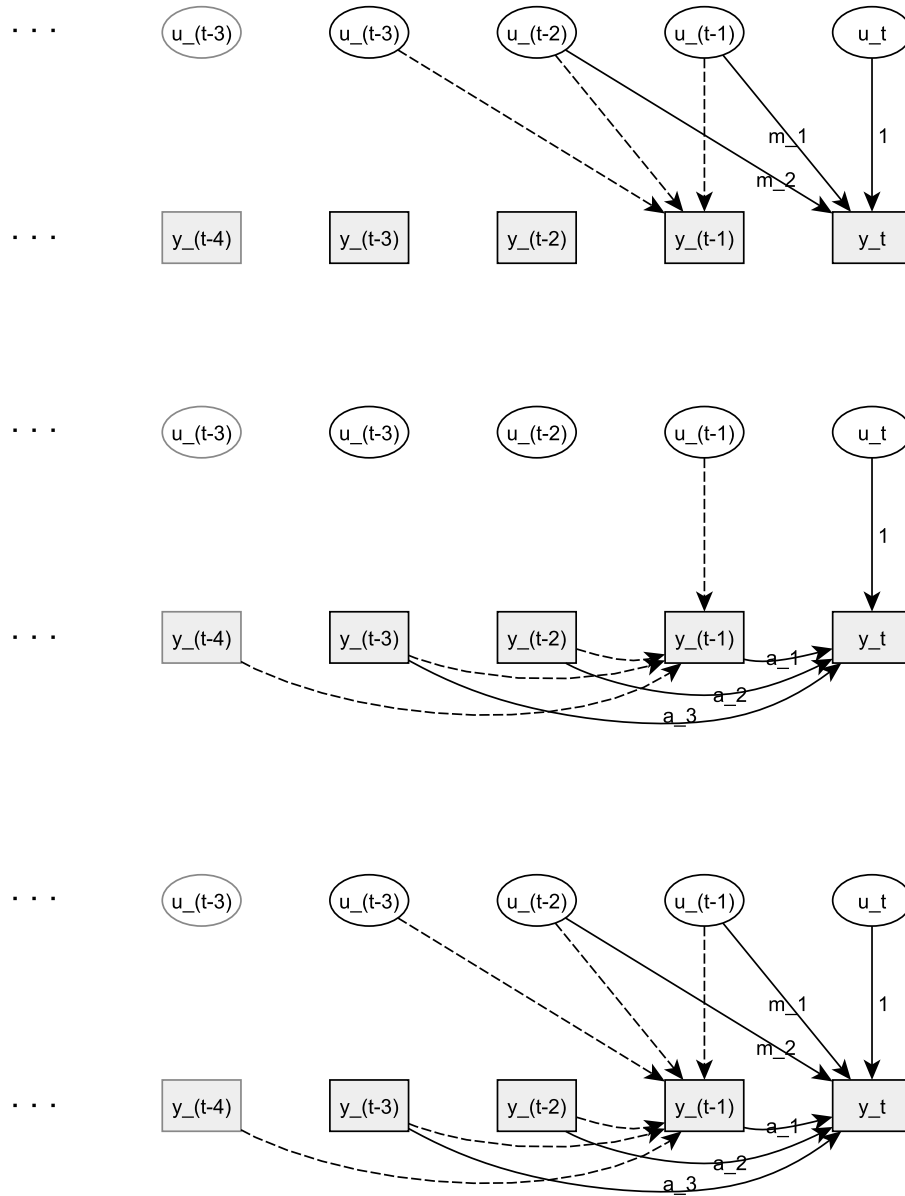
Egy VARMA folyamat egy vektor autoregresszív mozgóátlag folyamat. Ebben a bevezető részben azt szeretném bemutatni röviden, hogy egy ilyen folyamatot egyszerűbb részekből hogyan lehet összerakni, és néhány ábrával megmutatni, hogy én intuitívan hogyan értettem meg, hogy hogyan is működik egy VARMA folyamat.

Először az egy dimenziós folyamatokkal kezdjük. A legegyszerűbb talán egy mozgóátlag (MA) folyamat, ez egy diszkrét, idővel indexált sztochasztikus folyamat, és az értékei egy másik sztochasztikus folyamat elemeinek a lineáris kombinációi, ezért hívjuk mozgóátlagnak, mert minden eleme a másik folyamat néhány tagjának a súlyozott összege. Ez a mögöttes folyamat a zaj folyamat, amelyről feltesszük, hogy egy független értékű folyamat, melynek elemei azonos eloszlásúak, és 0 várható értékűek. Jelöljük ennek egy elemét u_t -vel, a mozgóátlag elemeit pedig y_t -vel. Ekkor y_t felírható

$$y_t = \mu + u_t + \sum_{i=1}^q m_i u_{t-i}$$

alakban, ahol μ a folyamat várható értéke; azt mondjuk, hogy y_t egy MA(q) folyamat. Ez azt jelenti, hogy a zaj folyamat értékeire q időegységet tekintünk vissza, és a zaj aktuális értékét is figyelembe vesszük. Ha az aktuális értéket nem $m_0 = 1$ súllyal akarjuk figyelembe venni, akkor a zaj folyamat egészét tudjuk súlyozni, ezzel egy új zaj folyamatot választva, és az összes többi súlyt visszaigazítva. Elképzelhető $q = \infty$ visszaléptetés is. Egy MA(2) folyamatot illusztrál a 1.1 ábra felső része.

Egy autoregresszív (AR) folyamat hasonlít az MA-ra abban, hogy itt is egy független, azonos eloszlású, 0 várható értékű zaj folyamat a mögöttes folyamat. Viszont itt a vizsgált y_t folyamat a saját korábbi értékeitől is lineárisan függ, és a zaj folyamatnak pedig csak



1.1. ábra. Egy dimenziós AR(3), MA(2), és ARMA(3,2) folyamat: a csúcsok a folyamat, és a zaj értékei, az élek a súlyokat jelzik. y_t súlyai feketével vannak jelölve, ezeknek értékei is szerepelnek, y_{t-1} súlyai szagatottal, de itt is ugyanazok a súlyok értékei.

az utolsó értéke számít. Így

$$y_t = \mu + \sum_{i=1}^p m_i y_{t-i} + u_t,$$

alakban írható fel pontosan, $p = \infty$ is lehet, ekkor végtelen AR folyamatról beszélünk. Egy AR(3) folyamatot illusztrál az 1.1 ábra középső része.

Egy ARMA folyamat egy AR és egy MA részből áll, és egy független, azonos eloszlású, 0 várható értékű zaj folyamat a mögöttes folyamat, mint mindkét részfolyamatnál. A következő képlettel írható fel:

$$y_t = \mu + \sum_{i=1}^p a_i y_{t-i} + u_t + \sum_{i=1}^q m_i u_{t-i}.$$

Ezt a 1.1 ábra alsó része szemlélteti, egy ARMA(3,2) folyamattal.

Egy VARMA folyamat egy olyan ARMA ahol a zaj és a kapott folyamat is nem egy hanem többváltozós. Ezt úgy képzelhetjük el, hogy ha például a zaj folyamatunk m dimenziós, akkor az alsó ábrán a felső sorban minden csúcsot kicserélünk egy vektorra, azaz m új csúcsra, és minden régi él helyett m új élet rakunk, az összes új csúcsból az eredeti régibe. És ugyanezt megteesszük a fő folyamatunkkal is az alsó sorban, ha egy n dimenziós folyamatunk van, akkor az összes alsó csúcs helyett egy n dimenziós vektort, azaz n új csúcsot rakunk, és minden eddigi él helyett n újat rakunk. Így ahol az ARMA folyamatábrán ment él, ott mostantól egy $m \cdot n$ élköteg megy. Ahogy a csúcs-csoportok vektorok, úgy az élkötegek súlymátrixok. Így egy VARMA(p,q) folyamat a következő módon írható fel:

$$\underline{y}_t = \underline{\mu} + \sum_{i=1}^p A_i \underline{y}_{t-i} + \underline{u}_t + \sum_{i=1}^q M_i \underline{u}_{t-i}$$

Itt \underline{u}_t a többdimenziós zajfolyamat, \underline{y}_t a szintén vektor vizsgált folyamatunk, $\underline{\mu}$ a folyamat átlaga, A_i , és M_i pedig a súlymátrixok. A továbbiakban az egyszerűség kedvéért a vektorok jelölését el fogom hagyni. Ha nincs mozgóátlag rész, akkor VAR folyamatról beszélünk.

Azt mondjuk, hogy egy VARMA folyamat stabil, ha a VAR része stabil, azaz

$$\det(I_K - A_1 z - \dots - A_p z^p) \neq 0 \text{ minden } |z| \leq 1.$$

Egy VARMA folyamatot invertálhatónak nevezünk, ha az MA részére igaz, hogy

$$\det(I_K + M_1 z + \dots + M_q z^q) \neq 0 \text{ minden } z \in \mathbb{C}, |z| \leq 1.$$

A VARMA folyamatokat becsülni általában körülményesebb, mint az egyszerűbb VAR modellekkel dolgozni, és így a VAR folyamatokat gyakran használják VARMA folyamatok becslésére is. Ezt Cooley es Dweyer sokszor idézett mondata is alátámasztja: "Ugyan a VARMA modellekkel több becslési és identifikációs probléma van, de ezek a komplikációk nem indokolják azt, hogy emiatt szisztematikusan figyelmen kívül hagyjuk a mozgóátlag részeket." [9] Sőt akkor is érdemes VARMA modellt használni, ha ugyan egy VAR folyamatot vizsgálunk, de vannak hiányzó változóink.

