

Универзитет у Београду
Математички факултет



МАСТЕР РАД

Неки нелинеарни модели временских серија са применом у финансијама

Ментор:
др Јелена Јоцковић

Студент:
Јована Недељковић 1035/2016

Београд, септембар 2019.

САДРЖАЈ

1. Увод	2
2. Финансијске временске серије и њихове особине	3
2.1. Принос активе.....	3
2.2. Карактеристике финансијских временских серија	5
3. Нелинеарни модели временских серија	8
3.1. Структура нелинеарних модела.....	8
3.2. Тестови нелинеарности.....	10
3.3. Општи нелинеарни модели	15
3.4. Уопштени <i>AR (GAR)</i> модели	16
3.5. Билинеарни модели.....	17
3.6. Још неки параметарски модели	21
3.7. Модели са случајним коефицијентима	22
3.8. Модели временских серија са праговима	26
3.8.1. <i>TAR</i> и <i>SETAR</i> модели	26
3.8.2. <i>STAR</i> модели.....	46
3.8.3. Још неки <i>TAR</i> модели	50
4. Закључак.....	52
5. Литература.....	53

1. Увод

Временске серије представљају уређени низ опсервација у односу на време. Оне су присутне у разним научним областима, а у овом раду посебна пажња ће бити посвећена финансијским временским серијама.

Финансијске временске серије описују промене вредности финансијских инструмената током времена, а у пракси ове временске серије имају најчешће бар једну од следећих пет особина:

1. Присуство сезонске компоненте и тренда
2. Нестабилна дисперзија
3. Постојање структурног лома
4. Асиметрија (на пример, негативне вредности се појављују чешће него позитивне)
5. Кластери (нестандардне опсервације се појављују у групама)

Основни циљ анализе временских серија јесте да се на основу историјских података прогнозира будуће понашање неке појаве. Да би се серија успешно користила за предвиђање понашања и контролу неке појаве, важно је одабрати адекватан модел. Серије код којих су присутни асиметрија и кластери не могу се добро описати помоћу стандардних линеарних модела (ауторегресиони модели покретних просека), па се зато уводе нелинеарни модели – они код којих је тренутна вредност серије нелинеарна функција претходних опсервација и белог шума.

Циљ овог мастер рада јесте упознавање са неким класама нелинеарних модела временских серија, као што су билинеарни модели, модели са праговима и модели са случајним коефицијентима. У раду су објашњене специфичности финансијских временских серија, дефинисани наведени нелинеарни модели и изведене њихове особине. Поред тога, рад је допуњен примерима који представљају симулације нелинеарних модела, али и примерима који представљају примене ових модела на конкретне финансијске временске серије.

2. Финансијске временске серије и њихове особине

2.1. Принос активе

Активa представља ресурс са економском вредношћу који поседује или контролише појединац, корпорација или држава, са очекивањем да ће му тај ресурс у будућности донети корист. **Финансијска активa** (финансијски инструменти) представља предмет инвестирања на финансијском тржишту и она може имати форму хартија од вредности и кредита. **Хартије од вредности (ХоВ)** представљају финансијску активу која се може прометовати на тржишту, а принос по основу ХоВ може бити у виду камате и купона (обвезнице), дивиденде и капиталног добитка (акције) и други облици добити за различите финансијске инструменте. **Обвезнице** (дуговне ХоВ) су вредносни папири са фиксим приходом и представљају обавезу за онога ко их је издао да плаћа новац ономе ко поседује обвезнице, у складу са правилима која су одређена у тренутку издавања обвезница. **Акције** (власничке ХоВ) представљају вредносне папире са приходом који није фиксан и означавају удео у капиталу неке компаније. Оне имају своју цену и са њима се може трговати на берзи. Власницима акција се уплаћују **дивиденде** које представљају део добити, односно профит акционарског друштва, а деле се процентуално броју акција. [6]

У већини истраживања која се односе на финансијске временске серије, уместо цена се користе приноси активе јер они дају потпуни увид инвеститорима у потенцијалне инвестиције. Осим тога, лакше је анализирати серије које се односе на принос активе јер имају боље статистичке особине у односу на исте серије које се односе на цене активе. [1]

Нека је P_t цена активе у временском тренутку t и претпоставимо да активa не исплаћује дивиденде.

Дефиниција 1. Ако имамо активу један временски период, од тренутка $t - 1$ до тренутка t , тада је стопа укупног приноса:

$$1 + R_t = \frac{P_t}{P_{t-1}}$$

односно:

$$P_t = P_{t-1}(1 + R_t)$$

Одговарајућа стопа приноса за један временски период дефинише се са:

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Дефиниција 2. Ако неку активу имамо k периода, тј. од тренутка $t - k$ до тренутка t , тада је стопа укупног приноса:

$$\begin{aligned} 1 + R_t[k] &= \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \cdot \frac{P_{t-1}}{P_{t-2}} \cdot \dots \cdot \frac{P_{t-k+1}}{P_{t-k}} \\ &= (1 + R_t) \cdot (1 + R_{t-1}) \cdot \dots \cdot (1 + R_{t-k+1}) \\ &= \prod_{j=0}^{k-1} (1 + R_{t-j}). \end{aligned}$$

Одавде уочавамо да је стопа укупног приноса за k периода једнака производу k стопа приноса за један период.

Стопа приноса за k периода је: $R_t[k] = \frac{P_t - P_{t-k}}{P_{t-k}}$.

У пракси је веома битан временски период за који се рачуна принос активе. Уколико период није прецизно дат, претпостављамо да је то једна година.

Дефиниција 3. **Стопа годишњег (просечног) приноса** дефинише се са:

$$\text{Годишња } R_t[k] = \left[\prod_{j=0}^{k-1} (1 + R_{t-j}) \right]^{\frac{1}{k}} - 1.$$

Ово је геометријска средина стопе укупног приноса за k периода, а она се може записати и на други начин:

$$\text{Годишња } R_t[k] = e^{\left[\frac{1}{k} \prod_{j=0}^{k-1} \ln(1 + R_{t-j}) \right]} - 1.$$

Помоћу Тејлорове формуле може да се изврши апроксимација за годишњу стопу приноса:

$$\text{Годишња } R_t[k] \approx \frac{1}{k} \sum_{j=0}^{k-1} R_{t-j}.$$

Ипак, у неким применама ова апроксимација није довољно прецизна.

Дефиниција 4. Природни логаритам простог приноса активе назива се **непрекидно камаћени принос** или **лог принос**, а рачуна се према следећој формули:

$$r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1},$$

где је $p_t = \ln P_t$.

Непрекидно камаћени принос за k периода је:

$$\begin{aligned}r_t[k] &= \ln(1 + R_t[k]) = \ln[(1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k+1})] \\ &= \ln(1 + R_t) + \ln(1 + R_{t-1}) + \dots + \ln(1 + R_{t-k+1}) \\ &= r_t + r_{t-1} + \dots + r_{t-k+1}.\end{aligned}$$

Дакле, непрекидно камаћени принос за више периода је заправо сума непрекидно камаћених приноса за један период. Статистичке особине лог приноса су погодније у анализи, па се самим тим и више користе.

2.2. Карактеристике финансијских временских серија

Финансијске временске серије описују промене вредности финансијских инструмената током времена, као што су, на пример, кретање цена на тржишту, годишња производња, извоз, кретање акција на берзи и слично. Ове временске серије у пракси поседује неку од следећих особина:

1. Тренд

Дугорочна компонента у кретању временске серије назива се тренд. Уколико серија системски расте или опада током времена, то значи да серија има растући или опадајући тренд који може бити детерминистички или стохастички, у зависности од тога да ли се кретање временске серије може предвидети на основу прошлих вредности.

2. Сезонска компонента

Уколико временска серија испољава правилности у кретању у току једне календарске године тада се она назива сезонска временска серија. Постојање ове компоненте указује да постоји већи степен корелације између опсервација истих месеци у различитим годинама него између суседних месеци.

3. Структурни лом

Структурни лом представља скуп опсервација које одступају од претходног тока временске серије. Он је најчешће резултат неке интервенције, у смислу догађаја који ће утицати на кретање временске серије.

4. Нестационарност временских серија цена

Кретање цена углавном прати путању случајног лутања, за разлику од кретања приноса. Временске серије приноса генерално задовољавају услов стационарности, при чему волатилност приноса представља степен осцилација приноса од средње вредности.

5. Нестабилна дисперзија

Цене финансијских инструмената најчешће нису стабилне. До промена цена акција на тржишту долази услед нових информација које стално пристижу, а степен промене цена акција зависи од тога како на новопрстиглу вест инвеститори гледају (уколико је вест негативна, варијабилност расте). Реакција на неку неочекивану новост може довести до настанка волатилности.

У економском смислу, волатилност финансијског инструмента говори о томе колико се мења цена тог инструмента у неком протеклом периоду. Волатилност најчешће рачунамо као стандардну девијацију промене цене.

Један од показатеља ризика јесте волатилност, а ризик је већи што је већа волатилност одређеног финансијског инструмента. Многи финансијски модели укључују независно оцењену или предвиђену волатилност (нпр. *Black-Scholes* формула). За оцењивање волатилности најбоље је користити информације о прошлим вредностима стопе приноса. Још једна од карактеристика волатилности је та да волатилност целог тржишта утиче на волатилност сваке појединачне акције.

6. Одсуство аутокорељације у временским серијама приноса и аутокорељација квадрата приноса

Финансијске временске серије које прате кретање дневних приноса генерално показују мали степен аутокорељације, док квадрати приноса, са друге стране, често имају више или мање изражену аутокорељацију која има тенденцију лаганог опадања.

7. Груписање волатилности - кластери

Приноси као финансијске временске серије имају изражену тенденцију груписања волатилности. Током периода које карактерише виша стопа ризика, стопе приноса испољавају високу волатилност (турбулентни период), да би се у периодима ниже стопе ризика волатилност смањила.

8. Одступање од нормалне расподеле

За разлику од нормалне расподеле, чији коефицијент спљоштености има вредност 3, емпиријска истраживања су показала да финансијске временске серије карактерише расподела са реповима расподеле који су тежи од репова нормалне расподеле (коефицијент спљоштености је већи од 3), који указују на већи ризик

услед веће вероватноће појаве екстремних вредности. Наведена појава израженија је код дневних приноса, док се расподела недељних и месечних приноса више приближава нормалној расподели.

Код финансијских временских серија уочено је да обично имају емпиријску расподелу асиметричну удесно, тј. да ове временске серије најчешће системски расту током времена.

9. Левериџ ефекат (енгл. *leverage effect*)

Финансијске временске серије које прате кретање приноса често показују различите реакције на позитивне, односно негативне вести. Појава нове вести утиче на повећање волатилности посматране временске серије, при чему је чест случај да негативне вести у већој мери утичу на волатилност него што је то случај са позитивним вестима. Овај вид асиметрије када је у питању испитивање финансијских временских серија назива се левериџ ефекат.

10. Календарски ефекти

Честа карактеристика финансијских временских серија је да се приноси могу разликовати у зависности од дана у недељи, дана у месецу, као и месеца у години. Календарски ефекти дана у недељи јављају се као последица чињенице да приноси понедељком, за разлику од осталих радних дана, представљају 72-часовне приносе, па би се могло очекивати да ће понедељком приноси бити јаче изражени. Међутим, у пракси се показало да просечни приноси понедељком често имају нижу вредност од приноса осталим данима.

3. Нелинеарни модели временских серија

3.1. Структура нелинеарних модела

Циљ анализе временске серије X_t јесте да се испита њена динамичка зависност. Временске серије код којих су присутни асиметрија или кластери, не могу добро да се опишу помоћу стандардних линеарних модела (ауторегресиони модели покретних просека), што је довело до потребе за развојем нелинеарних модела – они код којих је тренутна вредност серије нелинеарна функција претходних опсервација и белог шума. Нелинеарни модели су веома корисни, а у односу на линеарне моделе пружају већу прецизност у предвиђању.

У наставку ће бити представљени нелинеарни модели који се користе за анализу финансијских временских серија, као и тестови за испитивање нелинеарности. За писање овог поглавља коришћена је литература [1] и [16].

Посматрајмо временску серију X_t , чије су опсервације дате у временским тренуцима са једнаким временским размаком. Уочимо опсервације $\{x_t | t = 1, 2, \dots, T\}$, где је T обим узорка.

Дефиниција 5. За чисто стохастичку временску серију X_t кажемо да је линеарна временска серија ако можемо да је представимо на следећи начин:

$$X_t = \mu + \sum_{i=0}^{\infty} \psi_i e_{t-i} \quad (1)$$

где је μ константа, ψ_i су реални бројеви, при чему је $\psi_0 = 1$ и $\{e_t\}$ је низ независних једнако расподељених случајних величина са датом функцијом расподеле вероватноћа.

Претпоставићемо да је функција расподеле вероватноћа за $\{e_t\}$ непрекидна и да важи $E(e_t) = 0$. У многим сличајевима се претпоставља да је $D(e_t) = \sigma^2$, или чак да је e_t Гаусовски процес. Ако је $\sigma^2 \sum_{i=1}^{\infty} \psi_i^2 < +\infty$, тада је X_i слабо стационаран процес.

За сваки стохастички процес који не задовољава услов (1) кажемо да је нелинеаран.

Ова дефиниција нелинеарности важи за чисто стохастичке временске серије. Модел чисто стохастичке временске серије X_t је функција низа независних и једнако расподељених случајних величина:

$$X_t = f(e_t, e_{t-1}, \dots) .$$

Из линеарности модела у једнакости (1) следи да је $f(\cdot)$ линеарна функција својих аргумената. Било каква нелинеарност функције $f(\cdot)$ резултује нелинеарни модел.

Нелинеарни модел $X_t = f(e_t, e_{t-1}, \dots)$ није директно применљив јер садржи превише параметара. Да би нелинеарни модел био доступан, записујемо га преко условних момената.

Нека је \mathcal{F}_{t-1} σ –поље генерисано доступним информацијама у тренутку $t - 1$. Тада \mathcal{F}_{t-1} означава линеарну комбинацију елемената из $\{X_{t-1}, X_{t-2}, \dots\}$ и $\{e_{t-1}, e_{t-2}, \dots\}$.

Условно очекивање и дисперзија за X_t под условом \mathcal{F}_{t-1} су:

$$\mu_t = E(X_t | \mathcal{F}_{t-1}) \equiv g(\mathcal{F}_{t-1})$$

$$\sigma_t^2 = D(X_t | \mathcal{F}_{t-1}) \equiv h(\mathcal{F}_{t-1})$$

при чему су $g(\cdot)$ и $h(\cdot) > 0$ добро дефинисане функције.

Тако имамо модел:

$$X_t = g(\mathcal{F}_{t-1}) + \sqrt{h(\mathcal{F}_{t-1})} \epsilon_t, \quad \epsilon_t = e_t / \sigma_t .$$

Ако је $g(\cdot)$ нелинеарна функција, за X_t се каже да је нелинеарна у очекивању. Ако функција $h(\cdot)$ варира током времена, за X_t се каже да је нелинеарна у дисперзији.

Постоје дилеме када треба изабрати линеарни, а када нелинеарни модел. Одговарајућим тестовима може да се тестира присуство нелинеарности. Уколико се прихвати хипотеза о нелинеарности, потребно је направити низ избора како би се дошло до адекватног нелинеарног модела.

Не постоји утврђена процедура за избор адекватног модела, а дилеме које се јављају тичу се избора класе модела:

- параметарски или непараметарски модел
- стохастички или детерминистички модел
- временски или фреквентни домен.

У наставку су описани неки стандардни тестови за испитивање нелинеарности и уведени најзаступљенији нелинеарни модели.

За нелинеарну временску серију X_t , услов за ергодичност је много теже успоставити. Услови су обично специфични за модел, али често се могу извести помоћу теорије ланца Маркова, јер већина нелинеарних модела има својство Маркова.

3.2. Тестови нелинеарности

Пре увођења нелинеарних модела, важно је упознати се са неколико стандардних тестова за испитивање присуства нелинеарности. У наставку су описани следећи тестови: тест *McLeod-Li* (*Ljung-Box* статистика квадрата резидуала), тест регресионе спецификације грешке (*RESET*), *BDS* тест и тестови Лагранжовог множиоца. За писање овог поглавља је коришћена литература [1], [2] и [3].