E fejezetben a továbbiakban nagyrészt Lütkepohl [14] könyvének 12. fejezetét követem. Legyen y_t egy K -dimenziós, stacionárius folyamat, melynek VARMA(p, q) felírása

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t + M_1 u_{t-1} + \dots + M_q u_{t-q}. \quad (1.1)$$

Feltehetjük, hogy az átlag nulla, mivel ez a továbbiakban nem fog gondot okozni, ezért ezt a tagot már nem is tüntettük fel. A folyamatot felírhatjuk az L visszaléptetés (angolul *lag* vagy *backshift*) operátor segítségével is: $A(L)y_t = M(L)u_t$, ahol $A(L) := I_K - A_1 L - \dots - A_p L^p$ és $M(L) := I_K + M_1 L + \dots + M_q L^q$. Amennyiben a VARMA folyamat stabil és invertálható, akkor felírható végtelen VAR folyamatként is:

$$y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i} = \Phi(L)u_t = A(L)^{-1}M(L)u_t.$$

Van, hogy gyakorlatban ennél általánosabb formában felírható VARMA folyamatokat érdemes használni:

$$A_0 y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + M_0 v_t + M_1 v_{t-1} + \dots + M_q v_{t-q}$$

ahol v_t egy megfelelő független értékű folyamat, és az $A - 0$, M_0 nem feltétlenül identitás. Amennyiben A_0 és M_0 nonszinguláris, akkor ez a folyamat is átírható a standard VARMA(p, q) formára, ha balról szorozzuk A_0^{-1} -val és $u_t = A_0^{-1}M_0 v_t$ -t vesszük az új korrelálatlan zajnak.

$$y_t = A_0^{-1}A_1 y_{t-1} + \dots + A_0^{-1}A_p y_{t-p} + u_t + A_0^{-1}M_1 M_0^{-1}A_0 u_{t-1} + \dots + A_0^{-1}M_q M_0^{-1}A_0 u_{t-q}.$$

Ezzel a két változtatással egy olyan reprezentációt kapunk, ahol a t -beli (visszalépés nélküli) változókhoz egy egységmátrix tartozik. Az, hogy A_0 és M_0 nonszinguláris igazából nem ad megkötést, hogyha y_t egyik komponense sem írható fel a többi komponens lineáris kombinációjaként. Egy stabil és invertálható felírást ami olyan alakú, mint (1.1) standard VARMA reprezentációnak hívunk, megkülönböztetve azoktól ahol A_0 vagy M_0 nem az

identitás mátrix. Ebből már látszik, hogy egy VARMA felírás nem egyértelmű, mert egy nemszinguláris matrix-szal balról beszorozva szintén egy VARMA(p,q) felírást kapunk. Egy VAR folyamatnál már egyértelművé teszi a felírást, ha a standard formát ($A_0 = I_K$) választjuk. A VARMA folyamatoknál ez nincs így. Ugyan javít a helyzeten, hogyha a standard formát választjuk, de ez még nem ad elég megkötést az egyértelműséghez, a fejezet többi része a lehetséges megkötésekkel foglalkozik.

1.1. Ekvivalens reprezentáció, egyértelműség

Azt mondjuk, hogy egy y_t folyamat két reprezentációja ekvivalens, hogy ha ugyanazok a realizációik (egy nullmértékű halmaz kivételével), és az $y_t, y_{t+1}, \dots, y_{t+h}$ véges eloszlásaik megegyeznek minden $t, h \in \mathbb{N}$.

Legyen az y_t egy \mathcal{Q} várható értékű VARMA folyamat kononikus MA reprezentációja

$$y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i} = \Phi(L)u_t,$$

ahol $\Phi_0 = I_K$, és $\Phi(L) = \sum_{i=0}^{\infty} \Phi_i L^i$. Azt mondjuk, hogy akármelyik $A(L)y_t = M(L)u_t$ VARMA modell, amire $A(L)^{-1}M(L) = \Phi(L)$ teljesül, az egy ekvivalens reprezentációja y_t -nek. Azaz minden VARMA model ekvivalens amire $A(L)^{-1}M(L)$ ugyanazt a $\Phi(L)$ operátort adja. Így ahhoz, hogy egyértelmű VARMA reprezentációt kapjunk megkötéseket kell tenni az autoregresszív és a mozgóátlag részre úgy, hogy minden $\Phi(L)$ -hez csak egy $A(L), M(L)$ pár tartozhasson.

Azt már láttuk, hogy egy nemszinguláris matrix-szal balról szorozva nem változik a folyamat, így mostantól csak a standard reprezentációkkal foglalkozunk.

Vegyünk egy kétdimenziós VARMA(1,1) folyamatot:

$$y_t = A_1 y_{t-1} + u_t + M_1 u_{t-1}.$$

Ennek a kanonikus mozgóátlag reprezentációja:

$$y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i} = u_t + \sum_{i=0}^{\infty} (A_1^i + A_1^{i-1} M_1) u_{t-i}.$$

Így az összes VARMA(1,1) folyamat, amelyre igaz, hogy $M_1 = -A_1$ igazából ugyanaz az egyetlen folyamat. Persze ebben az esetben az u_t -n kívül minden tag kinullázódik, VARMA(0,0) folyamatot kapunk, ami korrelálatlan zaj. Ugyanez a visszaléptetés operátorral

felírva: $(I_2 - A_1 L)y_t = (I_2 + M_1 L)u_t$. Ez a jelenség, nevezetesen az, hogy az autoregresszív és a mozgóátlag rész kiüti egymást már az egydimenziós ARMA folyamatoknál is előfordul. Ott ennek az esetnek a kizárására elégséges feltenni, hogy az *AR* és *MA* operátoroknak nincs közös osztójuk.

Tegyük egy hasonló feltételt a több változós esetben: y_t ne legyen korrelálatlan, azaz $M_1 \neq -A_1$. Sajnos a többváltozós esetben ez még nem hoz egyértelműséget, ezt egy másik példán keresztül mutatom be: Legyen $A_1 = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}$ és $M_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, ahol $\alpha \neq 0$. Ekkor a kanonikus *MA* reprezentáció együtthatói a következők:

$$\Phi_1 = A_1, \quad \Phi_2 = \Phi_3 = \dots = 0, \quad (1.2)$$

mivel $A_1^i = 0$ ha $i > 1$. Ugyanezt az *MA* kanonikus reprezentációja annak a folyamatnak is, amelyre $A_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ és $M_1 = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}$.

Általánosabban a fenti (1.2) reprezentációja minden olyan ARMA folyamatnak, amelyre

$$A_1 = \begin{bmatrix} 0 & \alpha + m \\ 0 & 0 \end{bmatrix} \text{ és } M_1 = \begin{bmatrix} 0 & -m \\ 0 & 0 \end{bmatrix},$$

akármilyen m esetén, (és így minden m -re stabil és invertálható folyamatot kapunk).

Nézzük meg, honnan jön a paraméterek választásának szabadsága! Nézzük a következő VAR operátort, (erre még később vissza fogok utalni):

$$Z = I_2 - \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix} L,$$

ennek az inverze

$$Z^{-1} = I_2 + \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix} L,$$

azaz az inverze is egy véges VAR operátor. Az ilyen operátorok jelentenek problémát amikor egyértelműen akarunk paraméterezni egy VARMA folyamatot. Egy ilyen operátorral besorozva egy VARMA reprezentációt az egyik eredeti operátor (AR vagy MA) egy részét eltüntetheti, úgy hogy közben a másik eredeti operátornak megmarad a véges foka.

Nézzük ezt egy konkrét VARMA(1,1) modellen: $A(L)y_t = M(L)u_t$, ahol

$$A(L) := \begin{bmatrix} 1 - \alpha_{11}L & -\alpha_{12}L \\ 0 & 1 \end{bmatrix}$$

és

$$M(L) := \begin{bmatrix} 1 + m_{11}L & m_{12}L \\ 0 & 1 \end{bmatrix}$$

Ez a két operátor nem oltja ki egymást, ha $\alpha_{11} \neq -m_{11}$ és $\alpha_{12} \neq -m_{12}$, de attól még mindkét operátorból balról kifaktorálhatjuk a

$$D(L) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & \gamma \\ 0 & 0 \end{bmatrix} L = \begin{bmatrix} 1 & \gamma L \\ 0 & 1 \end{bmatrix}$$

operátort, anélkül hogy a struktúrájuk megváltozna:

$$A(L) = D(L) \begin{bmatrix} 1 - \alpha_{11}L & -(\gamma + \alpha_{12})L \\ 0 & 1 \end{bmatrix}$$

$$M(L) = D(L) \begin{bmatrix} 1 + m_{11}L & (m_{12} - \gamma)L \\ 0 & 1 \end{bmatrix}$$

a fentebbi operátorokból $D(L)$ -t elhagyva a következő operátorokat kapjuk, amiket újra szorzattáalakítva a következőket kapjuk:

$$\begin{bmatrix} 1 - \alpha_{11}L & -(\gamma + \alpha_{12})L \\ 0 & 1 \end{bmatrix} = D(L) \begin{bmatrix} 1 - \alpha_{11}L & -(2\gamma + \alpha_{12})L \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 + m_{11}L & (m_{12} - \gamma)L \\ 0 & 1 \end{bmatrix} = D(L) \begin{bmatrix} 1 + m_{11}L & (m_{12} - 2\gamma)L \\ 0 & 1 \end{bmatrix}$$

Azaz újra ki tudjuk faktorálni $D(L)$ -t, sőt akárhányszor megtehetjük ezt. Így attól még hogy mindkét operátor rangját nem csökkentjük egyszerre, attól még lehetséges hogy néhány elemet kinullázzunk, azaz hogy $A(L)$ -ből vagy $M(L)$ -ből néhány operátort kifaktoráljunk. A $D(L)$ operátorra szintén igaz, hogy az inverze véges rangú:

$$D(L)^{-1} = \begin{bmatrix} 1 & -\gamma L \\ 0 & 1 \end{bmatrix}.$$

Az olyan véges fokú operátorokat, amelyeknek az inverze is véges fokú unimoduláris operátoroknak nevezzük. Az ilyen operátorok karakterizálja, hogy a determinánsuk egy nemnulla konstans. Azaz a determináns nem tartalmaz visszaléptetéses tagot, azaz egyetlen L hatványt sem. Unimoduláris operátorra példa a fentebbi Z operátor, melynek determinánsa

$$\left| I_2 - \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix} L \right| = \left| \begin{bmatrix} 1 & -\alpha L \\ 0 & 1 \end{bmatrix} \right| = 1.$$

Az, hogy a véges fokú operátoroknál a véges fokú inverzből következik a nemnulla konstans determináns a következőből látható:

$$A(L)^{-1} = A(L)^{adj} / |A(L)| = |A(L)|^{-1} A(L)^{adj}.$$

Ebből a determináns egydimenziós operátor. Egy véges fokú invertálható egydimenziós operátornak pedig végtelen fokú inverze van, hacsak a fok nem 0, azaz ha csak nem konstans.

Ahhoz hogy egyértelműségi feltételt adhassunk egy VARMA reprezentációra, először is azt kell biztosítani, hogy a VAR és MA tagoknak ne legyen közös bal oldali osztójuk, kivéve az unimoduláris mátrixokat. Az ilyen tulajdonságú $A(L)$ $M(L)$ operátorpárokat bal relatív prímekeknek hívjuk. Azt mondjuk, hogy az $[A(L) : M(L)]$ mátrix operátor bal relatív prím, hogyha abban az esetben ha léteznek olyan $D(L)$, $\bar{A}(L)$ és $\bar{M}(L)$ operátorok amik kielégítik a

$$D(L) [\bar{A}(L) : \bar{M}(L)] = [A(L) : M(L)]$$

egyenletet, akkor $D(L)$ unimoduláris.

A fenti példából látszik, hogy általában a unimoduláris mátrixok kifaktorálásának a lehetősége elkerülhetetlen, ha nem adunk meg további megkötéseket. Így ahhoz, hogy egyértelmű felírást kapjunk olyan megkötéseket kell adnunk, hogy az egyedül az I_K mátrixszal lehessen balról faktorizálni, a megadott VARMA reprezentációt megtartva.