Ови тестови могу да потврде само присуство нелинеарности, али не и да утврде о ком облику нелинеарности је реч.

1. АУТОКОРЕЛАЦИОНА ФУНКЦИЈА (*ACF*) И *MCLEOD-LI* ТЕСТ

Код изградње *ARMA* (p, q) модела, аутокорелативна функција може да помогне приликом одређивања вредности редова p и q , а аутокорелативна функција резидуала је важан дијагностички алат.

Аутокорелативна функција (*ACF*) која се користи код линеарних модела не може да се примењује за нелинеарне моделе, јер коефицијенти аутокорелативне мере степен линеарне зависности између X_t и X_{t-i} , и као таква, аутокорелативна функција можда неће успети да открије важне нелинеарне везе које су присутне у подацима.

Размотримо следећи пример:

$$X_t = e_{t-1}^2 + e_t$$

где је $\{e_t\}$ бели шум са нормалном расподелом $\mathcal{N}(0, \sigma^2)$. Дакле, X_t је функција од e_{t-1} и вредност X_t зависи од X_{t-1} .

Како је $D(e_t) = D(e_{t-1}) = \sigma^2$, добијамо:

$$EX_t = EX_{t-1} = \sigma^2 .$$

Аутоковаријациони коефицијент са кашњењем i рачунамо на следећи начин:

$$\begin{aligned} \gamma_i &= E(X_t - \sigma^2)(X_{t-i} - \sigma^2) = E(e_{t-1}^2 + e_t - \sigma^2)(e_{t-1-i}^2 + e_{t-i} - \sigma^2) \\ &= E(e_{t-1}^2 e_{t-1-i}^2 + e_{t-1}^2 e_{t-i} - e_{t-1}^2 \sigma^2 + e_t e_{t-1-i}^2 + e_t e_{t-i} - e_t \sigma^2 - \sigma^2 e_{t-1-i}^2 - \sigma^2 e_{t-i} + \sigma^2 \sigma^2) \end{aligned}$$

Како је: $E(e_{t-i}^2 e_{t-j}^2) = \sigma^2 \sigma^2 = \sigma^4$, $E(e_t e_{t-i}^2) = 0$ и $E(e_t \sigma^2) = 0$, добијамо:

$$\gamma_i = \sigma^2 \sigma^2 + E(e_{t-1}^2 e_{t-i}) - \sigma^2 \sigma^2 + E(e_t e_{t-1-i}^2) - \sigma^2 \sigma^2 + \sigma^2 \sigma^2$$

$$\gamma_i = E(e_{t-1}^2 e_{t-i}) .$$

Јасно је да је $E(e_{t-1}^2 e_{t-i}) = 0$ за $i \neq 1$. Осим тога, када e_t има нормалну расподелу, важи:

$$E(e_t^3) = E(e_{t-1}^3) = 0$$

тако да су све вредности аутокорељационих коефицијената $\rho_i = 0$, $i \neq 0$.

Дакле, ако бисмо посматрали аутокорељациону функцију за $\{X_t\}$, а да не знамо да је $X_t = e_{t-1}^2 + e_t$, на основу запажања да су аутокорељације узорка мале, погрешно бисмо закључити да је серија бели шум. Иако су аутокорељације нула, вредност X_t јасно зависи од вредности X_{t-1} .

Како испитујемо нелинеарне односе временских серија, корисно је испитати и аутокорељациону функцију за квадратне или кубне вредности серије. На пример, аутокорељациона функција за X_t^2 може открити нелинеарни образац понашања временске серије. Дакле, свака занемарена нелинеарност у подацима се може проверити помоћу аутокорељационе функције за квадратне (или кубне) вредности серија.

Тест McLeod–Li настоји да утврди да ли постоје значајне аутокорељације код квадрата резидуала линеарног модела. Да би се извршио тест, потребно је да се изгради најпогоднији линеарни модел и затим да се одреде резидуали \hat{e}_t .

Нека ρ_i означава узорачки коефицијент корелације квадрата резидуала \hat{e}_t^2 и \hat{e}_{t-i}^2 и користећи *Ljung–Box* статистику, утврдићемо да ли квадрати резидуала показују серијску корелацију. Дакле, рачунамо:

$$Q = T(T+2) \sum_{i=1}^n \frac{\rho_i^2}{(T-i)},$$

где је T обим узорка, а n адекватно одабран број аутокорељација који се користи у тесту.

Вредност Q има асимптотску χ^2 расподелу са n степени слободе ако је $\{\hat{e}_t^2\}$ низ некорелисан. Одбацивање нулте хипотезе је еквивалентно прихватању алтернативне хипотезе да је модел нелинеаран.

Алтернативно, може да се посматра следећи линеарни модел:

$$\hat{e}_t^2 = \alpha_0 + \alpha_1 \hat{e}_{t-1}^2 + \dots + \alpha_n \hat{e}_{t-n}^2 + v_t.$$

Ако нема нелинеарности, $\alpha_1 = 0, \dots, \alpha_n = 0$. Код узорка од T резидуала, ако не постоји нелинеарност, тест статистика TR^2 ће конвергирати ка χ^2 расподели са n степени слободе. Код малих узорака може да се користи F -тест за нулту хипотезу $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$. Одбацивањем нулте хипотезе о линеарности не долазимо до облика нелинеарности која је присутна у подацима.

2. ТЕСТ СПЕЦИФИКАЦИЈЕ РЕГРЕСИОНЕ ГРЕШКЕ (*RESET* тест)

Тест спецификације регресионе грешке (*RESET*) такође поставља нулту хипотезу о линеарности наспрам опште алтернативне хипотезе о нелинеарности. Ако су резидуали из линеарног модела независни, они не би требало да буду у корелацији са регресорима који се користе или са фитованим вредностима модела.

Овај тест се састоји из два корака:

КОРАК 1: Бира се најпогоднији линеарни модел за X_t . Нека су $\{e_t\}$ резидуали модела и нека \hat{X}_t означава фитоване вредности.

На пример, нека је $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + e_t$. Након оцењивања непознатих параметара можемо да израчунамо фитоване вредности и резидуале:

$$\hat{X}_t = \phi_1 \hat{X}_{t-1} + \phi_2 \hat{X}_{t-2} + \dots + \phi_p \hat{X}_{t-p} \text{ и } \hat{e}_t = X_t - \hat{X}_t.$$

Суму квадрата резидуала означавамо са: $SSR_0 = \sum_{t=p+1}^T \hat{e}_t^2$, где је T обим узорка.

КОРАК 2: Бира се вредност H (обично 3 или 4) и посматра се следећи линеарни модел:

$$\hat{e}_t = \delta Z_t + \sum_{h=2}^H \alpha_h \cdot \hat{X}_t^h + v_t, \quad H \geq 2,$$

где је Z_t вектор који садржи променљиве укључене у модел из првог корака. У примеру из корака 1, Z_t ће садржати X_{t-1}, \dots, X_{t-p} . Одавде се добија:

$$\hat{v}_t = \hat{e}_t - \delta Z_t - \sum_{h=2}^H \hat{\alpha}_h \cdot \hat{X}_t^h \text{ и } SSR_1 = \sum_{t=p+1}^T \hat{v}_t^2.$$

Уколико је линеарни модел из првог корака адекватан, треба да важи $\delta = 0$ и $\alpha_h = 0$, за $h = 2, \dots, H$. Тест статистика:

$$F = \frac{\frac{SSR_0 - SSR_1}{g}}{\frac{SSR_1}{T-p-g}}, \quad g = H - 1 + p + 1,$$

под претпоставком нормалности и линеарности ће имати F расподелу са степенима слободе g и $T - p - g$.

RESET тест је једноставан за имплементацију јер не захтева оцењивање великог броја параметара и има разумну моћ да детектује неке типове нелинеарности. Међутим, како тест користи целокупну моћ фитованих вредности, он има малу моћ да детектује асиметричне моделе (као што је модел са прагом).

3. ОСТАЛИ *PORTMANTEAU* ТЕСТОВИ

Portmanteau тестови се обично односе на тестове засноване на резидуалима који немају специфичну алтернативну хипотезу. *Ljung–Box Q*-статистика је добар пример оваквог типа теста. Слично томе, популаран је *Brock, Dechert, Scheinkman* и *LaBarron* тест, назван *BDS* тест, који је *portmanteau* тест за испитивање независности. У суштини, тест испитује удаљеност између различитих парова резидуала.

Нека d представља дату раздаљину и нека су e_t и e_{t-1} две реализације $\{e_t\}$ серије. Ако су све вредности $\{e_t\}$ независне, онда вероватноћа да је растојање између било ког пара резидуала (e_i, e_j) мања од d треба да буде иста за све i и j . *BDS* тест је у стању да детектује серијску корелацију, нестабилност параметара, занемарену нелинеарност, структурне ломове и друге проблеме у вези са погрешном спецификацијом. Одбацавање нулте хипотезе о независности мало помаже у идентификовању природе проблема. Такође *BDS* тест нема посебно добре перформансе код малих узорака.

McLeod–Li тест, *RESET* и други *portmanteau* тестови имају веома општу алтернативну хипотезу. Као такви, тестови су корисни у одређивању да ли је нелинеарни модел прикладан, али не у одређивању природе нелинеарности. Промена параметара се појављује у многим облицима и може проузроковати значајну грешку у прогнози када се модели користе у пракси. Они такође утврђују да је тешко разликовати погрешне спецификације модела од проблема неконстантних параметара. Као такво, вредно је испитати тестове Лагранжовог множиоца за нелинеарност, јер они имају специфичну нулту хипотезу и специфичну алтернативну хипотезу.

4. ТЕСТОВИ ЛАГРАНЖОВОГ МНОЖИОЦА

Тестови Лагранжовог множиоца (*LM*) могу да се користе за тестирање специфичног типа нелинеарности. Дакле, *LM* тест може да помогне код избора одговарајућег функционалног облика који се користи код нелинеарних модела. Нека је $f(\cdot)$ нелинеарна функционална форма, нека α означава параметре функције $f(\cdot)$ и претпоставићемо да је $D(e_t) = \sigma^2$ константна.

Под овим претпоставкама, *LM* тест се може спровести на следећи начин:

КОРАК 1: Одређује се линеарни део модела како би се добиле вредности резидуала $\{e_t\}$.

КОРАК 2: Одређују се сви парцијални изводи $\frac{\partial f}{\partial \alpha}$ који се рачунају у оквиру нулте хипотезе о линеарности. Ови парцијални изводи ће бити нелинеарне функције регресора коришћених у првом кораку. Након тога се формира модел за e_t , где су регресори ови парцијални изводи.

КОРАК 3: Ако је нулта хипотеза тачна, TR^2 има χ^2 расподелу са степенима слободe једнаким броју регресора коришћених у кораку 2. Ако израчуната вредност TR^2 прелази критичну вредност (из таблице χ^2 расподеле), одбацује се нулта хипотеза о линеарности и прихвата се алтернативна. За мали узорак, користи се F -тест.

Једна од предности овог теста је да не мора да се одређује сам нелинеарни модел. Дакле, коришћење LM тестова може да помогне код избора облика нелинеарности.

На пример, ако желимо да проверимо да ли је X_t облика:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-1} X_{t-2} + e_t,$$

у првом кораку је потребно да се одреди линеарни део модела (то је $AR(2)$ модел) и да се одреде резидуали $\{e_t\}$.

У другом кораку се рачунају парцијали изводи:

$$\frac{\partial X_t}{\partial \phi_0} = 1, \quad \frac{\partial X_t}{\partial \phi_1} = X_{t-1}, \quad \frac{\partial X_t}{\partial \phi_2} = X_{t-2}, \quad \frac{\partial X_t}{\partial \phi_3} = X_{t-1} X_{t-2}.$$

Формира се модел за e_t на следећи начин:

$$e_t = a_0 + a_1 X_{t-1} + a_2 X_{t-2} + a_3 X_{t-1} X_{t-2} + v_t.$$

Рачуна се TR^2 и ако вредност тест статистике прелази критичну вредност за χ^2 расподелу са 4 степена слободe, одбацује се нулта хипотеза о линеарности и прихвата се алтернативна.

3.3. Општи нелинеарни модели

Нелинеарни AR модел (NAR) представља природно уопштење AR модела и може да се прикаже на следећи начин:

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-p}, e_t),$$

где је $f: \mathbb{R}^{p+1} \rightarrow \mathbb{R}$ нелинеарна функција, а $\{e_t\}$ строги бели шум (низ независних случајних величина са истом расподелом), при чему су e_t и X_s независне случајне величине за $s < t$.

Специјални случајеви су када је грешка e_t адитивна и тада је:

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-p}) + e_t,$$

односно када је мултипликативна:

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-p}) \cdot e_t.$$

Модели са адитивном грешком служе за моделирање условног очекивања, а модели са мултипликативном грешком за моделирање условне дисперзије која је променљива.

У моделе са адитивном грешком спадају $NAAR$ модели (*Nonlinear Additive AR*):

$$X_t = f_1(X_{t-1}) + f_2(X_{t-2}) + \dots + f_p(X_{t-p}) + e_t,$$

као и FAR (*Functional coefficient AR*):

$$X_t = f_1(X_{t-d})X_{t-1} + f_2(X_{t-d})X_{t-2} + \dots + f_p(X_{t-d})X_{t-p} + e_t.$$

Аналогно се дефинишу општи NMA (*Nonlinear MA*) и $NARMA$ (*Nonlinear ARMA*) модели.

Модел који обухвата све наведене моделе је облика:

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-p}, e_t, e_{t-1}, \dots, e_{t-q}).$$

Најједноставнији облик нелинеарног AR модела је NAR (1) који је дефинисан са:

$$X_t = f(X_{t-1}) + e_t,$$

и који је могуће параметризовати на следећи начин:

$$X_t = a_1(X_{t-1}) \cdot X_{t-1} + e_t,$$

где је $a_1(X_{t-1}) \cdot X_{t-1} \equiv f(X_{t-1})$.

Примећујемо да последња једнакост има исти облик као $AR(1)$ модел, при чему је допуштено да ауторегресиони коефицијент a_1 буде функција од X_{t-1} .

Посматрајмо сада општи нелинеарни модел $NAR(p)$:

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-p}) + e_t.$$

Параметризацијом функције f добијамо различите NAR моделе.

На пример, општи параметарски модел за условно очекивање и условну дисперзију дефинише се на следећи начин:

$$X_t = f_1(X_{t-1}, X_{t-2}, \dots, X_{t-p}, \theta_1) + f_2(X_{t-1}, X_{t-2}, \dots, X_{t-p}, \theta_2) \cdot e_t, \quad f_2 > 0$$

где су f_1 и f_2 познате функције, а $\theta_1 = (\theta_{1,1}, \dots, \theta_{1,p})^T$ и $\theta_2 = (\theta_{2,1}, \dots, \theta_{2,p})^T$ вектори непознатих параметара. Тада су:

$$E(X_t | X_{t-1} = x_1, \dots, X_{t-p} = x_p) = f_1(x_1, \dots, x_p, \theta_1)$$

$$D(X_t | X_{t-1} = x_1, \dots, X_{t-p} = x_p) = f_2^2(x_1, \dots, x_p, \theta_2) \cdot D(e_t).$$

У наставку су наведене најпознатије класе параметарских модела.

3.4. Уопштени AR (GAR) модели

Посматрајмо општи нелинеарни модел $NAR(p)$:

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-p}) + e_t.$$

Проблем који се јавља код одређивања овог модела јесте тај што је функција f непозната. У том случају користимо Тејлоров развој за непознату функцију f . На пример, за $NAR(2)$ модел $X_t = f(X_{t-1}, X_{t-2}) + e_t$ Тејлоров развој је следећег облика:

$$\begin{aligned} X_t = & a_0 + a_1 X_{t-1} + a_2 X_{t-2} + a_{12} X_{t-1} X_{t-2} + a_{11} X_{t-1}^2 + a_{22} X_{t-2}^2 + \\ & + a_{112} X_{t-1}^2 X_{t-2} + a_{122} X_{t-1} X_{t-2}^2 + a_{111} X_{t-1}^3 + a_{222} X_{t-2}^3 + e_t. \end{aligned}$$

У општем случају, Тејлоров развој функције f може да се напише у следећем облику:

$$X_t = a_0 + \sum_{i=1}^p a_i X_{t-i} + \sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^r \sum_{l=1}^s a_{ijkl} X_{t-i}^k X_{t-j}^l + e_t, \quad r, s \geq 1$$

где су r, s цели бројеви, а како би се избегао велики број параметара обично се ставља услов $r + s \leq 4$.