A továbbiakban 3 példa következik egyértelmű reprezentációra:

1.2. Végső forma

1.2.1. Definíció. Az $A(L)y_t = M(L)u_t$ VARMA egyenlet végső formában van megadva, ha $M_0 = I_K$ és $A(L) = \alpha(L)I_K$, ahol $\alpha(L) := 1 - \alpha_1 L - \dots - \alpha_p L^p$ egy (egydimenziós) skalár operátor és $\alpha_p \neq 0$.

1.2.2. Példa.

$$(1 - \alpha_1 L - \alpha_2 L^2 - \alpha_3 L^3) \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 1 + m_{11,1}L & m_{12,1}L \\ m_{21,1}L & 1 + m_{22,1}L \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

Ez a felírás egyértelmű, mivel a $\alpha(L)y_t = M(L)u_t$ egyenlet csak a $D(L) = I_K$ operátorral lehet balról szorozni, úgy, hogy az autoregresszív rész skalár együtthatójú maradjon. Ahhoz hogy a $D(L)\alpha(L)I_K$ operátor p rangú maradjon, a $D(L)$ -nek nulla fokúnak kell lennie, így a $D(L) = D$ egy skalár mátrix, L -től nem függ.

1.3. Echelon forma

1.3.1. Definíció. Az $A(L)y_t = M(L)u_t$ VARMA egyenlet echelon formában van megadva, ha a VAR és MA operátorok, $A(L) = [\alpha_{ki}(L)]_{k,i=1,\dots,K}$ és $M(L) = [m_{ki}(L)]$ bal relatív prímelek, és kielégítik az alábbi feltételeket. Az operátorokat mátrixként tekintve a sorok fokait vizsgáljuk először. Az $\alpha_{ki}(L)_{(i=1,\dots,K)}$ és az $m_{kj}(L)_{(j=1,\dots,K)}$ operátorok - amik $A(L)$ illetve $M(L)$ k -adik sorai - p_k fokúak, és a következő alakúak:

- $\alpha_{kk}(L) = 1 - \sum_{j=1}^{p_k} \alpha_{kk,j} L^j$, minden $k = 1, \dots, K$,
- $\alpha_{ki}(L) = - \sum_{j=p_k-p_{ki}+1}^{p_k} \alpha_{ki,j} L^j$ ha $k \neq i$,
- $m_{ki}(L) = \sum_{j=0}^{p_k} m_{ki,j} L^j$, minden $k, i = 1, \dots, K$,
- és $M_0 = A_0$

Az $\alpha_{ki}(L)$ VAR operátorokban

$$p_{ki} = \begin{cases} \min(p_k + 1, p_i) & \text{ha } k \geq i, \\ \min(p_k, p_i) & \text{ha } k < i, \end{cases} \quad \text{minden } k, i = 1, \dots, K. \quad (1.3)$$

Azaz a p_{ki} megadja a szabad paraméterek számát az $\alpha_{ki}L$ operátorban, $k \neq i$ esetén. A sorok fokait (p_1, \dots, p_K) Kronecker indexeknek nevezzük. Az összegük, $\sum_{i=1}^K p_i$ pedig a McMillan fok. Így mindkét VARMA rang $p = q = \max(p_1, \dots, p_K)$.

Egy (p_1, \dots, p_K) Kronecker indexekkel rendelkező VARMA folyamat jelölése néha $ARMA_E(p_1, \dots, p_K)$. Nézzünk egy példát egy echelon formában felírt kétdimenziós VARMA folyamatra, pontosabban egy $ARMA_E(2, 1)$ -re:

$$\begin{aligned} \begin{bmatrix} 1 - \alpha_{11,1}L - \alpha_{11,2}L^2 & -\alpha_{12,2}L^2 \\ -\alpha_{21,0} - \alpha_{21,1}L & 1 - \alpha_{22,1}L \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} \\ = \begin{bmatrix} 1 + m_{11,1}L + m_{11,2}L^2 & m_{12,1}L + m_{12,2}L^2 \\ -\alpha_{21,0} + m_{21,1}L & 1 + m_{22,1}L \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \end{aligned}$$

Ugyanez a visszaléptetés operátor nélkül, visszaléptetésenként csoportosítva, és balról az összes múltbeli elemet a jobb oldalra rendezve:

$$\begin{aligned}
\begin{bmatrix} 1 & 0 \\ -\alpha_{21,0} & 1 \end{bmatrix} \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} &= \begin{bmatrix} \alpha_{11,1} & 0 \\ \alpha_{21,1} & \alpha_{22,1} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{11,2} & \alpha_{12,2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} \\
&+ \begin{bmatrix} 1 & 0 \\ -\alpha_{21,0} & 1 \end{bmatrix} \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} + \begin{bmatrix} m_{11,1} & m_{12,1} \\ m_{21,1} & m_{22,1} \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} \\
&+ \begin{bmatrix} m_{11,2} & m_{12,2} \\ 0 & 1+0 \end{bmatrix} \begin{bmatrix} u_{1,t-2} \\ u_{2,t-2} \end{bmatrix}
\end{aligned}$$

Ebben a modellben a Kronecker indexek $p_1 = 2$ és $p_2 = 1$. A p_{ki} számok pedig a következők:

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$$

A diagonálison kívüli p_{12} és p_{21} azt jelzi, hogy mennyi paramétere van $\alpha_{12}(L)$ -nek illetve $\alpha_{21}(L)$ -nek. Mivel $\alpha_{12}(L)$ az első sorba tartozik, és így $p_1 = 2$ fokú, és csak egy szabad paramétere van, így $\alpha_{12}(L) = -\alpha_{12,2}L^2$ alakú. Ugyanezzel a gondolatmenettel, $\alpha_{21}(L)$ a második sorban van, így $p_2 = 1$ fokú, és $p_{21} = 2$ szabad paramétere van, így $\alpha_{21}(L) = -\alpha_{21,0} - \alpha_{21,1}L$.

Az echelon forma egy tulajdonsága, hogy A_0 egy alsó háromszög mátrix, és a diagonálisban pedig 1-ek szerepelnek. Mivel $M_0 = A_0$ ezért ugyanez igaz M_0 -ra is. A szabad paraméterek közül bármelyik lehet akár 0 is, így a p vagy q fok előfordulhat, hogy kisebb, mint $\max(p_1, \dots, p_K)$. Ha például a fenti modellben $m_{11,2}$ és $m_{12,2}$ is 0, akkor $q = 1 < \max(p_1, p_2) = 2$. Ahhoz, hogy a reprezentáció tényleg (p_1, \dots, p_K) Kronecker indexű echelon forma legyen, ahhoz kell, hogy $[A(L) : M(L)]$ k -adik sorában legalább egy operátornak p_k rangúnak kell lennie, úgy hogy a p_k visszaléptetéshez tartozó paraméter nem nulla.

Az echelon forma is egyértelmű reprezentációt ad, azaz ha egy VARMA folyamat echelon formában van felírva, akkor azt máshogy nem lehet felírni echelon formában. Az echelon forma ugyan komplikáltabb, mint a végső forma, de általában kevesebb szabad paramétert tartalmaz. Így könnyebben kezelhető, ha például becsülni szeretnénk egy VARMA folyamat paramétereit.

1.4. Skalár komponens modell

A skalár komponens modell (angolul scalar component modell, SCM) lényege, hogy az eredeti VARMA(p, q) modellnél egyszerűbb beágyazott struktúraalapján adja meg a folyamatot. Ezek az egyszerűbb struktúrák az úgynevezett skalár komponensek, melyek az eredeti változók olyan lineáris kombinációi, melyek alacsonyabb rendben függenek a múlttól. Az autoregresszív részük lehetőleg rövidebb, mint p időegységnyit tekint vissza. A mozgóátlag részük pedig kevesebb, mint q időegységnyit tekint vissza.

A SCM-t Tiao és Tsay vezette be 1989-ben [17], majd az ebben levő hiányosságokat javítva Athanasopoulos és Vahid fejlesztette tovább a modellt [7], később Poskitt-tel együtt is [6]. A továbbiakban az utóbbi két cikkegondolatmenetét követem.

Legyen $y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \epsilon_t - M_1 \epsilon_{t-1} - \dots - M_q \epsilon_{t-q}$ egy K -dimenziós VARMA(p, q) folyamat.

1.4.1. Definíció. Azt mondjuk hogy az $z_t = f^T y_t$ nemnulla lineáris kombináció egy $SCM(p_1, q_1)$, (ahol $f \in \mathbb{R}^K$), ha z_t teljesíti az alábbiakat:

- $f^T A_{p_1} \neq 0$, ahol $0 \leq p_1 \leq p$
- $f^T A_l = 0$, ha $l = p_1 + 1, \dots, p$
- $f^T M_{q_1} \neq 0$, ahol $0 \leq q_1 \leq q$
- $f^T M_n = 0$, ha $l = q_1 + 1, \dots, q$

Egy-egy ilyen skalár komponens, egy $ARMA(p, q)$ folyamat lesz, legfeljebb (p_1, q_1) ranggal, ugyanis lehet kifaktorálható közös tényezőjük.

Az SCM eljárás kanonikus korreláció tesztek sorozatával keres a fentiek megfelelő K darab független lineáris kombinációt, a legegyszerűbb $SCM(0, 0)$ modelltől indulva. Ha megvan a K lineárisan független skalár komponens, és a hozzájuk tartozó $F = (f_1, \dots, f_K)^T$ transzformációs mátrix. Ekkor az eredeti y_t -re vonatkozó egyenletet F -fel elforgatva

$$F y_t = A_1^* y_{t-1} + \dots + A_p^* y_{t-p} + \epsilon_t^* - M_1^* \epsilon_{t-1}^* - \dots - M_q^* \epsilon_{t-q}^*$$

kapjuk, ahol $A_i^* = F A_i$, $\epsilon_t^* = F \epsilon_t$, és $M_i^* = F M_i F^{-1}$. Az egyenlet jobb oldalán több mátrixsor is nulla lesz. Azonban ha vannak erősen egymásba ágyazott komponensek,

$SCM(p_r, q_r)$ és $SCM(p_s, q_s)$, ahol $p_r > p_s$ és $q_r > q_s$ akkor még nem értünk el egyértelműséget. Ekkor még $\min(p_r - p_s, q_r - q_s)$ darab AR vagy MA komponenst még ki kell nullázni. Az SCM során az MA rész ilyen együtthatóit szokás nullázni.

Ha F -ben is vannak szabad komponensek, az egyértelműséghez F -re is kell megkötéseket tenni. Adott skalár komponens struktúra esetén már egyértelmű a modellek felírása. Adott struktúra mellett maximum likelihood módszerrel becsülik az egész SCM együtthatóit.

2. fejezet

Hiba javító modellek

2.1. Integrált folyamatok

Ebben a fejezethez [14], [5] és [2] forrásokat követem vegyesen.