На овај начин се добија уопштени AR модел (GAR) који заправо представља полиномијални модел.

3.5. Билинеарни модели

Билинеарни модели су природно проширење $ARMA$ модела тако што се додају унакрсни производи величина X_{t-i} и e_{t-i} , односно укључују чланови другог реда Тејлоровог развоја. На тај начин се долази до нелинеарности у циљу побољшања апроксимације. Назив билинеарних модела произилази из чињенице да је модел линеаран по X_i као и по e_j . Ова класа модела је привлачна јер се превазилази једноставан линеарни облик, а ипак задржава већи део једноставне структуре линеарних $ARMA$ модела.

Као што $ARMA$ модел може да буде добра апроксимација $AR(p)$ модела великог реда p , тако у случају нелинеарних модела билинеарни модел представља добру апроксимацију GAR модела великог реда. [3]

Најједноставнији билинеарни модел је облика:

$$X_t = a_0 + a_1X_{t-1} + b_1e_{t-1} + c_1X_{t-1}e_{t-1} + e_t .$$

У општем случају, билинеарни модели $BL(p, q, r, s)$ се дефинишу на следећи начин:

$$X_t = a_0 + \sum_{i=1}^p a_i X_{t-i} + \sum_{i=1}^q b_i e_{t-i} + \sum_{i=1}^r \sum_{j=1}^s c_{ij} X_{t-i} e_{t-j} + e_t ,$$

при чему су p, q, r, s ненегативни цели бројеви који представљају параметре билинеарног модела. Уколико су сви коефицијенти $c_{ij} = 0$ добијамо линеарни $ARMA(p, q)$ модел.

Посматрајмо билинеарни модел $BL(1,0,1,1)$:

$$X_t = a_0 + a_1X_{t-1} + c_1X_{t-1}e_{t-1} + e_t ,$$

који може да се запише и као:

$$X_t = a_0 + (a_1 + c_1e_{t-1})X_{t-1} + e_t .$$

Уочавамо да је ово заправо израз за AR модел, при чему је ауторегресиони коефицијент функција $a_1 + c_1e_{t-1}$. Дакле, ауторегресиони коефицијент је случајна величина са очекивањем a_1 .

СТАЦИОНАРНОСТ

Својства билинеарних модела као што су стационарност, често се изводе изражавањем стања у тренутку t преко реализација серија у претходним тренуцима.

Како је билинеарни модел веома близу линеарном облику, можемо га изразити преко MA модела са бесконачним редом, као што то радимо за линеарне AR или $ARMA$ процесе. [4]

Заиста, понављањем n пута $X_t = aX_{t-1} + cX_{t-1}e_{t-1} + e_t$, добија се:

$$X_t = \left\{ \prod_{k=1}^n (a + c \cdot e_{t-k}) \right\} X_{t-n} + e_t + \sum_{j=1}^{n-1} \left\{ \prod_{k=1}^j (a + c \cdot e_{t-k}) \right\} e_{t-j}.$$

Ако сума на десној страни претходног израза конвергира у вероватноћи када $n \rightarrow \infty$, онда то значи да важи:

$$\prod_{k=1}^n (a + c \cdot e_{t-k}) \xrightarrow{P} 0.$$

Тада X_t може да се прикаже као:

$$X_t = e_t + \sum_{j=1}^{\infty} \left\{ \prod_{k=1}^j (a + c \cdot e_{t-k}) \right\} e_{t-j},$$

који је има облик $MA(\infty)$ модела са случајним коефицијентима $\prod_{k=1}^j (a + c \cdot e_{t-k})$.

Како је $\{e_t\}$ низ независних и једнако расподељених случајних величина, овако приказано $\{X_t\}$ је строго стационарно решење за модел:

$$X_t = \left\{ \prod_{k=1}^n (a + c \cdot e_{t-k}) \right\} X_{t-n} + e_t + \sum_{j=1}^{n-1} \left\{ \prod_{k=1}^j (a + c \cdot e_{t-k}) \right\} e_{t-j}.$$

Под условом да $E(e_t^4) < \infty$, сума на десној страни конвергира средњеквадратно ако и само ако је $a^2 + c^2\sigma^2 < 1$. То је неопходан и довољан услов да овај модел буде строго стационаран, тј. $E(X_t^2) < \infty$, под условом да је $E(e_t^4) < \infty$. [8]

Такође, важи:

$$\mu_x = EX_t = E\left(e_t + \sum_{j=1}^{\infty} \left\{ \prod_{k=1}^j (a + c \cdot e_{t-k}) \right\} e_{t-j}\right) = \sum_{j=1}^{\infty} a^{j-1} \cdot c E(e_{t-j}^2) = \frac{\sigma^2 c}{1-a}.$$

Дисперзија такође може да се изведе, али је израз компликован и није посебно „инспиративан“. Ипак, јасно је да је услов $E(e_t^4) < \infty$ неопходан за $E(X_t^2) < \infty$.

У изразу $X_t = aX_{t-1} + cX_{t-1}e_{t-1} + e_t$ се претпоставља да је $E(X_t e_t) = \sigma^2$, а ако применимо све наведене услове имамо:

$$X_t - \mu_x = a(X_{t-1} - \mu_x) + c(X_{t-1}e_{t-1} - \sigma^2) + e_t.$$

Ако се помноже обе стране са X_{t-k} за $k \geq 2$ и након тога израчуна очекивање, добија се *Yule-Walker*-ова једначина:

$$\gamma(k) = a \cdot \gamma(k-1), \quad k \geq 2,$$

где $\gamma(\cdot)$ означава аутоковаријациону функцију за $\{X_t\}$.

Тако можемо дефинисати *ARMA(1,1)* модел са ауторегресионим коефицијентом a и коефицијентом покретног просека и дисперзијом белог шума који је изабран тако да је аутоковаријациона функција модела исти као аутоковаријациона функција за $\{X_t\}$ за оба корака 0 и 1. Ово указује да је *BL(1,0,1,1)* процес ефективно *ARMA(1,1)* процес што се тиче његових својстава првог и другог момента.

ИНВЕРТИБИЛНОСТ

Пожељно је да модел буде глобално инвертибилан, тј. такав да на основу низа $\{e_t\}$ и почетне вредности могу да се одреде оригинални подаци X_t . Посматрајмо временску серију:

$$X_t = aX_{t-1} + cX_{t-1}e_{t-1} + e_t,$$

односно:

$$e_t = -(cX_{t-1})e_{t-1} + [X_t - aX_{t-1}].$$

Понављањем овог поступка j пута, добијамо:

$$e_t = \sum_{j=0}^{\infty} \left((-c)^{j-1} \prod_{i=0}^j X_{t-1-i} \right) [X_{t-j} - aX_{t-j-1}].$$

Ово инвертибилно представљање је корисно и за прогнозирање и оцењивање непознатих параметара.

Детаљније о билинеарним моделима може се наћи у литератури [7] и [8].

Пример 1. У овом примеру, кодом у програму *R* је представљена симулација $BL(1,0,1,1)$ модела помоћу функције BL_1011 .

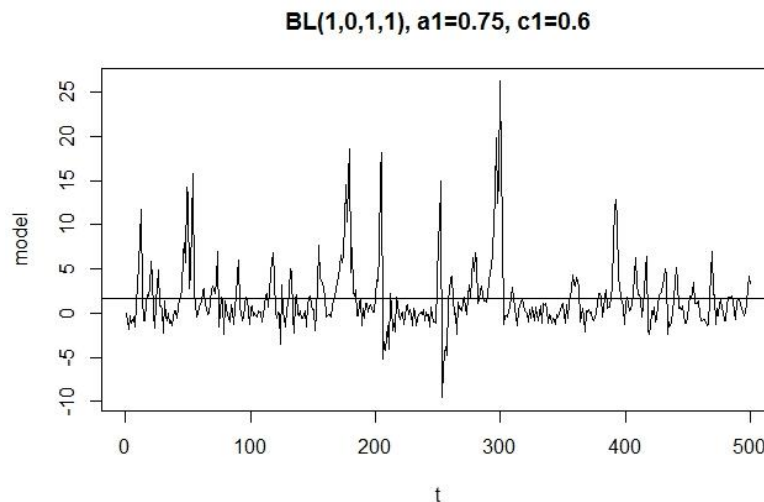
$$X_t = a_1 X_{t-1} + c_1 X_{t-1} e_{t-1} + e_t$$

где је $\{e_t\}$ низ независних једнако расподељених случајни величина са стандардном нормалном расподелом.

```
> BL_1011 <- function(n,a1,c1) {
+   X <- rnorm(n)
+   e <- rnorm(n)
+   for (t in 2:n) {
+     X[t] <- a1 * X[t-1] + c1 * e[t-1] * X[t-1] + e[t]
+   }
+   return(X[1:n])
+ }
```

Следећим кодом добијамо графички приказ $BL(1,0,1,1)$ модела са параметрима $a_1 = 0.75$, $c_1 = 0.6$.

```
> model=BL_1011(500, 0.75, 0.6)
> plot(model, xlab = "t", main = "BL(1,0,1,1), a1=0.75, c1=0.6", type = "l")
> abline(h=mean(model))
```

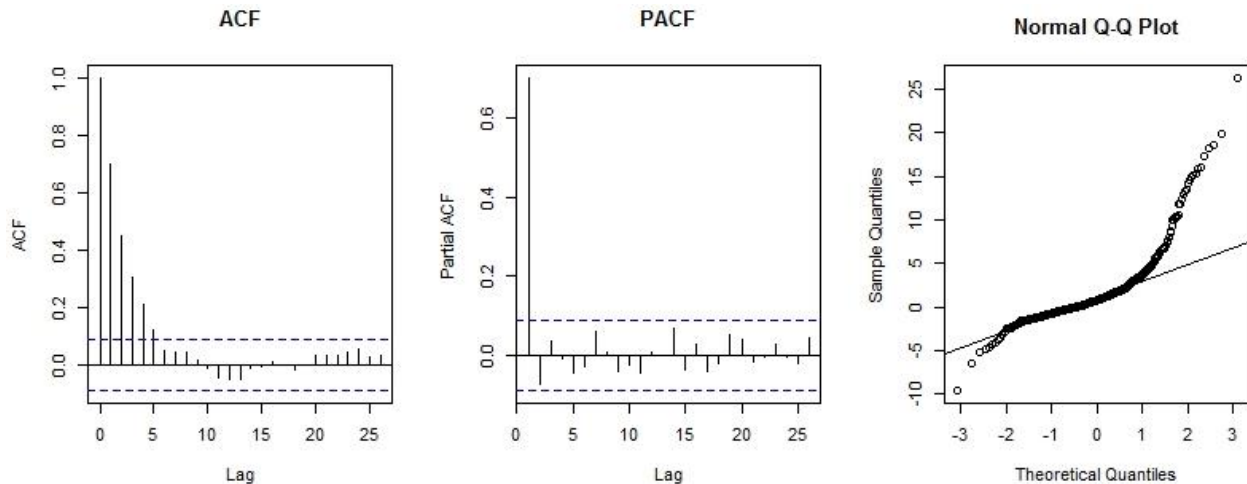


Слика бр. 1: Симулација билинеарног модела

У наставку су дати графички прикази ACF , $PACF$ и $Q-Q$ plot.

```
> par(mfrow=c(1, 3))
> acf(model, main="ACF")
> pacf(model, main="PACF")
> qqnorm(model)
```

Са слике број 2 уочавамо да вредности аутокорелационе функције брзо опадају, као и да је значајна само вредност парцијалне аутокорелација са кашњењем 1, што подсећа на својства $ARMA(1,1)$ модела.



Слика бр. 2: ACF, PACF и Q-Q plot билинеарног модела

3.6. Још неки параметарски модели

РАЦИОНАЛНИ МОДЕЛ

Рационални модели су модели који су дефинисани рационалном функцијом, односно као количник два полинома:

$$X_t = \frac{a_0 + \sum_{i=1}^k \sum_{j=1}^p a_{ij} X_{t-i}^j}{b_0 + \sum_{i=1}^k \sum_{j=1}^q b_{ij} X_{t-i}^j}, \quad 0 \leq p \leq q + 1.$$

ЕКСПОНЕНЦИЈАЛНИ МОДЕЛИ

Експоненцијални модели (*EXPAR*) су дефинисани на следећи начин:

$$X_t = \sum_{i=1}^k (a_i + b_i e^{-a X_{t-d}^2}) X_{t-d} + e_t, \quad a > 0.$$

За довољно велику или довољно малу вредност $|X_{t-d}|$ добија се $AR(k)$ модел.

ARCH МОДЕЛ

Код *ARCH* (енгл. *AR Conditional Heteroscedastic*) модела, условна дисперзија је линеарна функција квадрата претходних опсервација:

$$X_t = \sqrt{a_0 + a_1 X_{t-1}^2 + \dots + a_p X_{t-p}^2} \cdot e_t, \quad a_i > 0.$$

3.7. Модели са случајним коефицијентима

Предвиђање волатилности временских серија је веома корисно, посебно када се ради о финансијским временским серијама. Ауторегресиони модели са случајним коефицијентима су уведени како би се узела у обзир варијабилност између различитих предмета који се проучавају, слично као панел анализа у економетрији. Ови модели се класификују као условно хетероскедастични модели, а користе се како би се добио бољи опис једначине условног очекивања процеса омогућавајући да се параметри развијају током времена. [1]

Ауторегресиона временска серија $\{X_t\}$ представљена је на следећи начин:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + e_t ,$$

где су ϕ_1, \dots, ϕ_p реални параметри, а $\{e_t\}$ серија некорелисаних случајних величина са очекивањем 0 и дисперзијом 1. Поменути параметри ϕ_1, \dots, ϕ_p су фиксни и они се не мењају током времена. Међутим, ова претпоставка није баш реалистична јер посматрајући дужу финансијску временску серију, може се очекивати да се параметри споро мењају током времена. Због тога су и уведени модели са случајним коефицијентима.

За временску серију X_t се каже да прати ауторегресиони модел реда p са случајним коефицијентима, у ознаци $RCA(p)$ ако задовољава:

$$X_t = \sum_{i=1}^p (\phi_i + \delta_{it}) X_{t-i} + e_t ,$$

где је p позитиван цео број, ϕ_1, \dots, ϕ_p су фиксни коефицијенти, а $\{\delta_t\} = \{(\delta_{1t}, \dots, \delta_{pt})^T\}$ је низ независних случајних вектора са очекивањем 0 и коваријационом матрицом Ω_δ , $\{e_t\}$ је бели шум са очекивањем 0 и дисперзијом σ_e^2 , при чему важи да су $\{\delta_t\}$ и $\{e_t\}$ независни.

Условно очекивање и дисперзија $RCA(p)$ модела износе:

$$\mu_t = E(X_t | \mathcal{F}_{t-1}) = \sum_{i=1}^p \phi_i X_{t-i} ,$$

$$\sigma_t^2 = \sigma_e^2 + (X_{t-1}, \dots, X_{t-p}) \Omega_\delta (X_{t-1}, \dots, X_{t-p})^T .$$

Дакле, за $RCA(p)$ моделе, волатилност је квадратна функција посматраних претходних вредности X_{t-i} .

Специјално, за $p = 1$, добија се $RCA(1)$ модел:

$$X_t = (\phi_1 + \delta_t)X_{t-1} + e_t.$$

У општем случају, $RCA(1)$ модел није низ са коначним моментима другог реда. Претпоставка слабе стационарности је једна од полазних тачака у одређивању корелационе структуре и оцене непознатих параметара.