Azt mondjuk hogy, egy egydimenziós AR folyamat d rendben integrált, ezt $I(d)$ -vel jelöljük, ha d egységgyöke van a karakterisztikus polinomjának. Ha egy $I(1)$ folyamatunk van, akkor könnyű látni, hogyha az eredeti folyamat első differenciáit vesszük: $\Delta y_t := (1 - L)y_t = y_t - y_{t-1}$, akkor egy stabil folyamatot kapunk. A kezdeti értékek megfelelő választásával ez a folyamat választható stacionáriusnak. Általánosabban egy $I(d)$ folyamat d differenciálás után lesz ilyen tulajdonságú. A stabil, stacionárius folyamatokat szokás $I(0)$ -al jelölni.

Ennél egy pontosabb definíció az $I(1)$ folyamatokra:

2.1.1. Definíció. Azt mondjuk, hogy y_t egy $I(1)$ folyamat, ha $\Delta y_t = w_t$ egy stacionárius folyamat végtelen MA reprezentációval, $w_t = \sum_{j=0}^{\infty} \theta_j u_{t-j} = \theta(L)u_t$, ahol a mozgóátlag együtthatóira igaz, hogy $\sum_{j=0}^{\infty} j |\theta_j| < \infty$, $\theta(1) = \sum_{j=0}^{\infty} \theta_j \neq 0$ és $u_t \sim (0, \sigma_u^2)$ fehér zaj.

2.2. Kointegrált folyamatok

Vannak olyan közgazdasági idősor változók, melyekről úgy gondoljuk, hogy közöttük valamilyen lineáris egyensúlyi kapcsolat van. Például egy termék ára különböző területeken, vagy az állami kiadások és bevételek.

Gyűjtsünk egy ilyen változó csoportot egy vektorba $y_t = (y_{1t}, \dots, y_{Kt})^T$ és a hosszútávú

egyensúlyi kapcsolatuk legyen

$$\beta^T y_t = \beta_1 y_{1t} + \dots + \beta_K y_{Kt} = 0.$$

Egy-egy konkrét periódusban ennek a kapcsolatnak nem kell egzakt teljesülnie. Feltehető, hogy $\beta^T y_t = z_t$ egy sztochasztikus folyamatot vizsgálni, ami az egyensúlytól való eltérést mutatja. Ha valóban egyensúly van, akkor feltehetjük, hogy az y_t változók valamilyen értelemben együtt mozognak, és az z_t folyamat pedig stabil/stacionárius. Ez a modell nem zárja ki annak a lehetőségét, hogy az y_t változók egy csoportként, együtt mozogjanak. Ebben az esetben lehet, hogy egy közös sztochasztikus trend mozgatja őket. Azaz nem lehet kizárni, hogy ugyan a változók külön-külön egy-egy integrált folyamatként viselkednek, de közben van olyan lineáris kombinációjuk, ami stacionárius. Az ilyen tulajdonságú integrált folyamatokat kointegráltak nevezzük.

2.2.1. Definíció. *Legyen y_t egy K -dimenziós folyamat. Azt mondjuk, hogy a változói (d, b) -rendben kointegráltak – röviden $y_t \sim CI(d, b)$ –, ha y_t minden komponense $I(d)$, és létezik olyan $z_t = \beta^T y_t$ lineáris kombinációjuk ahol $\beta^T = (\beta_1, \dots, \beta_K) \neq 0$ és a z_t egy $I(d - b)$ folyamat.*

Az itt szereplő β vektort kointegráló vektornak hívjuk, az y_t folyamatot pedig kointegrált folyamatnak.

Az kointegrált folyamat modellt Granger [12] vezette be. A definíció a későbbi Englel közös [11] munkában jelent meg először.

Egy másik definíció egyszerűbbé teszi a kointegrált folyamatok kezelését:

2.2.2. Definíció. *Azt mondjuk, hogy egy K -dimenziós y_t folyamat $I(d)$, ha $\Delta^d y_t$ stabil, de $\Delta^{d-1} y_t$ még nem stabil.*

Egy y_t $I(d)$ folyamat kointegrált, ha létezik olyan $z_t = \beta^T y_t$ lineáris kombinációja, ahol $\beta \neq 0$ és z_t alacsonyabb, mint d rendben integrált.

Ez a második definíció megengedőbb, mert nem kell minden komponensnek $I(d)$ -nek lennie, lehetnek kisebb rendű komponensek is; mivel ha kisebb rendű komponenseket az integráció rendjénél tovább integrálunk, attól, az még $I(0)$ marad.

A kointegráló vektor nem feltétlenül egyértelmű.

2.3. Hiba javító modellek

A hibajavító modellek (angolul Error Correction Model, ECM) a kointegrált folyamatok előfutárai voltak.

Egy hibajavító modell több változóból és egy azok közti egyensúlyi kapcsolatból áll, és egy-egy változó mozgása az egyensúlyi állapottól való eltérésektől függ.

Tegyük fel hogy van egy termékünk amit két piacon is árulnak; jelölje az árát a t . időintervallumban az egyik piacon y_{1t} , míg a másik piacon y_{2t} , és legyen az egyensúlyi kapcsolat köztük $y_{1t} = \beta_1 y_{2t}$. Tegyük fel, hogy az egyik piacon a termék árváltozása attól függ, hogy a megelőző időszakban mekkora volt az eltérés az egyensúlytól:

$$\Delta y_{1t} = \alpha_1(y_{1,t-1} - \beta_1 y_{2,t-1}) + u_{1t}.$$

És egy hasonló reláció áll fent a másik piacon is:

$$\Delta y_{2t} = \alpha_2(y_{1,t-1} - \beta_1 y_{2,t-1}) + u_{2t}.$$

Egy általánosabb hibajavító modellben az árváltozás az azt megelőző periódusok között megtörtént árváltozástól is függ, és itt mindkét piac árváltozását figyelembe kell venni:

$$\Delta y_{1t} = \alpha_1(y_{1,t-1} - \beta_1 y_{2,t-1}) + \gamma_{11,1} \Delta y_{1,t-1} + \gamma_{12,1} \Delta y_{2,t-1} + u_{1t},$$

$$\Delta y_{2t} = \alpha_2(y_{1,t-1} - \beta_1 y_{2,t-1}) + \gamma_{21,1} \Delta y_{1,t-1} + \gamma_{22,1} \Delta y_{2,t-1} + u_{2t},$$

Egynél több időbeli visszaléptetés is figyelembe vehető Δy_{it} -kre vonatkozó modellben. Olyan modell is elképzelhető, ahol egy konstans tag is van a lineáris becslésben. A hibajavító modell neve abból ered, hogy figyelembe veszi, hogy az előző időszakban hogyan tértünk el a hosszútávú egyensúlytól, és ennek segítségével igyekszik az egyes változókat visszakorrigálni.

2.4. Hibajavító modellek különböző becslései

Az első hibajavító modell, egy két lépéses eljárás, amit Engle és Granger vezetett be. [11] Először mindkét idősort (pl x_t , y_t) meg kell vizsgálni stacionaritás szempontjából, ezt az egységgyökök jelenlétének tesztelésével lehetséges. Erre alkalmas a (kiterjesztett) Dickey-Fuller teszt (angolul Augmented Dickey Fuller test, ADF), a Phillips-Perron (PP) teszt,

Kwiatkowski–Phillips–Schmidt–Shin (KPSS) teszt. A KPSS-en teszten kívül a nullhipotézis az hogy létezik egységgyök, míg a KPSS-nél a stacionaritás a nullhipotézis. Az ADF-ről később kicsit részletesebben írok.

Ha az eredeti idősorok nem stacionáriusak, viszont az első differenciasorozataik $(\Delta x_t \Delta y_t)$ már stacionáriusak, akkor lehet próbálkozni ECM modell illesztésével.

A két lépéses Engle-Grenger eljárásban a legkisebb négyzetek módszerével keresünk lineáris modellt az eredeti két változó kapcsolatára $y_t = \beta_1 + \beta_2 x_t + u_t$, ebben a lépésben az adat idősor jellegét nem vesszük figyelembe. Viszont a lineáris modellből kapott reziduálisokat (\hat{u}_t) megint idősorként vizsgáljuk. Ahhoz, hogy az eljárást folytatni tudjuk arra van szükség, hogy a reziduálisok idősora stacionárius legyen. Ekkor tudhatjuk, hogy az eredeti változók kointegráltak, közgazdasági értelemben pedig valamilyen hosszútávú (egyensúlyi) kapcsolat áll fent köztük.

Amennyiben ez a feltétel teljesül, akkor a rövid távra vonatkozó rész becslése következik, ami szintén egy OLS lineáris illesztés, ekkor az egyik differencia sorozat (Δy_t) elemeit becsljük a másik differenciasorozat (Δx_t) elemeivel, az előző becslésből kapott reziduálisok visszaléptetettjeivel (Δu_{t-1}) :

$$\Delta y_t = \beta_3 + \beta_4 \Delta x_t - \gamma \hat{u}_{t-1} + \epsilon_t.$$

Ebben a modellben β_3 egy konstans, β_4 rövidtávú együtttható azt mutatja, hogy az x_t azonnali változása milyen hatással van y_t változására, γ -t pedig szokás visszacsatoló, vagy hibajavító hatásnak hívni, mivel azt mutatja meg, hogy az előző időszakban az egyensúlytól való eltérés mennyit javul a következő időszakra.

Visszaírva \hat{u}_t -t

$$\Delta y_t = \beta_3 + \beta_4 \Delta x_t - \gamma (y_{t-1} - \hat{\beta}_1 - \hat{\beta}_2 x_{t-1}) + \epsilon_t$$

Az egy lépéses modellben rögtön az utóbb kapott egyenlet egy átrendezettjének az együttthatóit becsljük. Itt is teljesülni kell az eredeti sorozatok nem-stacionaritásának, és a differenciasorozatok stacionaritásának. Ha ezek teljesülnek, akkor a következő modellt illesztjük:

$$\Delta y_t = \alpha_1 + \alpha_2 y_{t-1} - \alpha_3 x_{t-1} + \alpha_3 \Delta x_t + \zeta_t$$

Ennek az illesztésnek az eredmény valamelyest különbözik a két lépéses modellétől, erre a következő fejezetben is lesz példa.

Kibővíthetjük az egy lépéses ECM modellt, ha a lineáris becslésnél a magyarázó változók közé bevehetjük a az x_t és y_t differencia sorozat visszaléptetettjeit is:

$$\Delta y_t = \alpha_1 + \alpha_2 y_{t-1} - \alpha_3 x_{t-1} + \sum_{i=1}^{n-1} \delta_i \Delta y_{t_i} + \sum_{i=0}^{m-1} \gamma_i \Delta x_{t_i} + \zeta_t$$

Az ECM modellek és a kointegrált folyamatok kapcsolatát Granger reprezentációs tétele mutatja meg, mely röviden azt mondja ki, hogy ha egy $I(1)$ vektor elemei pontosan akkor kointegráltak, hogyha van ECM reprezentációjuk. A tétel pontos kimondásához szükségünk van a következő jelölésekre: Legyen $m \leq n$, és M egy $m \times n$ -es mátrix, aminek a rangja n , jelölje ekkor M_\perp az M egy ortogonális kiegészítőjét. Legyen továbbá y_t egy K -dimenziós kointegrált folyamat, ahol a kointegráció rangja r .