$RCA(1)$ модел је слабо стационаран и ергодичан ако важи:

$$E[(\phi_1 + \delta_t)^2] = \phi_1^2 + \delta_t^2 < 1,$$

а строго стационаран и ергодичан са бесконачним другим моментом ако важи:

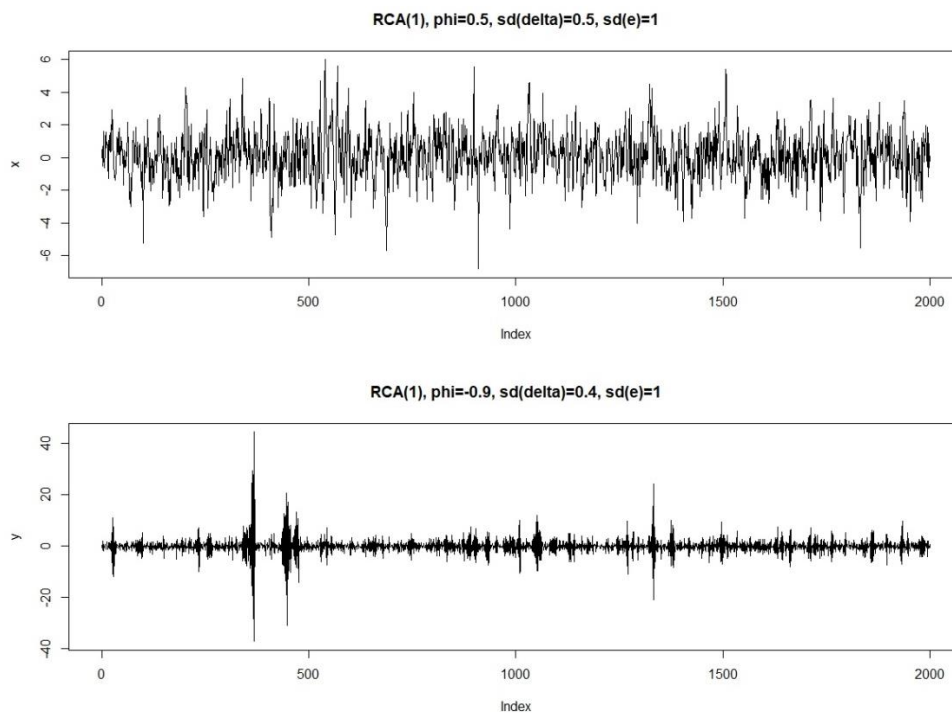
$$-\infty \leq E(\log|\phi_1 + \delta_t|) < 0.$$

Детаљније о моделима са случајним коефицијентима може се наћи у литератури која је коришћена за писање овог поглавља: [1], [9], [10] и [13].

Пример 2. У овом примеру је извршена симулација два $RCA(1)$ модела са различитим параметрима:

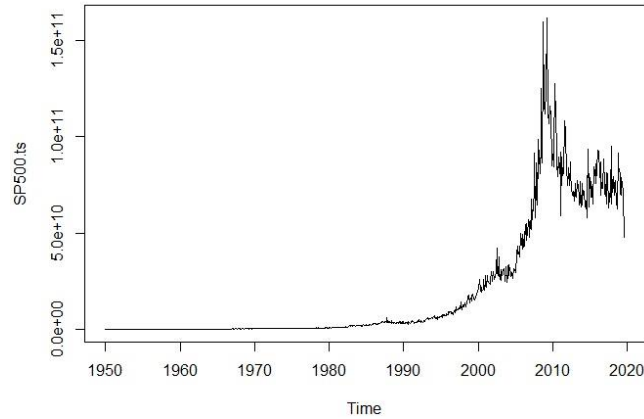
$$X_t = (0.5 + \delta_t)X_{t-1} + e_t, \quad \sigma_\delta^2 = 0.25, \sigma_e^2 = 1$$

$$Y_t = (-0.9 + \delta_t)Y_{t-1} + e_t, \quad \sigma_\delta^2 = 0.25, \sigma_e^2 = 1$$



Слика бр. 3: Симулација два модела са случајним коефицијентима

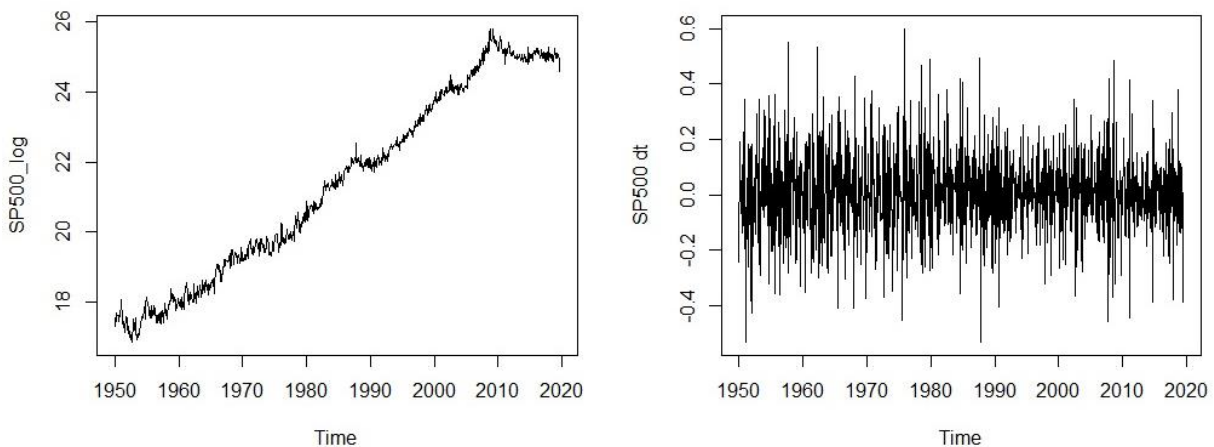
Пример 3. У овом примеру се посматра *Standard & Poor's 500* берзански индекс¹, тј. обим месечних трансакција у периоду од јануара 1950. до августа 2019. године, са укупно 836 опсервација.



Слика бр. 4: Графички приказ обима месечних трансакција

На основу слике број 4 можемо да претпоставимо да посматрана временска серија има приближно експоненцијални тренд до краја 2008. године, а након тога се уочава пад као последица светске економске кризе. Такође, примећује се да варијабилност података расте са временом, због чега је погодно посматрати логаритмовану серију.

Нека X_t означава логаритам месечног обима трансакција у тренутку t . Са графичког приказа временске серије X_t уочава се постојање растућег линеарног тренда. Након уклањања тренда, добија се временска серија која је прилично стабилна око своје средње вредности, осим краткорочних шокова када се уочава присуство веће волатилности. Због тога је погодно за ову временску серију користити $RCA(1)$ модел.



Слика бр. 5: Графички приказ логаритмоване детрендиране серије

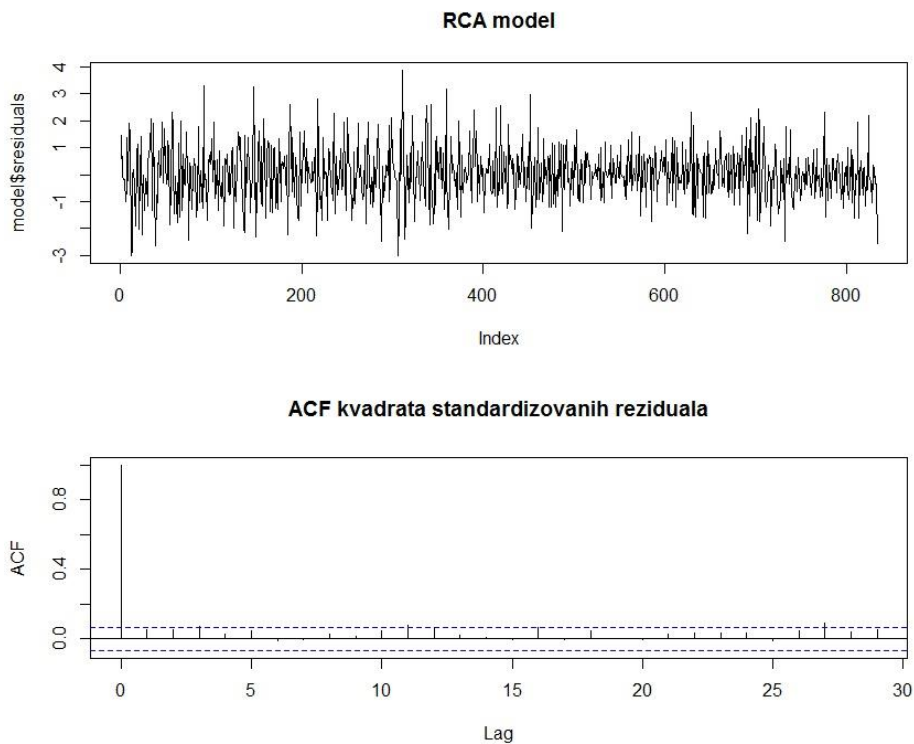
¹ База података је доступна на сајту: <https://finance.yahoo.com/quote/%5EGSPC/>

У програму *R* коришћењем функције *rcAR()* из пакета *NTS* добија се да је узорачка средина временске серије X_t једнака 0.01, тако да $\tilde{X}_t = X_t - 0.01$ представља прилагођену временску серију. Након уклапања података у *RCA(1)* модел добија се:

$$\tilde{X}_t = (-0.38 + \delta_t)\tilde{X}_{t-1} + e_t$$

где је δ_t низ независних случајних величина са очекивањем 0 и коваријацијом (дисперзијом) 0.02, а дисперзија за e_t је 0.02.

На основу слике број 6, која представља графички приказ стандардизованих резидуала уклопљеног *RCA(1)* модела и аутокорелациону функцију квадрата стандардизованих резидуала, може да се закључи да модел адекватно описује остварену волатилност посматране временске серије.



Слика бр. 6: Графички приказ стандардизованих резидуала и аутокорелациона функција квадрата стандардизованих резидуала

3.8. Модели временских серија са праговима

Модели временских серија са праговима су класа нелинеарних модела где се увођењем прагова модел разлаже на коначан број подмодела. Сваки од подмодела важи између нека два суседна прага, а избор подмодела зависи од вредности неке случајне величине Z_t , која се назива *threshold* променљива. У зависности од случајне величине Z_t разликујемо различите моделе временских серија са праговима. Уколико је Z_t нека од претходних вредности серије, ради се о *SET* (енгл. *Self-Exciting Threshold*) моделима.

Такође, у зависности од врсте подмодела разликујемо различите типове модела временских серија са праговима.

TAR (енгл. *Threshold AR*) модели су модели где је *threshold* променљива функција више претходних опсервација, а подмодели су *AR* модели. Уколико су подмодели између прагова *AR* модели и ако је Z_t нека од претходних вредности серије, онда се ради о *SETAR* (енгл. *Self-Exciting Threshold AR*) моделима.

Ако су између прагова *MA* или *ARMA* модели, ради се о *SETMA* (енгл. *Self-Exciting Threshold Moving Average*), односно *SETARMA* (енгл. *Self-Exciting Threshold ARMA*) моделима. *SET* модели који имају случајне коефицијенте називају се *SETRC* (енгл. *Self-Exciting Threshold Random Coefficient*) модели.

У наставку ће бити представљени *TAR*, *SETAR* и *STAR* модели, а за писање овог поглавља коришћена је литература [2], [3], [4], [5], [11], [12], [14] и [15].

3.8.1. *TAR* и *SETAR* модели

TAR модели су настали због неколико нелинеарних карактеристика које се често јављају у пракси, као што су асиметрија у обрасцу пада или раста процеса. Ови модели користе део по део линеарне моделе како би се добила боља апроксимација једначине условног очекивања. *TAR* модели допуштају промене које се јављају у одређеним временским интервалима (режимима) користећи праг.

Дакле, основни *TAR* модел може да се посматра као део по део линеарни *AR* модел са променама из једног режима у други који зависи од односа вредности прага у односу на променљиву Z_{t-d} .

1. Дворежимски *TAR (SETAR)* модели

Временска серија X_t прати дворежимски *TAR* модел редова p_1 и p_2 , у ознаци $TAR(2; p_1, p_2)$ са *threshold* променљивом Z_{t-d} ако важи:

$$X_t = \begin{cases} \phi_0 + \sum_{i=1}^{p_1} \phi_i X_{t-i} + \sigma_1 e_t, & Z_{t-d} \leq r \\ \theta_0 + \sum_{i=1}^{p_2} \theta_i X_{t-i} + \sigma_2 e_t, & Z_{t-d} > r \end{cases},$$

где је e_t низ независних, једнако расподељених случајних величина са очекивањем нула и дисперзијом један, θ_i и ϕ_j су реални параметри, тако да је $\theta_i \neq \phi_i$ за неко i , d је позитиван цео број који означава параметар кашњења, а r означава вредност прага.

Дакле, у зависности од Z_{t-d} генерише се један од два *AR* модела. Уколико је $Z_{t-d} = X_{t-d}$, реч је о *SETAR* моделу са параметром кашњења d .

У наставку је разматран случај када је $Z_{t-d} = X_{t-d}$. Често се претпоставља да e_t има нормалну $\mathcal{N}(0,1)$ расподелу. У наставку се користи исти ред p за оба режима ради једноставности. Модел *TAR* који је претходно дефинисан, такође може да се запише на следећи начин:

$$X_t = \left(\phi_0 + \sum_{i=1}^p \phi_i X_{t-i} + \sigma_1 e_t \right) (1 - I(X_{t-d} > r)) + I(X_{t-d} > r) \left(\theta_0 + \sum_{i=1}^p \theta_i X_{t-i} + \sigma_2 e_t \right)$$

$$X_t = \phi_0 + \sum_{i=1}^p \phi_i X_{t-i} + \sigma_1 e_t + I(X_{t-d} > r) \left(\beta_0 + \sum_{i=1}^p \beta_i X_{t-i} + \gamma e_t \right),$$

где је:

$$I(X_{t-d} > r) = \begin{cases} 1 & , X_{t-d} > r \\ 0 & , X_{t-d} \leq r \end{cases}$$

$$\beta_i = \theta_i - \phi_i, i = 0, \dots, p$$

$$\gamma = \sigma_2 - \sigma_1.$$

На овај начин, структурна промена модела када је $X_{t-d} = r$ постаје очигледна са β_i , које означава промену i -ог коефицијента.

Ова врста модела се често јавља у многим научним областима. Са економске тачке гледишта, оправданост за коришћење *TAR* и *SETAR* модела је да када случајна величина

која представља праг прекорачи одређену вредност, динамика временске серије се мења. На пример, ако инфлација оде изнад одређеног нивоа, онда може да настане хиперинфлација која мења динамику саме инфлације, или када буџетски дефицит премaши граничну вредност, динамика раста БДП-а може да се промени јер потрошачи и тржишта могу да очекују масовна будућа повећања опорезивања или смањења јавних трансфера. Међутим, како је потребно времена да се његово деловање покаже, то резултира одлагањем од d јединица времена.

У наставку су упоређени AR , GAR , BL и TAR модели.

Пример 4. У програму R је генерисана серија од 200 опсервација са стандардном нормалном расподелом која представља серију $\{e_t\}$. На првом графику је представљен $AR(1)$ процес (слика број 7) који је описан следећом формулом:

$$X_t = 0.7X_{t-1} + e_t, \text{ при чему је } X_1 = e_1.$$

На другом графику је представљен GAR процес облика:

$$X_t = 0.7X_{t-1} - 0.06X_{t-1}^2 + e_t,$$

који може да се запише и као $AR(1)$ процес са случајним коефицијентима:

$$X_t = (0.7 - 0.06X_{t-1})X_{t-1} + e_t.$$

Примећује се да, што је већа вредност X_{t-1} то је мањи ауторегресиони коефицијент. Такође, уколико се упореде AR и GAR серије, примећује се да GAR серија има више негативних вредности и да се спорије враћа ка нули од AR серије.

Следећи график представља билинеарни модел који је описан следећом формулом:

$$X_t = 0.7X_{t-1} - 0.3X_{t-1}e_{t-1} + e_t, \text{ односно}$$

$$X_t = (0.7 - 0.3e_{t-1})X_{t-1} + e_t.$$

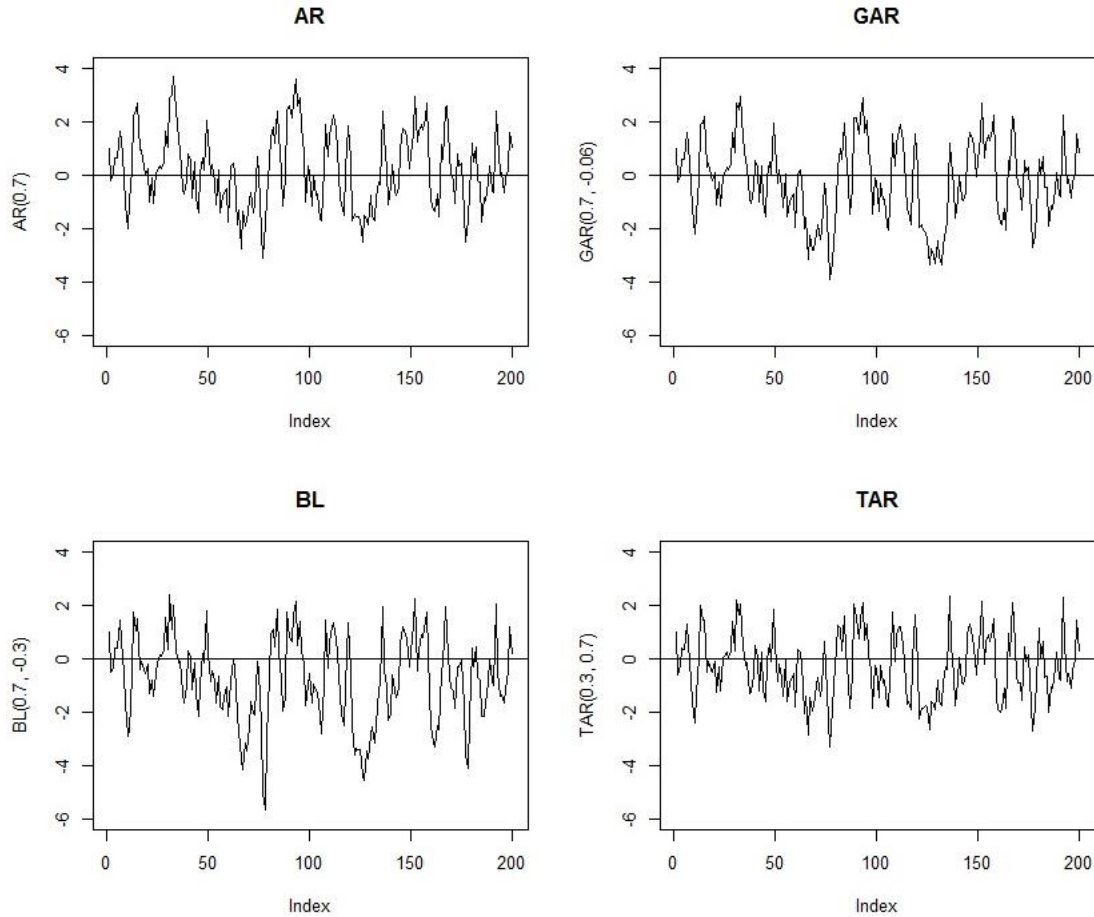
Код BL модела, степен постојаности зависи од вредности e_{t-1} , тако да што је веће e_{t-1} то је мањи степен постојаности. На графику могу да се уоче и екстремна кретања BL процеса.

На крају је представљен TAR процес помоћу следеће формуле:

$$X_t = 0.3X_{t-1}I_t + 0.7X_{t-1}(1 - I_t) + e_t$$

где је $I_t = 1$ за $X_{t-1} > 0$, а у супротном је $I_t = 0$.

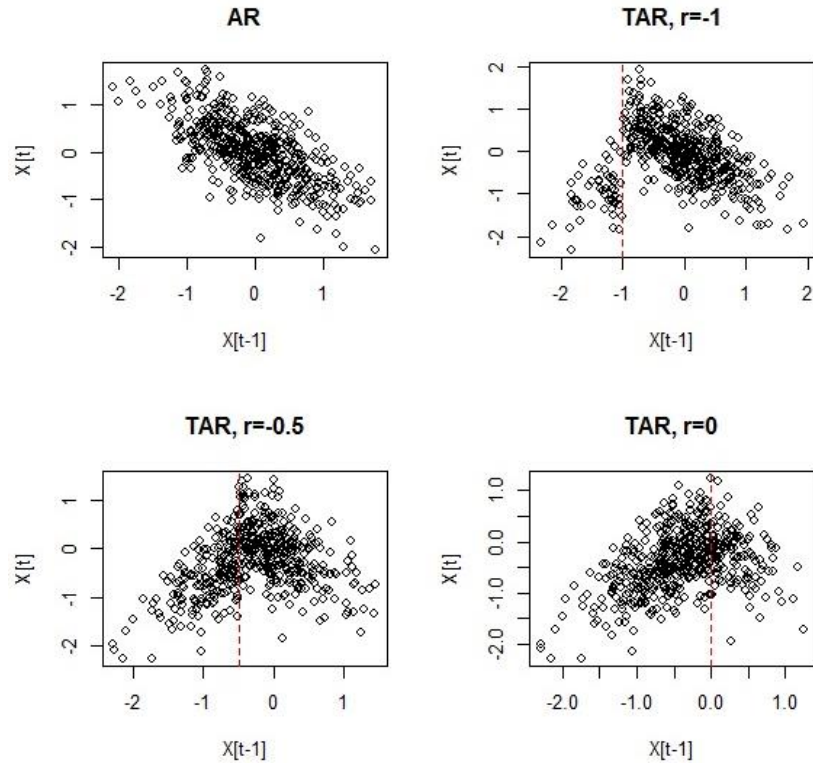
За $X_{t-1} \leq 0$, овај TAR процес је исти као први $AR(1)$ процес. Међутим, за TAR процес, само 30% тренутне вредности X_t има тенденцију да се задржи у наредном периоду када је $X_{t-1} > 0$.

Слика бр. 7: Поређење AR , GAR , BL и TAR модела

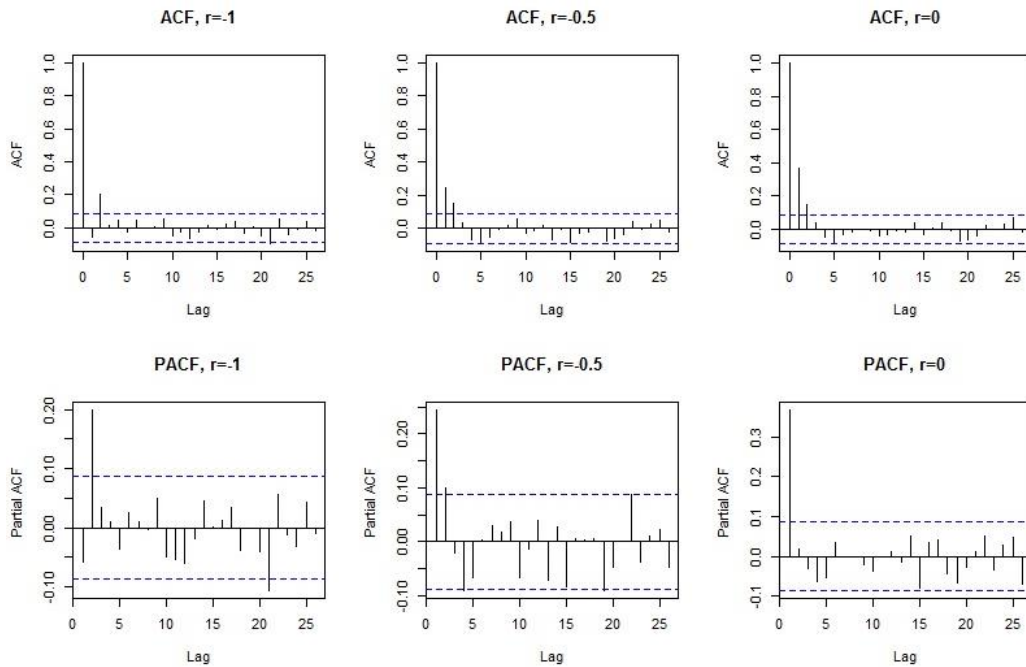
Пример 5. Нека је једноставан TAR (тј. $SETAR$) модел облика:

$$X_t = \begin{cases} -0.7X_{t-1} + e_t, & X_{t-1} \geq r \\ 0.7X_{t-1} + e_t, & X_{t-1} < r \end{cases},$$

где су e_t независне случајне величине са нормалном $\mathcal{N}(0, 0.5^2)$ расподелом. У програму R су генерисане четири серије X_t са по 500 опсервација, које ће имати различите прагове. На првом графику је представљена серија код које је праг $r = -\infty$, што значи да се ради о линеарном $AR(1)$ моделу, док када је праг $r = -1$, $r = -0.5$ или $r = 0$, примећује се присуство нелинеарности.

Слика бр. 8: Поређење различитих *TAR* модела

Графички прикази *ACF* и *PACF* указују на то да скоро да и нема значајне аутокорељације када је $r = -1$, упркос зависности између X_t и X_{t-1} .

Слика бр. 9: *ACF* и *PACF* *TAR* модела

2. Својства дворежимског TAR модела

Најједноставнији дворежимски модел $TAR(1)$ може да се запише:

$$X_t = \begin{cases} \phi_1 X_{t-1} + \sigma_1 e_t, & X_{t-1} \leq 0 \\ \theta_1 X_{t-1} + \sigma_2 e_t, & X_{t-1} > 0 \end{cases},$$

где је параметар кашњења $d = 1$, праг је нула, а e_t је низ независних једнако расподељених случајних величина са очекивањем нула, дисперзијом 1 и са позитивном функцијом густине расподеле.

Функција $f(\cdot)$ овог једноставног модела је:

$$f(X_{t-1}) = \begin{cases} \phi_1 X_{t-1}, & X_{t-1} \leq 0 \\ \theta_1 X_{t-1}, & X_{t-1} > 0 \end{cases}.$$

Како X_0 може да представља било који број, да би низ X_t био добро дефинисан, потребно је да важи $\phi_1 < 1$ и $\theta_1 < 1$. У супротном, могу да се пронађу неке X_0 за које рекурзивне замене $f(\cdot)$ воде до експлозивних вредности. То и не чуди с обзиром да је линеарни $AR(1)$ модел стационаран када је $|\phi_1| < 1$. С друге стране, прелази између два режима стварају нову прилику за TAR моделе.

Потребан и довољан услов да дворежимски $TAR(1)$ модел буде ергодичан [15]:

- a) $\phi_1 < 1$
- b) $\theta_1 < 1$
- c) $\phi_1 \theta_1 < 1$.

Услов $c)$ осигурава да одговарајућа функција $f(\cdot)$ модела остаје стабилна за било коју произвољну почетну вредност.

За $d > 1$, ови услови нису довољни за ергодичност процеса. Ако су ϕ_1 и θ_1 истог знака, ови услови су довољни, међутим ако су супротног знака, онда поред њих треба имати и неке додатне услове. Због тога су уведени услови за геометријску ергодичност, а више о њима се може наћи у [15].

3. Вишережимски TAR модели

Нека су m , d и p_1, p_2, \dots, p_m природни бројеви, а r_1, r_2, \dots, r_{m-1} реални бројеви за које важи $r_1 < \dots < r_{m-1}$.

Временска серија X_t прати m -режимски $TAR(p_1, p_2, \dots, p_m)$ модел са *threshold* променљивом Z_{t-d} где је $d > 0$ параметар кашњења, ако задовољава:

$$X_t = \begin{cases} \phi_{1,0} + \sum_{i=1}^{p_1} \phi_{1,i} X_{t-i} + \sigma_1 e_t, & Z_{t-d} \leq r_1 \\ \phi_{2,0} + \sum_{i=1}^{p_2} \phi_{2,i} X_{t-i} + \sigma_2 e_t, & r_1 < Z_{t-d} \leq r_2 \\ \dots & \dots \\ \phi_{m,0} + \sum_{i=1}^{p_m} \phi_{m,i} X_{t-i} + \sigma_m e_t, & r_{m-1} < Z_{t-d} \end{cases}$$

где су σ_j позитивни реални бројеви, e_t је из $\mathcal{N}(0,1)$ расподеле, а $\phi_{j,i}$, $j = 0, 1, \dots, m$, $i = 1, 2, \dots, p_j$ су реални параметри.

Овакав модел, у ознаци $TAR(m; p_1, p_2, \dots, p_m)$ је дефинисан са m AR модела, чији су редови p_1, p_2, \dots, p_m респективно, а случајна величина Z_{t-d} одређује актуелни подмодел. Ако су сви подмоделни $AR(p)$, онда се користи ознака $TAR(m; p)$. За $m = 1$ се добија обичан AR модел.

За $Z_t = X_t$ у моделу добија се $SETAR(l; p_1, p_2, \dots, p_m)$ модел:

$$X_t = \sum_{j=1}^m \left(\phi_{j,0} + \sum_{i=1}^{p_m} \phi_{j,i} X_{t-i} \right) I(X_{t-d} \in R_j) + e_t,$$

где је $R_j = (r_{j-1}, r_j]$, $r_0 = -\infty$, $r_m = +\infty$.

Ако је θ вектор параметара у моделу,

$$\theta = (\theta_1^T, \theta_2^T, \dots, \theta_m^T, r_1, r_2, \dots, r_m, d)^T,$$

где је $\theta_j = (\phi_{j,0}, \phi_{j,1}, \dots, \phi_{j,p})^T$, $j = 1, 2, \dots, m$, модел има облик

$$X_t = F(X_{t-1}, X_{t-2}, \dots, X_{t-p}, \theta) + e_t,$$

при чему је:

$$F(X_{t-1}, X_{t-2}, \dots, X_{t-p}, \theta) = \sum_{j=1}^m \left(\phi_{j,0} + \sum_{i=1}^{p_j} \phi_{j,i} X_{t-i} \right) I(X_{t-d} \in R_j) .$$

Нелинеарност *SETAR* модела је обезбеђена са $m > 1$. Већина примена *TAR* модела се односи на $m = 2$ или $m = 3$.

Потешкоће које се могу јавити приликом коришћења *TAR* модела су одређивање вредности прага (односно прагова) као и редова одвојених *AR* модела. Процена нивоа прага је углавном субјективна. [3]

ОЦЕЊИВАЊЕ НЕПОЗНАТИХ ПАРАМЕТАРА

Оцењивање параметара *TAR* модела често се врши методом нелинеарних најмањих квадрата. Међутим, тешко је утврдити својства добијених оцена непознатих параметара.

Chan [14] је разматрао оцењивање параметара дворежимског *TAR(p)* модела који је дат следећом једначином:

$$X_t = \begin{cases} \phi_0 + \sum_{i=1}^p \phi_i X_{t-i} + \sigma_1 e_t, & X_{t-d} \leq r \\ \theta_0 + \sum_{i=1}^p \theta_i X_{t-i} + \sigma_2 e_t, & X_{t-d} > r \end{cases} .$$

Нека су $\phi = (\phi_0, \phi_1, \dots, \phi_p)^T$ и $\theta = (\theta_0, \theta_1, \dots, \theta_p)^T$ вектори параметара два режима, респективно и нека је $\Theta = (\phi^T, \theta^T, r, d)^T$ вектор параметара модела у претходној једначини, а Θ_0 прави вектор параметара.

Претпоставимо да су доступне реализације временске серије $\{X_t | t = 1, 2, \dots, T\}$, где T означава величину узорка. Нека је $E_{\Theta}(X_t | \mathcal{F}_{t-1})$ условно очекивање модела који има вектор параметара Θ и нека је \mathcal{F}_{t-1} σ -поље генерисано доступним информацијама у тренутку $t - 1$. Функција условних најмањих квадрата је:

$$L_T(\Theta) = \sum_{t=p+1}^T [X_t - E_{\Theta}(X_t | \mathcal{F}_{t-1})]^2 .$$

У пракси, функција $L_T(\Theta)$ се минимизира, тако што се минимизира $S(r, d)$, где је $S(r, d)$ минимум суме квадрата резидуала модела са датим фиксним прагом r и кораком кашњења d . $S(r, d)$ се може добити минимизирањем сума квадрата резидуала користећи податке у сваком режиму.