2.4.1. Tétel. *Granger reprezentációs tétele* Legyen $\Delta y_t = \alpha \beta^T y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t$, $t > 0$ esetén, és $y_t = 0$ $t \geq 0$, u_t 0 várható értékű, független értékű zaj $t > 0$ esetén, és $u_t = 0$ ha $t \geq 0$. Legyen

$$C(z) := (1 - z)I_K - \alpha \beta^T z - \sum_{i=1}^{p-1} \Gamma_i (1 - z) z^i,$$

ahol a következő feltételek teljesülnek a paraméterekre:

- ha $\det C(z) = 0$, akkor $|z| > 1$ vagy $z = 1$.
- Pontosán $K - r$ egységgyöke van z -nek.
- α és β $K \times r$ mátrixok, és mindkettőnek a rangja r .

Ekkor y_t felírható a következő reprezentációban:

$$y_t = \Xi \sum_{i=1}^t u_i + \Xi^*(L) u_t + y^* 0,$$

ahol

$$\Xi = \beta_\perp \left[\alpha_\perp^T (I_K - \sum_{i=1}^{p-1} \Gamma_i) \beta_\perp \right]^{-1} \alpha_\perp^T,$$

$$\Xi^*(L) u_t = \sum_{j=0}^{\infty} \Xi_j^* u_{t-j} \text{ egy } I(0) \text{ folyamat,}$$

és y_0^* a kezdeti értékeket tartalmazza.

A VECM lényegében különbözik az Engle-Granger modellektől, itt nem kell előre megvizsgálni a folyamatokat stacionaritás szempontjából, és nem kell a változók között különbséget tenni aszerint, hogy melyiket kívánjuk magyarázó, és melyiket magyarázott változónak tekinteni.

A kointegrációra vonatkozó tesztet Johansen publikálta. [13]

2.5. Dickey-Fuller teszt

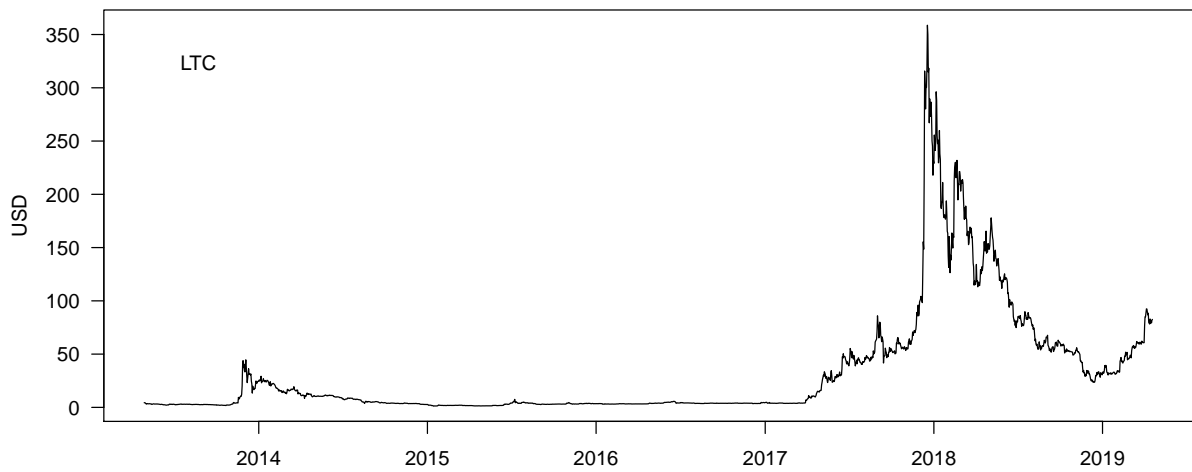
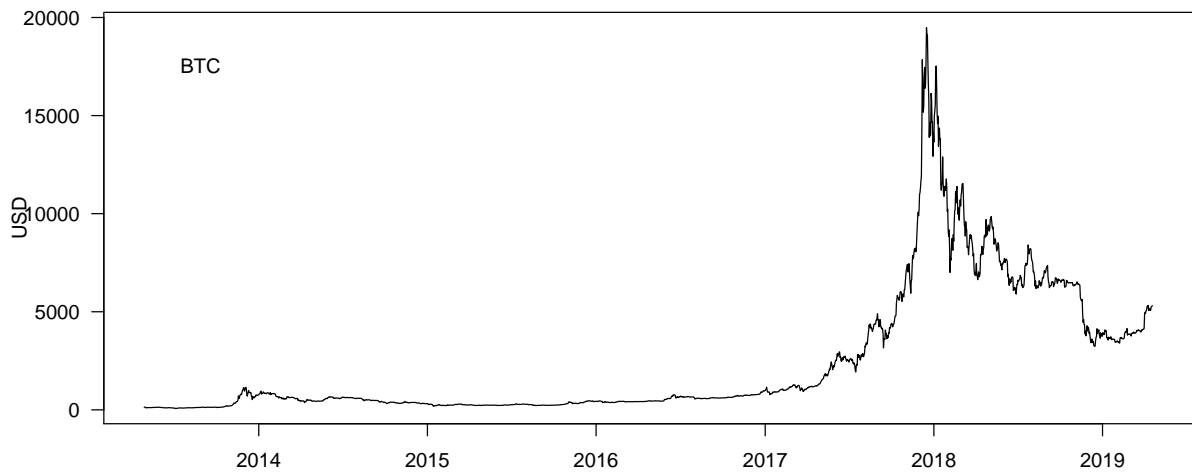
A Dickey-Fuller teszt [8, 10] egy egy dimenziós y_t idősrőről vizsgálja meg, hogy az integrált folyamat-e. Az $y_t = ay_{t-1} + u_t$ egyenletben teszteli az a együtthatót; ahol a nullhipotézis $a = 1$ azaz, hogy tartalmaz egységgyököt a folyamat; az ellenhipotézis pedig, az hogy $a < 1$, nincs egységgyök. Gyakorlatban inkább az átalakított $\Delta y_t = by_{t-1} + u_t$ egyenletet vizsgálják, itt a nullhipotézis $b = 0$, és az ellenhipotézis $b < 0$. Ha egy konstanst, és determinisztikus trendet is meg szeretnénk engedni, akkor az ezekkel kiegészített $\Delta y_t = by_{t-1} + \mu + \lambda t + u_t$ egyenletre kell a $b = 0$ nullhipotézist vizsgálni.

A tesztstatisztika a következő: $T = \hat{b}/SE(\hat{b})$, ahol \hat{b} a becült b , $SE(\hat{b})$ pedig a standard hibája. Ez ugyan T-statisztikának tűnik, de nem a Student T-eloszlást követi, ezért a kritikus értékei eltérnek. A Dickey-Fuller teszt csak akkor működik jól hogyha a u_t korrelálatlan; ha úgy gondoljuk, hogy u_t autokorrelált, akkor a kiegészített (augmented) Dicke-Fuller tesztet érdemes használni, ami y_t differenciasorozatának késleltetettjeit is beleveszi a tesztelt egyenletbe: $\Delta y_t = by_{t-1} + u_t + \sum_{i=1}^p c_i y_{t-1}$ és ugyanez determinisztikus trendet, és konstans tagot is megengedve: $\Delta y_t = by_{t-1} + \mu + \lambda t + u_t + \sum_{i=1}^p c_i y_{t-1}$. A statisztika megegyezik a korábban felírt Dickey-Fuller teszt statisztikájával.

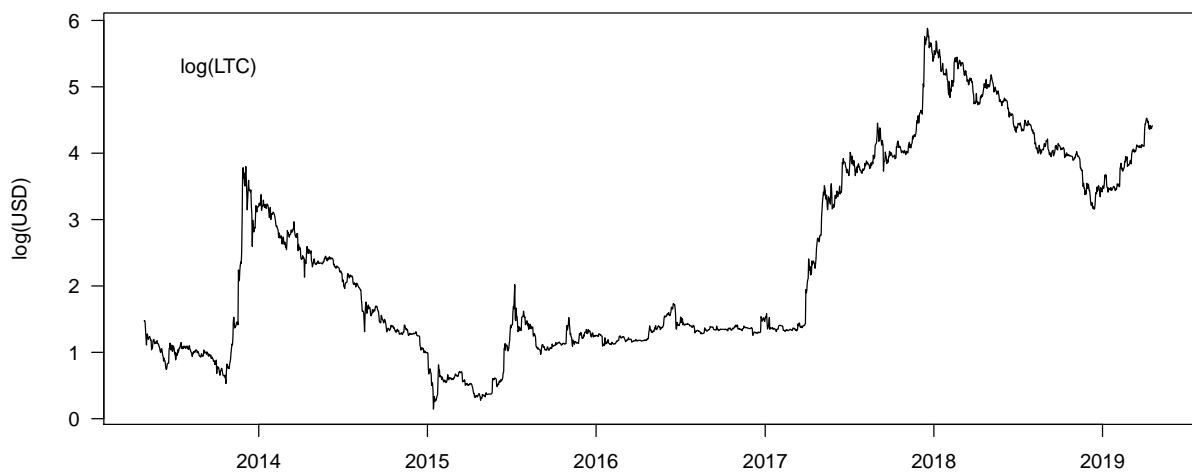
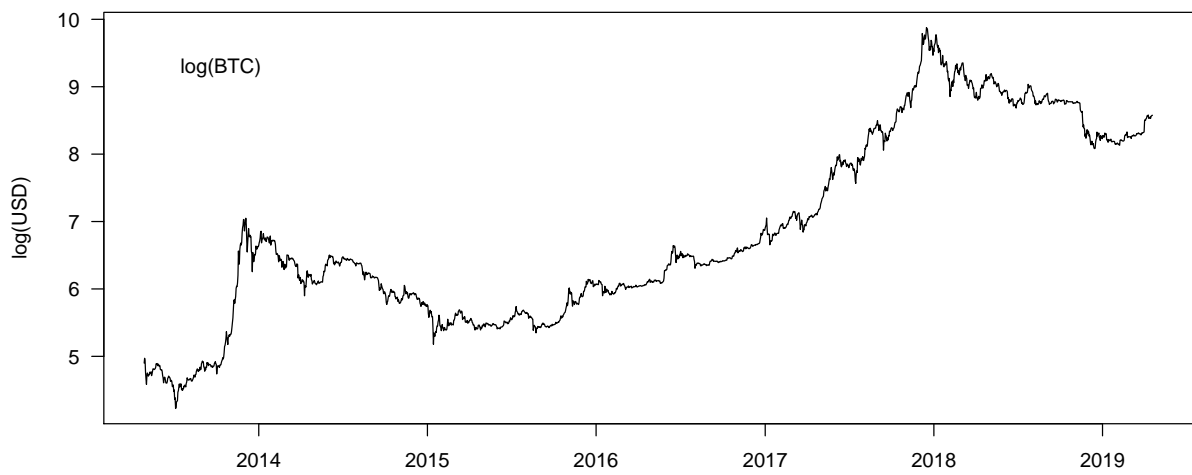
3. fejezet