Конкретно, за дате r и d , имамо $S_1 = \{t | X_{t-d} \leq r\}$ и $S_2 = \{t | X_{t-d} > r\}$ колекције индекса времена за режиме 1 и 2, респективно. Затим се методом најмањих квадрата оцењују непознати параметри ϕ линеарног модела:

$$X_t = \phi_0 + \sum_{i=1}^p \phi_i X_{t-i} + e_t, \quad t \in S_1,$$

и параметри θ линеарног модела:

$$X_t = \theta_0 + \sum_{i=1}^p \theta_i X_{t-i} + e_t, \quad t \in S_2,$$

где e_t означава појам грешке. Дисперзије σ_1^2 и σ_2^2 се затим оцењују на основу средње квадратних грешака резидуала из претходне две линеарне регресије.

Тада се функција $L_T(\Theta)$ може минимизирати минимизирањем $S(r, d)$.

У већини случајева, вредност прага је непозната и треба да се процени заједно са другим параметрима TAR модела. *Chan* претпоставља да је $r = X_{(n)}$, где је n унутар унапред одређеног распона, нпр. $0.05T \leq n \leq 0.95T$. Поред тога, претпоставља се да $d \in \{1, \dots, r\}$. Стога се минимизирање $S(r, d)$ врши у коначним броју корака.

Генерално, праг мора да се налази између максималне и минималне вредности серије. У пракси, 15% највиших и најнижих вредности су искључене из претраге како би се осигурао адекватан број запажања на свакој страни прага. Ако постоји веома велики број опсервација, искључују се само 10% највиших и најнижих опсервација као потенцијални прагови. За сваку од 70% унутрашњих опсервација се проверава да ли је праг. Регресија која садржи најмању суму квадрата резидуала садржи конзистентну процену прага.

Постоји неколико процедура које су доступне за избор вредности параметра кашњења d . Стандардна процедура је да се формира TAR модел за сваку потенцијалну вредност $d \in \{1, \dots, r\}$. Онај са најмањом вредношћу сума квадрата резидуала даје конзистентну процену параметра кашњења. Алтернативно, може се одабрати параметар кашњења који води до најмање вредности AIC .

Нека је $\hat{\Theta}$ оцена за Θ добијена методом најмањих квадрата. Претпоставимо да је посматрани дворежимски $TAR(p)$ модел стационаран и ергодичан, са коначним другим моментима и да је функција густине стационарне расподеле $(X_1, X_2, \dots, X_p)^T$ је свуда позитивна. *Chan* [14] је показао да је оцена $\hat{\Theta}$ конзистентна када $T \rightarrow \infty$. То јест, $\hat{\Theta} \rightarrow \Theta_0$ скоро сигурно када $T \rightarrow \infty$. Поред тога, и оцене $\hat{\sigma}_i^2$, ($i = 1, 2$) су такође конзистентне.

ИЗГРАДЊА МОДЕЛА

За дворежимске TAR моделе, поступак изградње модела се често започиње избором ауторегресионог модела реда p и претпоставља се да параметар кашњење d задовољава $1 \leq d \leq p$. Након тога, претпоставља се да праг r узима одређену вредност статистике поретка. Нека је статистика поретка дата са:

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)} \leq X_{(n+1)} \leq \dots \leq X_{(T)} .$$

Претпоставља се да је $r \in \{X_{(l)} \leq X_{(l+1)} \leq \dots \leq X_{(u)}\}$, где је $l = \lfloor \alpha T \rfloor$ и $u = \lfloor (1 - \alpha) T \rfloor$ са $0 < \alpha < 0.3$ и $\lfloor z \rfloor$ који означава цео део реалног броја z .

Са унапред одређеним максималним редом p AR модела и α , могу да се одреде параметар кашњења d , праг r , редови p_i AR модела у сваком режиму, као и ауторегресиони коефицијенти. Редови AR модела унутар сваког режима j добијају се минимизирањем AIC критеријума за сваки од режимских линеарних модела ($0 \leq p_j \leq p$).

Нека је T_i број опсервација у режиму S_i , где су $i = 1$ и 2 . Под претпоставком нормалности, оцена дисперзије методом максималне веродостојности је:

$$\tilde{\sigma}_i^2 = \frac{1}{T_i} \sum_{t \in S_i} \left(X_t - \hat{\phi}_{0,i} - \sum_{j=1}^{p_i} \hat{\phi}_{j,i} X_{t-j} \right)^2 ,$$

где је p_i AR ред режима i , $\hat{\phi}_{j,i}$ оцена параметра $\phi_{j,i}$ методом најмањих квадрата.

Тада је укупни AIC за фитовани дворежимски TAR модел:

$$AIC = T_1 \ln(\tilde{\sigma}_1^2) + T_2 \ln(\tilde{\sigma}_2^2) + T[1 + \ln(2\pi)] + 2n_p ,$$

где n_p означава број параметара који се користе у моделу.

На пример, $n_p = (p_1 + 1) + (p_2 + 1) + 1$, ако оба режима садрже слободан члан.

Изградња TAR модела са више режима је тежа. Интуитивно, што је више режима, то је нижи AR ред појединог режима. У пракси се често претпоставља да је број режима мали као што је $m \leq 3$.

ПРЕДВИЋАЊЕ

Предвиђања код TAR модела могу да се добију симулацијом. У наставку је разматран дворежимски TAR модел.

Нека је $n = T$ величина узорка и нека је l хоризонт предвиђања. Желимо да предвидимо X_{n+l} , на основу познатих вредности $\{X_n, X_{n-1}, \dots\}$. Нека $X_n(l)$ означава тачку предвиђања за X_{n+l} , а $e_n(l)$ грешку предвиђања. Предвиђање се врши методом најмањих квадрата, а по потреби се могу користити и друге функције губитка.

Предвиђање за један корак

С обзиром на то да је $d \geq 1$, предвиђање TAR модела за један корак је лако спровести. То је зато што се посматра вредност *threshold* променљиве X_{t+1-d} која је позната. Ако је $X_{t+1-d} \leq r$, тада имамо:

$$X_n(1) = \phi_0 + \sum_{i=1}^p \phi_i X_{n+1-i} , \quad e_n(1) = \sigma_1 e_{n+1} .$$

Под претпоставком нормалности, предвиђање 95% интервала за X_{n+1} је $X_n(1) \pm 1,96\sigma_1$. С друге стране, ако је $X_{t+1-d} > r$, тада имамо:

$$X_n(1) = \theta_0 + \sum_{i=1}^p \theta_i X_{n+1-i} , \quad e_n(1) = \sigma_2 e_{n+1} .$$

и предвиђање 95% интервала за X_{n+1} је $X_n(1) \pm 1,96\sigma_2$.

Предвиђање за више корака

Прво разматрамо предвиђање за два корака унапред. Ако је позната вредност *threshold* променљиве X_{t+2-d} и $X_{t+2-d} \leq r$, тада је:

$$X_n(2) = \phi_0 + \phi_1 X_n(1) + \sum_{i=2}^p \phi_i X_{n+2-i} , \quad e_n(2) = \phi_1 e_n(1) + \sigma_1 e_{n+2} .$$

Ако је $X_{t+2-d} > r$, тада имамо:

$$X_n(2) = \theta_0 + \theta_1 X_n(1) + \sum_{i=2}^p \theta_i X_{n+2-i} , \quad e_n(2) = \theta_1 e_n(1) + \sigma_2 e_{n+2} .$$

С друге стране, ако није позната вредност *threshold* променљиве X_{t+2-d} , тада је потребна симулација за предвиђање. Вишеструке итерације су потребне за добијање поузданих предвиђања.

Процедура предвиђања:

Нека је m број итерација. За $j = 1, 2, \dots, m$:

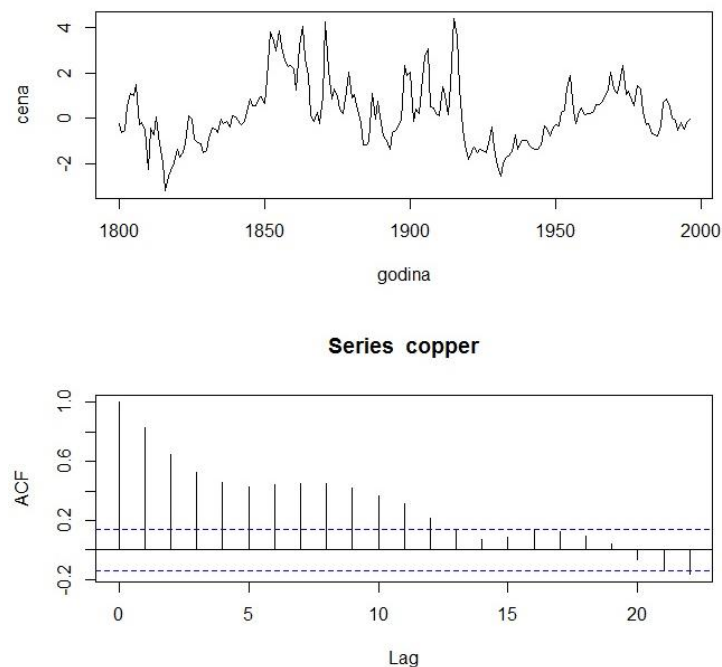
1. Генерише се l случајних величина које имају исту расподелу као e_t . Обично се претпоставља да има $e_t \sim \mathcal{N}(0,1)$ расподелу. Добијају се случајне величине: $e_{n+1}^{(j)}, \dots, e_{n+l}^{(j)}$.
2. За $t = n + 1, \dots, n + l$, користи се TAR модел и $\{e_{n+i}^{(j)}\}$ да би се генерисале реализације $X_{n+1}^{(j)}, \dots, X_{n+l}^{(j)}$ за j -ту итерацију.

Коначно, за $i = 1, \dots, l$ се рачунају предвиђања вредности серије и грешке предвиђања:

$$\tilde{X}_n(i) = \frac{1}{m} \sum_{j=1}^m X_{n+i}^{(j)}$$

$$D[e_n(i)] = \frac{1}{m-1} \sum_{j=1}^m [X_{n+i}^{(j)} - \tilde{X}_n(i)]^2 .$$

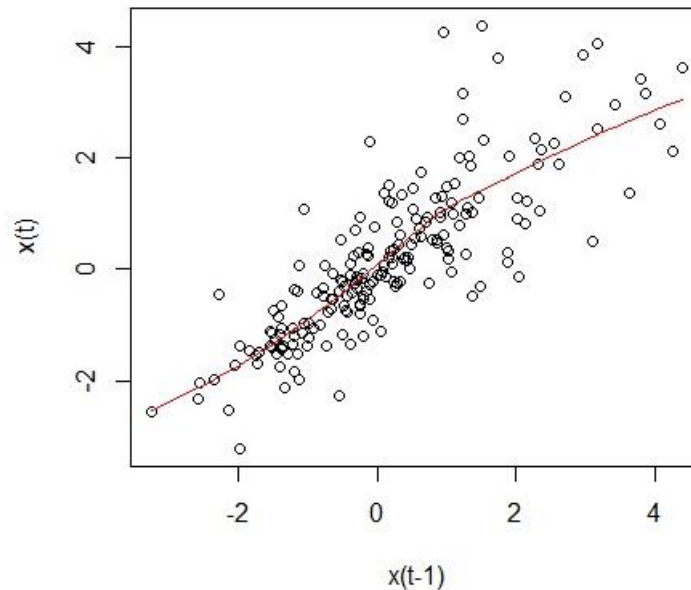
Пример 6. У овом примеру се користи временска серија која представља годишње цене бабра у периоду од 1800. до 1996. године. Цене су прилагођене након уклањања дугорочног тренда.² Нека X_t означава цену бабра у години t . На слици број 10 је најпре дат графички приказ временске серије и аутокорелационе функције.



Слика бр. 10: Графички приказ временске серије и аутокорелационе функције

² Подаци се могу наћи на: <https://faculty.chicagobooth.edu/ruey.tsay/teaching/nts/>

На слици број 11 је приказан дијаграм расипања за X_{t-1} и X_t . Црвена линија означава уклопљене вредности помоћу локалне методе изравнавања (*loess* регресија). Примећујемо да је веза између X_{t-1} и X_t нелинеарна, а да се нагиб постављене линије мења када је X_{t-1} нешто мањи од 1.



Слика бр. 11: Дијаграм расипања за X_{t-1} и X_t

Дакле, са слике се види да би одговарајући нелинеарни модел за годишњу цену бакра могао да буде дворежимски *TAR* модел.

Уколико бисмо користили линеарни модел, на основу *AIC* критеријума најпогоднији би био *AR(12)* модел, са следећим карактеристикама:

	Estimate	Std. Error	z value	Pr(> z)
ar1	0.899884	0.068903	13.0602	< 2.2e-16 ***
ar2	-0.198050	0.093010	-2.1294	0.0332251 *
ar3	0.039901	0.093763	0.4255	0.6704365
ar4	0.045908	0.093624	0.4903	0.6238921
ar5	-0.085532	0.093641	-0.9134	0.3610281
ar6	0.193084	0.094506	2.0431	0.0410439 *
ar7	-0.069076	0.094829	-0.7284	0.4663501
ar8	0.141798	0.094882	1.4945	0.1350543
ar9	-0.041572	0.094980	-0.4377	0.6616101
ar10	-0.051811	0.095420	-0.5430	0.5871408
ar11	0.203687	0.096332	2.1144	0.0344794 *
ar12	-0.232393	0.070570	-3.2931	0.0009909 ***
intercept	0.169478	0.335981	0.5044	0.6139622

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Уочавамо да нису сви ауторегресиони коефицијенти статистички значајни, већ само *ar1*, *ar2*, *ar6*, *ar11* и *ar12*. Када мало коригујемо почетни модел, добијамо следећи *AR* модел:

$$X_t = 0.9X_{t-1} - 0.16X_{t-2} + 0.16X_{t-6} + 0.17X_{t-11} - 0.22X_{t-12} + e_t,$$

где је $\sigma_e^2 = 0.57$, а $AIC=461.13$.

```
Call:
arima(x = copper, order = c(12, 0, 0), fixed = c1)

Coefficients:
      ar1      ar2  ar3  ar4  ar5      ar6  ar7  ar8  ar9  ar10      ar11      ar12  intercept
 0.8984 -0.1596  0    0    0  0.1559  0    0    0    0  0.1671 -0.2166      0
s.e.  0.0674  0.0696  0    0    0  0.0456  0    0    0    0  0.0704  0.0696      0

sigma^2 estimated as 0.5666:  log likelihood = -224.57,  aic = 461.13
```

Алтернативни избор модела био би $ARMA(3,2)$.

$$X_t = -0.59X_{t-1} + 0.31X_{t-2} + 0.66X_{t-3} + e_t + 1.57e_{t-1} + 0.96e_{t-2},$$

где је $\sigma_e^2 = 0.59$, а $AIC=465.92$.

```
> auto.arima(copper, stepwise = FALSE, approximation = FALSE)
Series: copper
ARIMA(3,0,2) with zero mean

Coefficients:
      ar1      ar2      ar3      ma1      ma2
-0.5852  0.3070  0.6622  1.5709  0.9630
s.e.    0.0710  0.0962  0.0850  0.0564  0.0265

sigma^2 estimated as 0.5919:  log likelihood=-226.74
AIC=465.48  AICc=465.92  BIC=485.18
```

На основу слике број 11 приметили смо да би одговарајући модел могао да буде дворежимски TAR модел. Користећи функцију $tar()$ из пакета TSA , добијамо следећи TAR модел реда 12 који садржи само статистички значајне ауторегресионе коефицијенте:

```
> model<-tar(copper, 12, 12, 1)
> model$thd
0.9459653
> model$qr1$coefficients
intercept-copper      lag1-copper
 0.0809485          0.9519411
> model$qr2$coefficients
intercept-copper      lag1-copper      lag2-copper
 0.8258597          0.7184221      -0.3023816
> model$AIC
1
391.9
```

Уочавамо да су статистички значајна само прва два ауторегресиона коефицијента, тако да ћемо формирати дворежимски TAR модел реда 2:

$$X_t = \begin{cases} 0.08 + 0.93X_{t-1} + \sigma_1 e_t, & X_{t-1} \leq 0.946 \\ 0.73 + 0.73X_{t-1} - 0.29X_{t-2} + \sigma_2 e_t, & X_{t-1} > 0.946 \end{cases},$$


```

> model2<-tar(copper,2,2,1)
> model2$thd

0.9459653
> model2$sqr1$coefficients
intercept-copper      lag1-copper
      0.07536253      0.92867483
> model2$sqr2$coefficients
intercept-copper      lag1-copper      lag2-copper
      0.7303377      0.7336741      -0.2881161
> model2$AIC
      1
427.8
> model2$rms1
      1
0.3628672
> model2$rms2
      1
1.303754
> length(model2$y.regime1)
[1] 144
> length(model2$y.regime2)
[1] 51

```

Величине узорка за два режима су 144 и 51, респективно, а укупни AIC модела је 427.8.