Kriptovaluták kointegráltságának vizsgálata

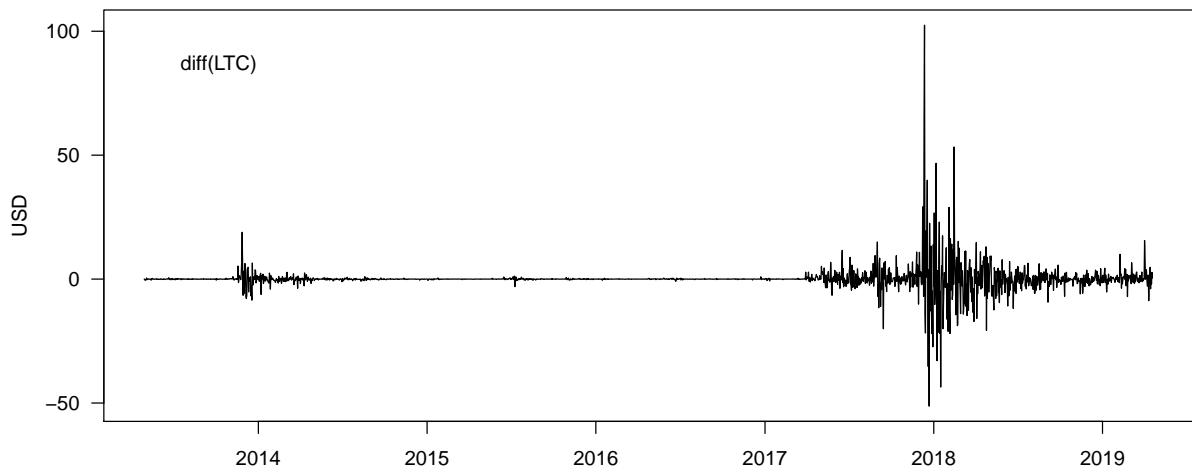
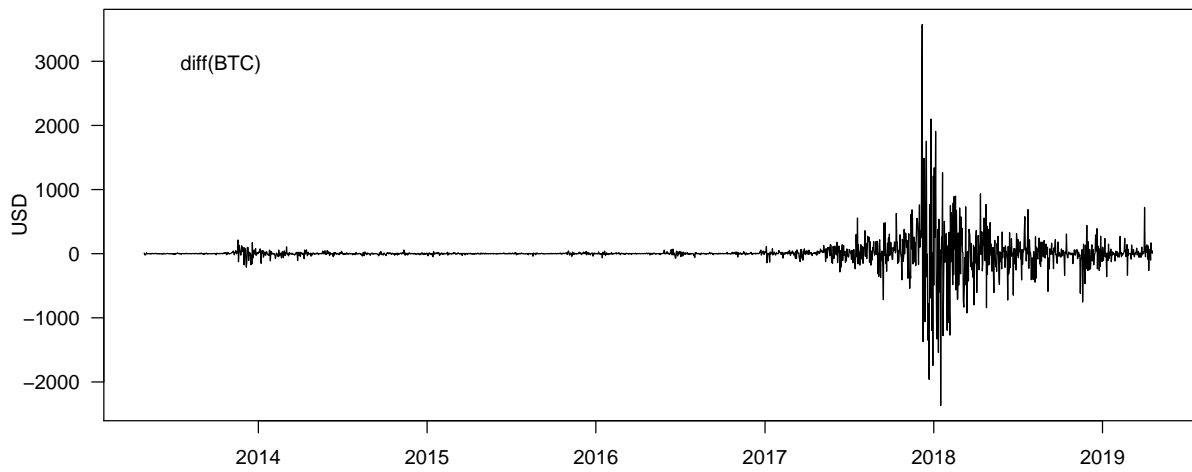
A hibajavító modellt három kriptovaluta árváltozásának a vizsgálatára alkalmaztam az R nyelv [16] és az aTSA [15] programcsomagjának a használatával. Ehhez a Bitcoin (BTC), Litecoin (LTC) és az Ethereumot (ETH) választottam ki, és az USD-ben kifejezett napi árakkal dolgoztam. Az adatok 2013. április 28. és 2019. április 20. közötti napi árak, kivéve az Ethereumnál, mert ott az első nap 2018. augusztus 9. Kezdetben csak a Bitcoinnal és Litecoinnal foglalkoztam, az Ethereumot csak később vettem be a modellekbe. A Bitcoin, Litecoin és Ethereum idősor adatai a coinmarketcap.com weboldalról származnak. [1, 4, 3] A kriptovaluták esetében nincsenek banki szünnapok, egy tranzakció akármi-kor megtörténhet, így minden napra vannak adataink. A kiinduló adat tartalmazott napi nyitó, napi záró, maximum és minimum árat. Mivel a nyitó és a záró ár egyszerűen UTC éjfél előtti utolsó és azutáni első tranzakcióra vonatkozó árat tartalmazza, így napi átlag helyett az előző napi záró és a másnapi nyitó ár átlagát vettem, és továbbiakban ezt fogom napi átlagnak hívni. A 3.1 ábrán ezeknek a mennyiségeknek a változása látható Bitcoin és Litecoin esetén. Az adat intuitív megismerésére ábrázoltam ennek a logaritmusát (3.2 ábra), és az eredeti idősorok változását (3.3 ábra) is. Az ábrák alapján az eredeti folyamatok, és a logaritmusok lehetnek stacionáriusok, míg a differenciák fehér zajhoz hasonló folyamatok.



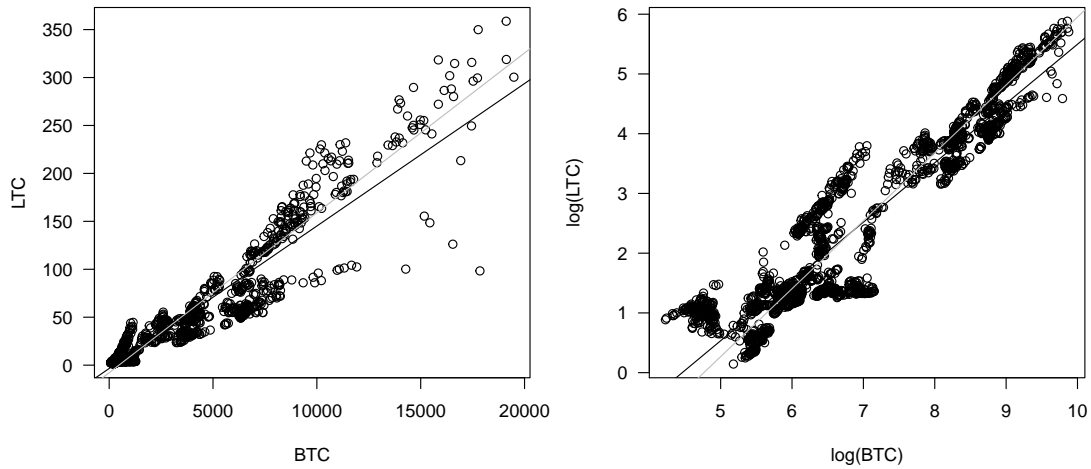
3.1. ábra. Bitcoin és Litecoin ára USD-ben 2013 áprilisától 2019 áprilisáig



3.2. ábra. Bitcoin és Litecoin USD árának logaritmusa 2013 áprilisától 2019 áprilisáig



3.3. ábra. Bitcoin és Litecoin USD árának változása 2013 áprilisától 2019 áprilisáig



3.4. ábra. BTC és LTC napi árai, illetve a logaritmusaik, mindkettő a két-két lehetséges legkisebb négyzetek módszerével illesztett lineáris modellel

3.1. Stacionaritás vizsgálata

Ahhoz, hogy hibajavító modellt illeszthessünk a két idősorra, először is az kell, hogy integráltak legyenek. Ezt a hipotézist az aTSA csomag Augmented Dickey-Fuller tesztjével és a Phillips-Perron tesztével is megvizsgáltam. A log-idősorok stacionaritását mindkettő teszt alapján erősen el lehetett vetni $p = 0.7$ és 0.9 körüli szinteken (a vizsgált modell változatok: drift és trend nélkül, drifttel és trend nélkül, drifttel és trenddel; és 0-7 visszalépést megengedő VAR modellek), míg az eredeti idősorokat stacionaritását is elvetette a teszt, de kisebb konfidenciaszinteken.

Az eredeti és a logaritmált napi áraknak is megvizsgáltam a differenciasorozatait, ezek mind stacionáriusnak bizonyultak minden fent felsorolt esetben $p \leq 0.01$ értékekkel.

Következő lépésként egy egyszerű lineáris modellt illesztettem az eredeti idősorra két-féleképpen, (a BTC és az LTC függő illetve magyarázó változóként), és ugyanezt a logaritmált adatokra. Mind a négy modell helyesen működött (statisztikailag szignifikáns volt), az illesztett modellek a 3.4 ábrán láthatók.

Ezután -az előző szakaszban leírt módszertan szerint- a lineárisan illesztett modellekből adódó reziduálisok vizsgálata következik. Erre szintén az aTSA beépített Dickey-Fuller tesztjét használtam, és emellett ábrázoltam is a reziduálisokat, az ábrák nagyjából egymás tükörképei, ha a függő és a magyarázó változót felcseréljük, így itt csak az az eset látható,

amikor az LTC a függő változó. (3.5 ábra) Itt az eredeti és a logaritmált adatra nagyon különböző eredmények születtek: az eredeti idősorra illesztett mindkét lineáris modell reziduálisai stacionáriusnak bizonyultak, míg a logaritmált adatokból származó reziduálisokról csak akkor fogadhattuk volna el a stacionaritást, ha trendet és driftet nem veszünk figyelembe. Így továbbiakban a logaritmált BTC és LTC idősorokkal nem foglalkoztam.

3.2. ECM modell illesztés

A két lépéses Engle-Granger teszt első lépésének eredményét már meg is kaptuk a lineáris modellekkel. A következő modelleket kaptam a hosszútávú kapcsolatra, $\approx 0.89 R^2$ statisztikával (mindkét esetben).