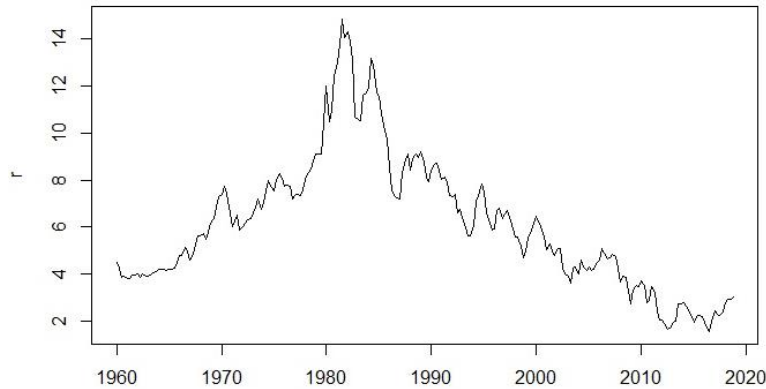
Сада ћемо упоредити формиране моделе за серију годишњих цена бакра. Уколико упоредимо два линеарна модела AR и $ARMA$, модел $AR(12)$ је мало бољи јер су има ниже вредности AIC и σ^2 . Уколико $AR(12)$ модел упредимо са $TAR(2,2,1)$ моделом, потребно је да коригујемо AIC модела $TAR(2,2,1)$ јер модели користе различите величине узорка. Код $AR(12)$ модела $T = 197$, док је код $TAR(2,2,1)T = 195$. Зато ћемо да прилагодимо величину ефективног узорка која се користи за израчунавање AIC када се користе различите класе модела. За $TAR(2,2,1)$ модел вредност AIC треба да буде:

$$AIC = 427.8 \cdot \frac{197}{195} = 432.2$$

а то је опет мања вредност од AIC вредности за линеарни $AR(12)$ који износи $AIC=461.13$.

Дакле, закључујемо да је најбољи $TAR(2,2,1)$ модел.

Пример 7. Дугорочна каматна стопа је важна променљива у актуарству. У овом примеру ће бити приказана техника изградње *TAR (SETAR)* модела користећи серију дугорочних каматних стопа америчке владе. Квартални подаци о приносу на дугорочне (десетогодишње) државне обвезнице САД-а (r_t) се могу наћи на сајту *FRED*-а³ у периоду од *Q1 1960.* до *Q4 2018.* године, са укупно 236 опсервација.



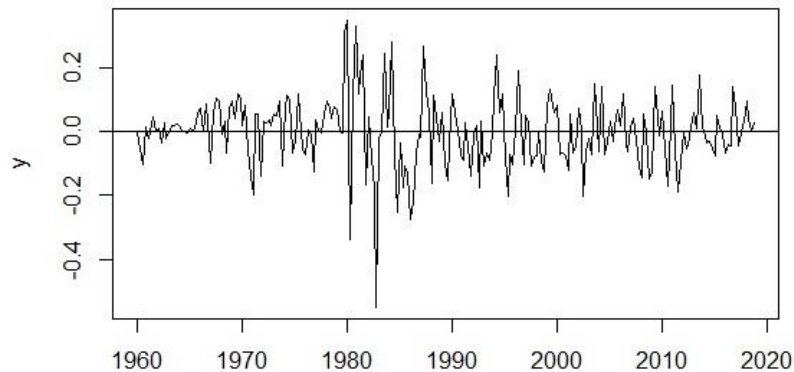
Слика бр. 12: Квартални подаци о приносу на дугорочне државне обвезнице, у процентима

Ефективна квартална каматна стопа за период t је дефинисана као:

$$i_t = \left[\left(1 + \frac{r_t}{100} \right)^{\frac{1}{4}} - 1 \right] \cdot 100 .$$

Анализу ће бити базирана на разлици логаритмованих вредности каматних стопа, која се дефинише са:

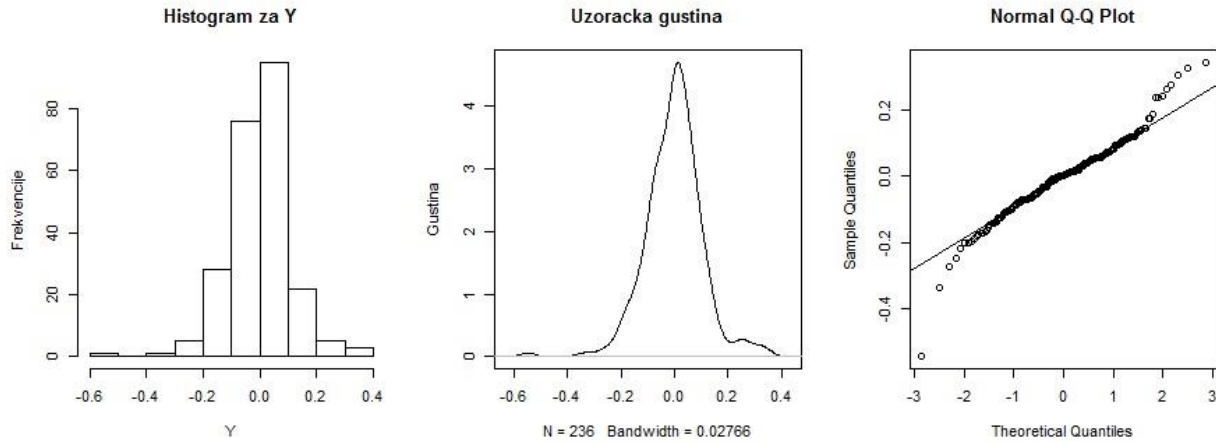
$$Y_t = \left[\ln \left(1 + \frac{i_t}{100} \right) - \ln \left(1 + \frac{i_{t-1}}{100} \right) \right] \cdot 100 .$$



Слика бр. 13: Разлике логаритмованих вредности каматних стопа

³Federal Reserve Bank of St. Louis
<https://fred.stlouisfed.org/series/IRLTLT01USM156N>

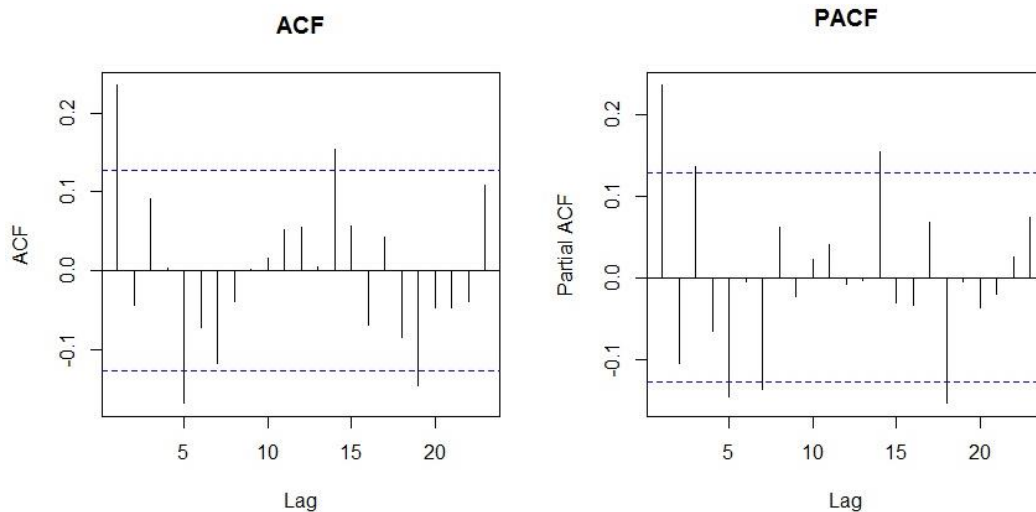
Емпиријска расподела посматране временске серије одступа од нормалне расподеле, а то може да се провери графички, као и помоћу коефицијента асиметрије (-0.33) и спљоштености (6.35).



Слика бр. 14: Хистограм, узорачка густина и $Q-Q$ дијаграм за временску серију Y_t

```
> skewness(Y)
[1] -0.3287837
> kurtosis(Y)
[1] 6.353632
```

Са графика $PACF$ за серију Y_t видимо да би $p = 5$ могао да буде одговарајући ред почетног AR модела. То се потврђује позивањем функције $auto.arima()$, при чему се фиксира $max.q=0$ како бисмо добили најбољи AR модел.



Слика бр. 15: ACF и $PACF$ за временску серију Y_t

```
> modell<-auto.arima(Y,max.q = 0,stepwise = FALSE, approximation = FALSE)
> modell
Series: Y
ARIMA(5,0,0) with zero mean

Coefficients:
      ar1      ar2      ar3      ar4      ar5
      0.2723 -0.1255  0.1312 -0.0234 -0.1444
s.e.  0.0643  0.0665  0.0664  0.0665  0.0640

sigma^2 estimated as 0.01046:  log likelihood=205.64
AIC=-399.27  AICc=-398.91  BIC=-378.49
```

Следеће што треба да одредимо је параметар кашњења d , који може да узима вредности $1 \leq d \leq 5$. У програму R покрећемо тестове за различите вредности $d = 1, \dots, 5$. Комбинација $(p, d) = (5, 1)$ даје најзначајнију F -статистику, што значи да ћемо прихватити $d = 1$.

```
> thr.test(Y,5,1)
SETAR model is entertained
Threshold nonlinearity test for (p,d):  5 1
F-ratio and p-value:  2.749741 0.01390068
> thr.test(Y,5,2)
SETAR model is entertained
Threshold nonlinearity test for (p,d):  5 2
F-ratio and p-value:  1.027835 0.408609
> thr.test(Y,5,3)
SETAR model is entertained
Threshold nonlinearity test for (p,d):  5 3
F-ratio and p-value:  1.913459 0.08072296
> thr.test(Y,5,4)
SETAR model is entertained
Threshold nonlinearity test for (p,d):  5 4
F-ratio and p-value:  1.655592 0.1342358
> thr.test(Y,5,5)
SETAR model is entertained
Threshold nonlinearity test for (p,d):  5 5
F-ratio and p-value:  0.2332143 0.9653069
```

Следећи корак је одређивање броја режима k , вредности прагова и оцењивање непознатих параметара. Као што је већ описано, праг мора да се налази између максималне и минималне вредности серије, а у пракси се добија када се из избора искључи 15% највиших и најнижих вредности. Након тога се за сваку од 70% унутрашњих опсервација проверава да ли је праг. Регресија која има најмању суму квадрата резидуала даје конзистентну процену прага. Редови AR модела унутар сваког режима j добијају се минимизирањем AIC критеријума за сваки од режимских линеарних модела ($0 \leq p_j \leq p$).

Када имамо познате вредности $p=5$ и $d=1$, покрећемо функцију $setar()$ за формирање $SETAR$ модела у програму R .

Добијени $SETAR$ модел за промене логаритамске серије дугорочних каматних стопа је:

$$Y_t = \begin{cases} 0.02 + 0.41Y_{t-1} + 0.02Y_{t-2} - 0.03Y_{t-3} + 0.10Y_{t-4} - 0.16Y_{t-5} + \sigma_1 e_t, & x_{t-1} \leq 0.00 \\ 0.14Y_{t-1} - 0.15Y_{t-2} + 0.26Y_{t-3} - 0.08Y_{t-4} - 0.1Y_{t-5} + \sigma_2 e_t, & x_{t-1} > 0.00 \end{cases}$$

Први режим представља окружење када се дугорочне каматне стопе смањују, а други режим када се оне повећавају. У првом режиму се налази 48%, а у другом 52% опсевација. Овако дефинисан *SETAR* модел пружа механизам за описивање асиметричног понашања дугорочне каматне стопе током различитих фаза пословног циклуса.

У наставку ће бити приказане карактеристике *SETAR* модела, као и графички приказ режима.

```
> model<-setar(Y,5,thDelay = 1)
> summary(model)

Non linear autoregressive model

SETAR model ( 2 regimes)
Coefficients:
Low regime:
  const.L    phiL.1    phiL.2    phiL.3    phiL.4    phiL.5
  0.01731159  0.40702647  0.02386473 -0.02867621  0.09768423 -0.15670889

High regime:
  const.H    phiH.1    phiH.2    phiH.3    phiH.4    phiH.5
 -0.004290141  0.138916200 -0.149479543  0.256236049 -0.077989896 -0.101018784

Threshold:
-Variable: Z(t) = + (0) x(t)+ (1)x(t-1)+ (0)x(t-2)+ (0)x(t-3)+ (0)x(t-4)
-Value: 0.001528
Proportion of points in low regime: 48.05%      High regime: 51.95%

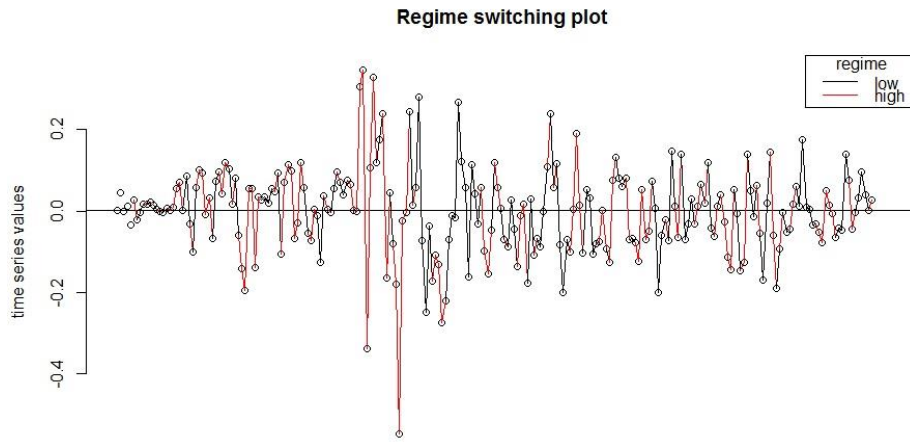
Residuals:
  Min       1Q   Median       3Q      Max
-0.4348330 -0.0649755  0.0068085  0.0548053  0.2901300

Fit:
residuals variance = 0.009623,  AIC = -1070,  MAPE = 193.2%

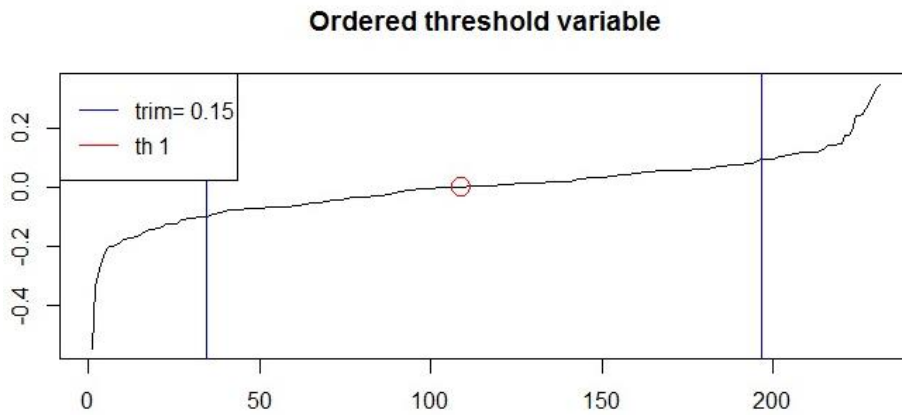
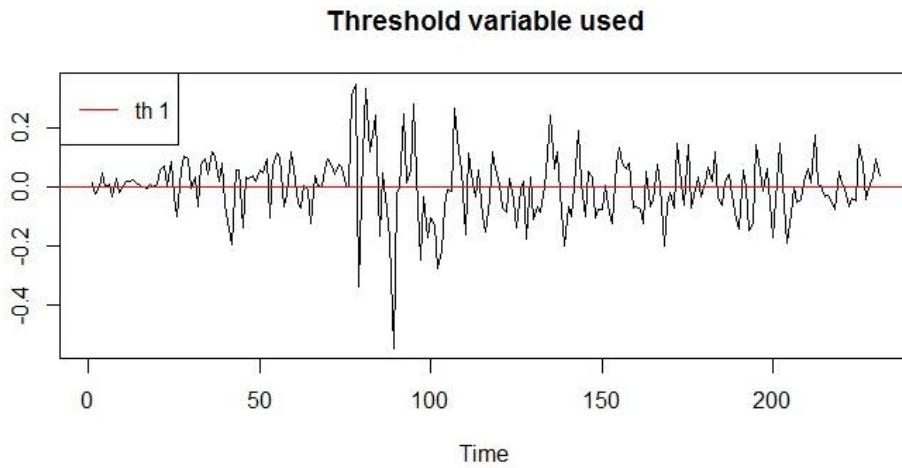
Coefficient(s):

      Estimate Std. Error t value Pr(>|t|)
const.L  0.0173116  0.0140198  1.2348  0.218201
phiL.1   0.4070265  0.0868914  4.6843  4.872e-06 ***
phiL.2   0.0238647  0.1265319  0.1886  0.850572
phiL.3  -0.0286762  0.0964854 -0.2972  0.766583
phiL.4   0.0976842  0.0957779  1.0199  0.308874
phiL.5  -0.1567089  0.0930855 -1.6835  0.093673 .
const.H -0.0042901  0.0133834 -0.3206  0.748844
phiH.1   0.1389162  0.0964389  1.4405  0.151134
phiH.2  -0.1494795  0.1356663 -1.1018  0.271723
phiH.3   0.2562360  0.0964112  2.6577  0.008433 **
phiH.4  -0.0779899  0.0951336 -0.8198  0.413205
phiH.5  -0.1010188  0.0902972 -1.1187  0.264451
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold
Variable: Z(t) = + (0) x(t) + (1) x(t-1)+ (0) x(t-2)+ (0) x(t-3)+ (0) x(t-4)
Value: 0.001528
```