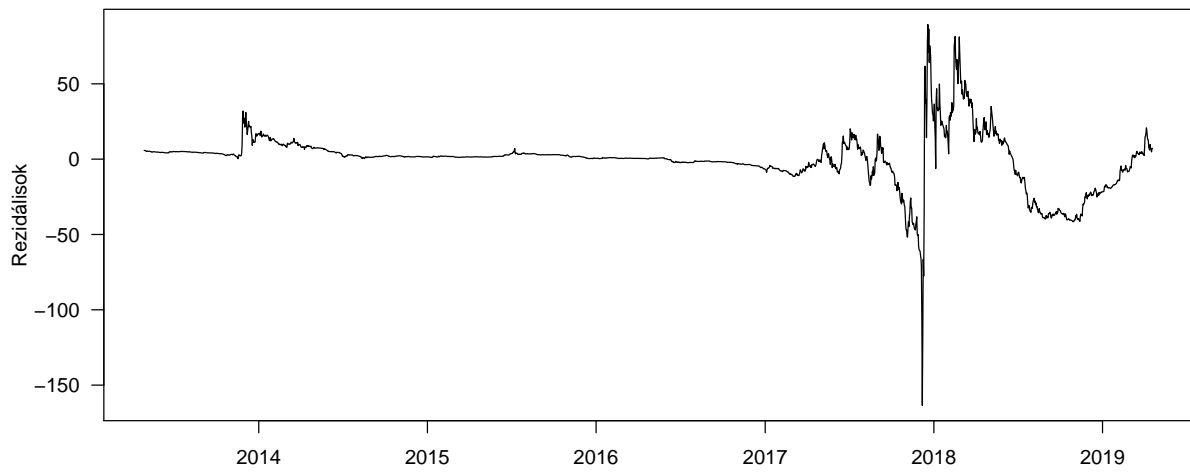
$$\begin{aligned} LTC_t &= -3.496 + 0.01486 \cdot BTC_t + U_t, \\ BTC_t &= 469.8 + 59.98 \cdot LTC_t + U'_t. \end{aligned}$$

A rövidtávú kapcsolatra az egyik kriptovaluta differenciasorozatát becsültem: a másik differenciasorozatával és az előző időszak reziduálisával. Itt az először a konstans becslésre kapott érték nem különbözött szignifikánsan nullától, így azt inkább 0-nak állítva megnéztem az eredményt. Ezzel a feltétellel a következő modelleket kaptam:

$$\begin{aligned} \Delta LTC_t &= 0.01211 \cdot \Delta BTC_t - 0.03300 \cdot \hat{U}_{t-1} + W_t \\ \Delta BTC_t &= 26.04 \cdot \Delta LTC_t + 0.0007176 \cdot \hat{U}'_{t-1} + W'_t \end{aligned}$$

Az \hat{U} és \hat{U}' az előző lineáris modellek reziduálisai, $\Delta BTC_t = BTC_t - BTC_{t-1}$, és $\Delta LTC_t = LTC_t - LTC_{t-1}$. Itt már jóval alacsonyabb R^2 statisztikákat kaptam, ≈ 0.33 mindkét esetben. A második egyenletben a reziduálisra vonatkozó becsült konstans nem tűnt szignifikánsnak, az összes többi pedig $p \leq 0.001$ szinten volt szignifikáns. A két egyenletpárt összedolgozva az egy lépéses modellnek megfelelő formába a következő egyenleteket kapjuk.

$$\begin{aligned} \Delta LTC_t &= 0.01211 \cdot \Delta BTC_t + -0.03300(LTC_{t-1} + 3.496 - 0.01486 \cdot BTC_{t-1}) + W_t \\ &= 0.1154 + 0.01211 \cdot \Delta BTC_t - 0.03300 \cdot LTC_{t-1} + 0.0004904 \cdot BTC_{t-1} + W_t \\ \Delta BTC_t &= 26.04 \cdot \Delta LTC_t + 0.0007176(BTC_{t-1} - 469.8 - 59.98 \cdot LTC_{t-1}) + W'_t \\ &= -0.3371 + 26.04 \cdot \Delta LTC_t + 0.0007176 \cdot BTC_{t-1} - 0.04304 \cdot LTC_{t-1} + W'_t \end{aligned}$$



3.5. ábra. Egyszerű lineáris modellek reziduálisai az eredeti adata, és a logaritmizált adatra, a litecoin a függő változó

Az egy lépéses eljárás a következő regressziókat hozta:

$$\begin{aligned}\Delta LTC_t &= -0.2276 + 0.01213\Delta BTC_t + -0.03295LTC_{t-1} - 0.0005397 \cdot BTC_{t-1} + W_t'' \\ \Delta BTC_t &= 7.414 + 25.94\Delta LTC_t + 0.000799BTC_{t-1} - 0.2459 \cdot LTC_{t-1} + W_t'''\end{aligned}$$

Ennek a modellnek a reziduálisok nem egyeznek meg az előző modell W_t és W_t' reziduálisával. Tekintettel arra, hogy a továbbakban számos modell felírásában lesz még szükség hibatagok jelölésére, a különbséget nem fogom jelölni, így minden egyenletben mást fog jelenteni W_t ; ami remélhetőleg az olvashatóságot inkább könnyíti, mint nehezíti. Ebben az esetben is R^2 értéke 0.33 körüli. Az első egyenletben a konstans együtthatója csupán 0.0365 p -érték mellett szignifikáns. A második egyenletben az LTC_{t-1} -en kívül semmi nem volt szignifikáns.

Az első modellt tovább bővítettem, még a differencia sorozat első visszaléptetettjeinek bevonásával. Így a következő modellt kaptam:

$$\begin{aligned}\Delta LTC_t &= -0.2116 + 0.01247 \cdot \Delta BTC_t + -0.03318 \cdot LTC_{t-1} + 0.0005340 \cdot BTC_{t-1} \\ &\quad - 0.0004816 \cdot \Delta BTC_{t-1} + 0.1354 \cdot \Delta LTC_{t-1} + W_t\end{aligned}$$

Ez a modell csak egy kicsivel mutat magasabb R^2 értéket, 0.35-öt, ΔBTC_{t-1} nem szignifikáns.

Ha továbbá a differenciasorozaton még jobban visszamegyünk a múltba, a 2. és 3. visszaléptetéssel, akkor R^2 még kicsit nő 0.37-re, illetve 0.39-re, és a ΔBTC_{t-1} tag is szignifikánssá válik.

Az egyenletek 2 illetve 3 visszaléptetéssel:

$$\begin{aligned}\Delta LTC_t &= -0.1952 + 0.01269 \cdot \Delta BTC_t + -0.02628 \cdot LTC_{t-1} + 0.0004337 \cdot BTC_{t-1} \\ &\quad - 0.0001163 \cdot \Delta BTC_{t-1} + 0.1452 \cdot \Delta LTC_{t-1} \\ &\quad + 0.003593 \cdot \Delta BTC_{t-2} - 0.1549 \cdot \Delta LTC_{t-2} + W_t\end{aligned}$$

$$\begin{aligned}\Delta LTC_t &= -0.2046 + 0.01240 \cdot \Delta BTC_t + -0.03201 \cdot LTC_{t-1} + 0.0005157 \cdot BTC_{t-1} \\ &\quad - 0.0001734 \cdot \Delta BTC_{t-1} + 0.1760 \cdot \Delta LTC_{t-1} \\ &\quad + 0.00417 \cdot \Delta BTC_{t-2} - 0.1652 \cdot \Delta LTC_{t-2} \\ &\quad - 0.003269 \cdot \Delta BTC_{t-3} + 0.1334 \cdot \Delta LTC_{t-3} + W_t\end{aligned}$$

Látható, hogy a BTC-hez tartozó tagok együttthatói nagyságrenddel kisebbek. Ebből ugyan úgy tűnik, hogy a Bitcoinnak sokkal kisebbbb hatása van, de az valójában a két kriptovaluta árfolyamának az arányából következik. A Bitcoin hozzávetőlegesen 60-szor erősebb, mint a Litecoin.

Mivel az eddig kapott modellek nem túl erős magyarázó erejűek, ezért két módosítással próbálkoztam:

- csak egy rövidebb intervallumot vizsgáltam, ahol stabilabbnak tűnt a két valuta kapcsolata
- még egy kriptovalutát, az Ethereumot is belevettem a modellbe.

A rövidített intervallumnak az (900 : 1400) elemeket választottam, ami a 2015.10.14. és 2017.02.25. közötti időszakot jelenti.

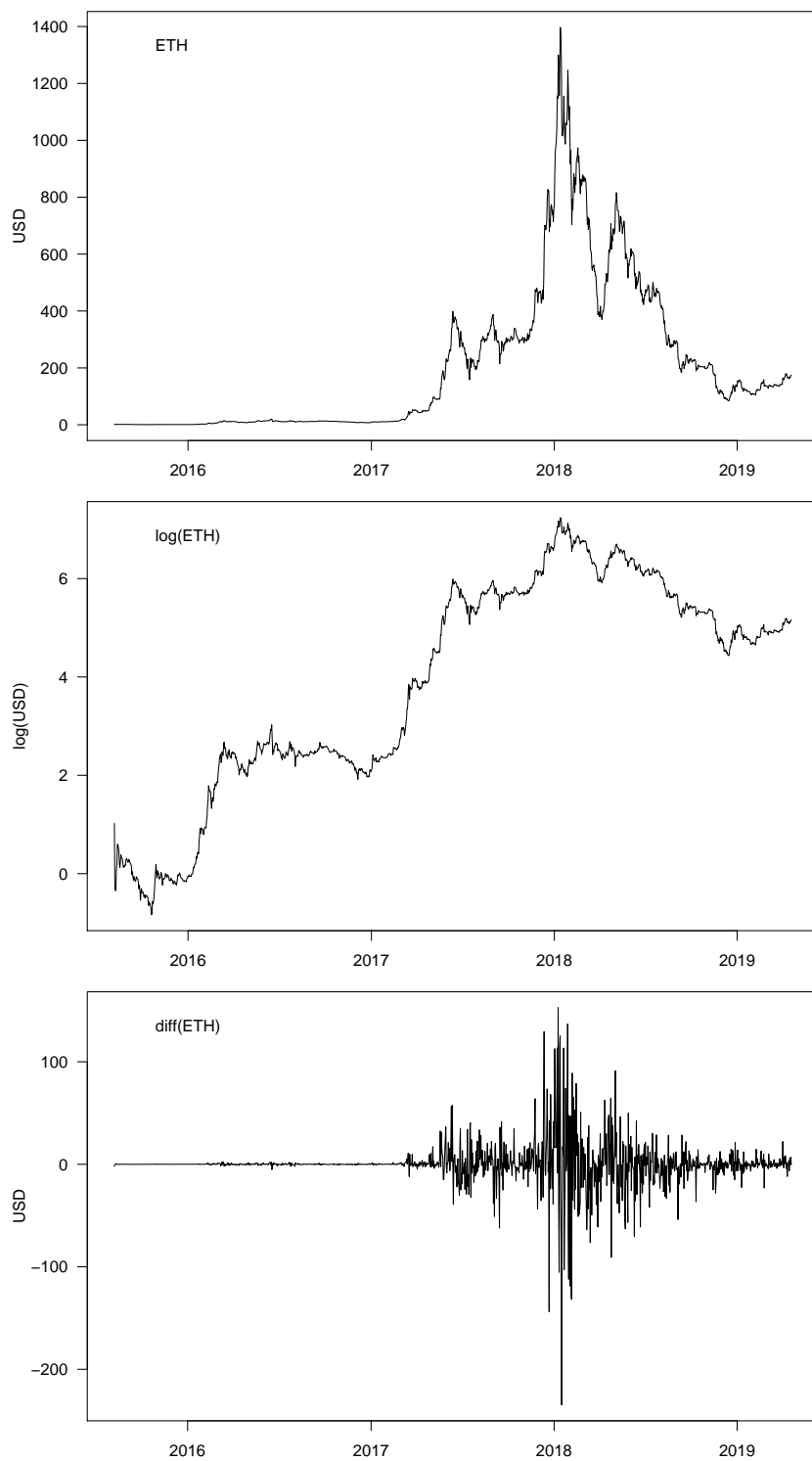
Itt a várakozással ellentétben még rosszabb eredményeket kaptam, ugyan ha csak az eredeti idősorokat, és a sima differencia sorozatokat vettem figyelembe, akkor már $R^2 = 0.54$ lett, de a Bitcoin ára (BTC_{t-1}) már egyáltalán nem volt szignifikáns. Ez a modell:

$$\Delta LTC_t = 0.1171 + 0.005054 \cdot \Delta BTC_t - 0.03189 \cdot LTC_{t-1} - 0.0000089 \cdot BTC_{t-1} + W_t$$

További visszaléptetéseket megvizsgálva az tapasztalható, hogy a visszaléptetett differencia sorozatnak egyáltalán nincsen hatása.

Modellbővítés az Ethereum bevonásával. Tekintettel az eddigi modellek rossz magyarázóerejére, egy harmadik kriptovaluta, az Ethereum, bevonásának a hatását vizsgáltam. Az Ethereum adatsorai viszont csak 2015. augusztusában kezdődnek, így az Ethereumot is magában foglaló modellt nem lehet az eddig használt 2013-ban kezdődő adatokon alapuló modellhez hasonlítani. A BTC és LTC idősorát is megrövidítettem, hogy csak 2015.08.07-től kezdődjenek, és a rövidített idősoron kapott két kriptovalutás modellt hasonlítottam a 3 kriptovalutát használóhoz. A 3 eredeti idősorra 3 féle módon inearis modellt illesztettem, hogy eldöntsem, hogy melyik legyen a függő változó, és a Litecoin mellett döntöttem a reziduálisokra alkalmazott ADF teszt alapján.

Az Ethereum nélküli egy visszaléptetési modellnek a következőt kaptam:



3.6. ábra. Az ETH USD árai, az ár logaritmusa és az ár változása 2015 augusztusa és 2019 áprilisa között

$$\begin{aligned}\Delta LTC_t &= -0.5117 + 0.012356\Delta BTC_t \\ &\quad + 0.0006248BTC_{t-1} - 0.03689LTC_{t-1} \\ &\quad - 0.0005328\Delta BTC_{t-1} + 0.1342845\Delta LTC_{t-1} + W_t\end{aligned}$$

Itt a ΔBTC_{t-1} együtthatója 0.1 szinten sem volt szignifikáns, a konstans p -értéke a második legnagyobb: 0.011, a többi együttható erősen szignifikáns, az R^2 értéke 0.35 erre a modellre.

Az Ethereummal bővített modell a következő:

$$\begin{aligned}\Delta LTC_t &= -0.3717 + 0.005867\Delta BTC_t + 0.1505\Delta ETH_t \\ &\quad + 0.0005050BTC_{t-1} - 0.03606LTC_{t-1} + 0.001259ETH_{t-1} \\ &\quad + 0.002433\Delta BTC_{t-1} + 0.1269\Delta LTC_{t-1} - 0.05667\Delta ETH_{t-1} + W_t\end{aligned}$$

Az Ethereum bevétele a modellbe nagyon megnövelte R^2 értékét: 0.53-ra, és érdekes módon az előző időszakbeli Ethereum ár (ETH_{t-1}) nem volt szignifikáns, míg a többi változó - a konstans kivéve - erősen szignifikáns. Jól látszik, hogy szinte minden, mindkét modellben szereplő változó együtthatója lecsökkent, de irányában és nagyságrendjében nem változott. Kivétel ez alól ΔBTC_{t-1} ami az ETH nélküli modellben nem is volt szignifikáns.

Az ETH nélküli modellbe még egy visszaléptetést belevéve a következő egyenletet kapjuk:

$$\begin{aligned}\Delta LTC_t &= -0.4623 + 0.01257\Delta BTC_t \\ &\quad + 0.0005155BTC_{t-1} - 0.02964LTC_{t-1} \\ &\quad - 0.001199\Delta BTC_{t-1} + 0.1438\Delta LTC_{t-1} \\ &\quad + 0.003536\Delta BTC_{t-2} - 0.1534\Delta LTC_{t-2} + W_t\end{aligned}$$

Itt az $R^2 \approx 0.37$, a konstans és a ΔBTC_{t-1} csak $p = 0.05$ szinten szignifikáns, viszont a többi változó erősen az.

Nézzük ugyanezt a bővítést, az ETH-s modell esetén:

$$\begin{aligned}
\Delta LTC_t = & -0.3583 + 0.006138\Delta BTC_t + 0.1474\Delta ETH_t \\
& + 0.0004604BTC_{t-1} - 0.02929LTC_{t-1} + 0.0004170ETH_{t-1} \\
& + 0.001709\Delta BTC_{t-1} + 0.1170\Delta LTC_{t-1} - 0.04705\Delta ETH_{t-1} \\
& + 0.002201\Delta BTC_{t-2} - 0.1034\Delta LTC_{t-2} - 0.01383\Delta ETH_{t-2} + W_t
\end{aligned}$$

Itt $R^2 \approx 0.54$ -et kaptam, és ETH_{t-1} egyáltalán nem volt szignifikáns, ΔETH_{t-2} p -értéke $\approx \leq 0.076$ esetén). Ezenkívül a konstans, és ΔBTC_{t-1} volt még gyengébben szignifikáns, minden más erősen szignifikánsnak bizonyult.

Megvizsgáltam még, hogy hogyan változik az Ethereummal bővített modell, ha a differencia még egy késleltetettjét hozzáveszem, itt az R^2 0.56-ra nőtt, de ΔETH_{t-3} már egyáltalán nem volt szignifikáns, viszont ΔETH_{t-2} szignifikanciaszintje megjavult $p = 0.006$ körülire.

$$\begin{aligned}
\Delta LTC_t = & -0.3898 + 0.005928\Delta BTC_t + 0.1451\Delta ETH_t \\
& + 0.0005181BTC_{t-1} - 0.03628LTC_{t-1} + 0.001177ETH_{t-1} \\
& + 0.001227\Delta BTC_{t-1} + 0.1455\Delta LTC_{t-1} - 0.04841\Delta ETH_{t-1} \\
& + 0.002998\Delta BTC_{t-2} - 0.1044\Delta LTC_{t-2} - 0.02142\Delta ETH_{t-2} \\
& - 0.002825\Delta BTC_{t-3} - 0.1559\Delta LTC_{t-3} - 0.008319\Delta ETH_{t-3} + W_t
\end{aligned}$$

Összefoglalva a fenti elemzés eredményét, az alábbi következtetésre juthatunk. A statisztikák alapján egy ECM Modell alkalmazása indokolt. Csak Bitcoin és Litecoin adatok használatával nem érhetünk el túl nagy magyarázóerőt. Viszont az Ethereummal bővített modell differenciáiban két és három visszaléptetési változata az adatok érvényes modelljének tekinthető.

Irodalomjegyzék

- [1] *Bitcoin price data*, <https://coinmarketcap.com/currencies/bitcoin/historical-data/?start=20130428&end=20190420>, letöltve: 2019.04.22.
- [2] *Cointegration and error correction model*, <https://www.researchgate.net/file.PostFileLoader.html?id=4f6053ca80e5828651000000&assetKey=AS%3A271750879744001%401441801866418>, letöltve:2019.05.08.
- [3] *Ethereum price data*, <https://coinmarketcap.com/currencies/ethereum/historical-data/?start=20130428&end=20190420>, letöltve:2019.05.10.
- [4] *Litecoin price data*, <https://coinmarketcap.com/currencies/litecoin/historical-data/?start=20130428&end=20190420>, letöltve: 2019.04.22.
- [5] Bernardina Algieri, *Time series 5: ECM (lecture notes)*, http://www.ecostat.unical.it/Algieri/Didattica/Financial%20Markets/Tutorials/Lecture_L5_Time%20Series%20ECM.pdf, letöltve:2019.05.18.
- [6] George Athanasopoulos, DS Poskitt, and Farshid Vahid, *Two canonical varma forms: Scalar component models vis-à-vis the echelon form*, *Econometric Reviews* **31** (2012), no. 1, 60–83.
- [7] George Athanasopoulos and Farshid Vahid, *A complete varma modelling methodology based on scalar components*, *Journal of Time Series Analysis* **29** (2008), no. 3, 533–554.
- [8] Yin-Wong Cheung and Kon S Lai, *Lag order and critical values of the augmented Dickey-Fuller test*, *Journal of Business & Economic Statistics* **13** (1995), no. 3, 277–280.

- [9] Thomas F Cooley and Mark Dwyer, *Business cycle analysis without much theory a look at structural vars*, Journal of econometrics **83** (1998), no. 1-2, 57–88.
- [10] David A Dickey and Wayne A Fuller, *Distribution of the estimators for autoregressive time series with a unit root*, Journal of the American Statistical Association **74** (1979), no. 366, 427–431.
- [11] Robert F Engle and Clive WJ Granger, *Co-integration and error correction: representation, estimation, and testing*, Econometrica: Journal of the Econometric Society (1987), 251–276.
- [12] Clive WJ Granger, *Some properties of time series data and their use in econometric model specification*, Journal of Econometrics **16** (1981), no. 1, 121–130.
- [13] Søren Johansen, *Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models*, Econometrica: Journal of the Econometric Society (1991), 1551–1580.
- [14] Helmut Lütkepohl, *New introduction to multiple time series analysis*, Springer Science & Business Media, 2005.
- [15] Debin Qiu, *aTSA: Alternative time series analysis*, 2015, R package version 3.1.2.
- [16] R Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018.
- [17] George C Tiao and Ruey S Tsay, *Model specification in multivariate time series*, Journal of the Royal Statistical Society: Series B (Methodological) **51** (1989), no. 2, 157–195.