Слика бр. 16: Графички приказ промене резима



Слика бр. 17: Графички приказ вредности прага

3.8.2. STAR модели

Недостатак *SETAR* модела је тај што функција условног очекивања није непрекидна, а режими могу да се дешавају у непрекидним интервалима. Тада прагови представљају тачке прекида функције условног очекивања μ_t . У циљу превазилажења овог недостатка, прекидност прагова се може заменити глатком функцијом и тада се предлаже употреба *STAR* модела.

Главна разлика између *TAR* (односно *SETAR*) и *STAR* модела је механизам који управља транзицијом између режима. За разлику од *SETAR* модела, где су промене параметара строго скоковите, овде оне могу бити највише глатко скоковите и због тога се ови модели називају *STAR* модели (енгл. *Smooth Transition AR*), тј. глатко прелазни *AR* модели са праговима.

Да би веза између различитих класа нелинеарних модела била јасна, користимо следећи *SETAR(2; p)* модел:

$$X_t = \begin{cases} \phi_{1,0} + \phi_{1,1}X_{t-1} + \dots + \phi_{1,p_1}X_{t-p} + \sigma_1 e_t, & X_{t-d} \leq r \\ \phi_{2,0} + \phi_{2,1}X_{t-1} + \dots + \phi_{2,p_2}X_{t-p} + \sigma_2 e_t, & X_{t-d} > r \end{cases},$$

који може да се запише и у облику:

$$X_t = \left(\phi_{1,0} + \sum_{i=1}^p \phi_{1,i}X_{t-i} + \sigma_1 e_t \right) (1 - I_t) + \left(\phi_{2,0} + \sum_{i=1}^p \phi_{2,i}X_{t-i} + \sigma_2 e_t \right) I_t,$$

где је:

$$I_t = \begin{cases} 1, & X_{t-d} \leq r \\ 0, & X_{t-d} > r \end{cases}.$$

Ако се у овом моделу индикаторска функција I замени непрекидном и глатком функцијом F , тада се промена свих параметара дешава истовремено и генерисана је функцијом F .

Углавном се узима функција F са скупом вредности $[0,1]$, као што су функције расподела вероватноћа случајних променљивих. Различити избори функције преласка F резултирају различитим врстама *STAR* модела.

Нека је $Z_t = X_{t-d}$ *threshold* променљива за $d > 0$. Ако је F функција локалне густине дефинисана са:

$$F(Z_t; \gamma, r) = \frac{1}{1 + e^{-\gamma(Z_t - r)}}, \quad \gamma > 0$$

резултирајући модел се зове логистички *STAR* (*LSTAR*) модел. Параметар γ одређује глаткоћу, тако да када $\gamma \rightarrow 0$ или када $\gamma \rightarrow \infty$, модел постаје сличан линеарном *AR*(p) моделу јер је функција $F(Z_t; \gamma, r)$ константна. За остале вредности γ , степен ауторегресионог расипања зависи од Z_t . Када $Z_t \rightarrow -\infty$, $F(Z_t; \gamma, r) \rightarrow 0$, а када $Z_t \rightarrow +\infty$ важи $F(Z_t; \gamma, r) \rightarrow 1$.

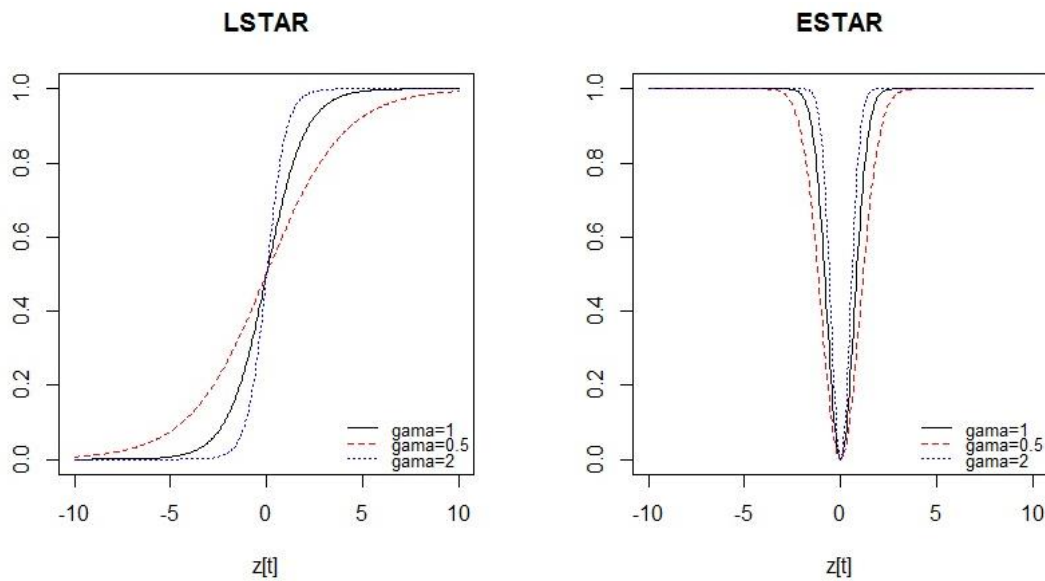
Још један тип *STAR* модела који се често помиње у литератури је експоненцијални *STAR* (*ESTAR*) модел код кога је експоненцијална функција дефинисана са:

$$F(Z_t; \gamma, r) = 1 - e^{-\gamma(Z_t - r)^2}.$$

Функција $F(Z_t; \gamma, r)$ садржи квадрат у експоненту, тако да су коефицијенти *ESTAR* модела симетрични око $Z_t = r$. Када $Z_t \rightarrow r$, тада $F(Z_t; \gamma, r) \rightarrow 0$, а када се Z_t удаљава од r , тада $F(Z_t; \gamma, r) \rightarrow 1$. Такође, када $\gamma \rightarrow 0$ или када $\gamma \rightarrow \infty$, модел постаје сличан линеарном *AR*(p) моделу јер је функција $F(Z_t; \gamma, r)$ константна.

Пример 8. У програму *R* су генерисане функције прелаза за *LSTAR* и *ESTAR* моделе. Вредност прага је $r = 0$, вредности Z_t се крећу у интервалу $[-10, 10]$, а параметар γ узима различите вредности $\gamma = \{0.5, 1, 2\}$.

Са графика *LSTAR* функције се примећује да *S* облик постаје оштрији како се параметар γ повећава, док са графика *ESTAR* функције *U* облик постаје оштрији како се γ повећава. Разлика између модела *LSTAR* и *ESTAR* је у томе што се логистичка функција приближава вредности 1 само кад $\gamma(Z_t - r) \rightarrow \infty$. С друге стране, експоненцијална функција је симетрична и приближава се 1 када је $Z_t \rightarrow \pm\infty$.

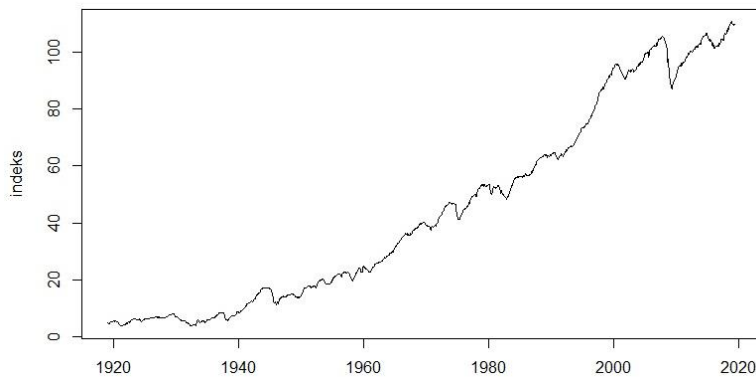


Слика бр. 18: *LSTAR* и *ESTAR* функције

Идеја о глатком прелазу је привлачна јер ће се промене динамичке зависности временских серија вероватно развијати постепено током времена. Предност *STAR* модела у односу на *TAR* моделе је диференцијабилност функције условног очекивања, међутим параметри функције F се тешко оцењују (γ и r) без обзира на избор функције преласка.

Дакле, код *STAR* модела се промена ауторегресионих параметара уводи функцијом F , а ако се за сваки параметар уведе нека нова функција која дефинише промену, добија се *FAR* модел. Дакле, *STAR* је специјални случај *FAR* (енгл. *Functional AR*) модела.

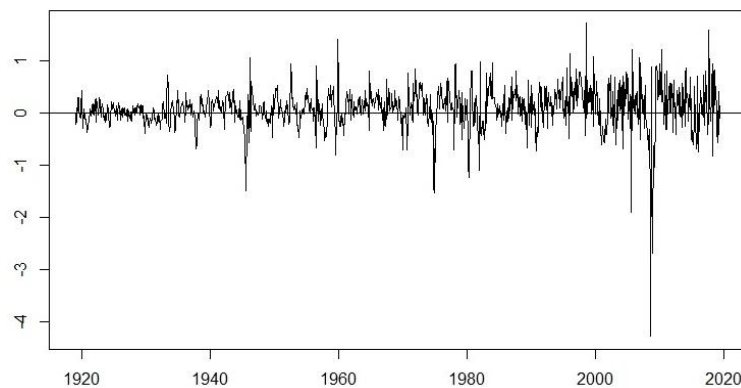
Пример 9. Индекси индустријске производње представљају краткорочне показатеље који могу да се користе као показатељи развоја тржишта индустријских производа и услуга произведених унутар индустрије. У наставку се користи *LSTAR* модел за индексе индустријске производње у САД-у, у периоду од јануара 1919. до јуна 2019. године⁴.



Слика бр. 19: Индекс индустријске производње у САД-у (десезонирани подаци)

Примећује се да је присутан растући тренд и због тога се у наставку рачунају месечне промене индекса индустријске производње на следећи начин:

$$X_t = IP_t - IP_{t-1}, \text{ где је } IP_t \text{ индекс у тренутку } t.$$



Слика бр. 20: Промене индекса индустријске производње

⁴ Подаци се могу наћи на: <https://fred.stlouisfed.org/series/INDPRO>

Прво ћемо да тестирамо присуство нелинеарности. Резултати тестова показују да можемо да одбацимо нулту хипотезу, што значи да се подаци о променама индекса индустријске производње могу моделирати нелинеарним моделима (у овом случају *LSTAR* моделом).

```

** McLeod-Li test **
Null hypothesis: The time series follows some ARIMA process
Maximum p-value = 0.009263207

** Tsay's Test for nonlinearity **
Null hypothesis: The time series follows some AR process
F-stat = 2.437 p-value = 0.0002104

```

Поступак за изградњу *LSTAR* модела је веома сличан изградњи *TAR (SETAR)* модела. Након одређивања реда $p = 4$ полазног *AR* модела, тестирањем нелинеарности за различите вредности параметра кашњења d , добија се да је најбоље да се користи параметар кашњења $d = 3$.

```

Non linear autoregressive model

LSTAR model
Coefficients:
Low regime:
  const.L      phiL.1      phiL.2      phiL.3      phiL.4
0.153468537 -0.004415883  0.079411783 -0.222959642 -0.037767206

High regime:
  const.H      phiH.1      phiH.2      phiH.3      phiH.4
-0.13264809  0.31453300  0.02751424  0.21442691  0.08696848

Smoothing parameter: gamma = 16.23

Threshold
Variable: Z(t) = + (0) X(t) + (0) X(t-1)+ (0) X(t-2)+ (1) X(t-3)
Value: -0.1651

Residuals:
      Min          1Q      Median          3Q          Max
-4.1275507 -0.1646015  0.0057336  0.1894469  1.5835136

Fit:
residuals variance = 0.1351, AIC = -2390, MAPE = 238.7%
Coefficient(s):
      Estimate Std. Error t value Pr(>|z|)
const.L  0.1534685  0.0481034  3.1904 0.0014208 **
phiL.1  -0.0044159  0.0905925 -0.0487 0.9611229
phiL.2   0.0794118  0.0686597  1.1566 0.2474361
phiL.3  -0.2229596  0.0537155 -4.1508 3.314e-05 ***
phiL.4  -0.0377672  0.0692132 -0.5457 0.5852962
const.H  -0.1326481  0.0641582 -2.0675 0.0386855 *
phiH.1   0.3145330  0.0945881  3.3253 0.0008833 ***
phiH.2   0.0275142  0.0822689  0.3344 0.7380455
phiH.3   0.2144269  0.0786549  2.7262 0.0064073 **
phiH.4   0.0869685  0.0785494  1.1072 0.2682154
gamma   16.2307728 13.7639310  1.1792 0.2383086
th      -0.1651224  0.0650542 -2.5382 0.0111415 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Non-linearity test of full-order LSTAR model against full-order AR model
F = 4.3914 ; p-value = 0.0015795

Threshold
Variable: Z(t) = + (0) X(t) + (0) X(t-1)+ (0) X(t-2)+ (1) X(t-3)

```

У програму R помоћу функције $lstar()$ добија се следећи модел:

$$X_t = (0.15 + 0.08X_{t-2} - 0.22X_{t-3} - 0.04X_{t-4})(1 - F(X_{t-3}; \gamma, r)) \\ + (-0.13 + 0.31X_{t-1} + 0.03X_{t-2} + 0.21X_{t-3} + 0.09X_{t-4})F(X_{t-3}; \gamma, r)$$

где је $F(X_{t-3}; \gamma, r) = \frac{1}{1 + e^{-16.23(X_{t-3} + 0.17)}}$.

3.8.3. Још неки TAR модели

$MTAR$ модели

Ако се избор подмодела врши у зависности од тога да ли X_{t-1} расте ($\Delta X_{t-1} \geq 0$) или опада ($\Delta X_{t-1} < 0$), онда се ради о $MTAR$ (енгл. *Momentum TAR*) моделу:

$$\Delta X_t = \begin{cases} a_0 + a_1 X_{t-1} + \dots + a_p X_{t-p}, & \Delta X_{t-1} < 0 \\ b_0 + b_1 X_{t-1} + \dots + b_p X_{t-p}, & \Delta X_{t-1} \geq 0 \end{cases}.$$

Овакви модели се користе за моделирање асиметричности која је честа појава код економских временских серија.

MA модели са праговима

Аналогно AR моделима са праговима могу да се дефинишу и MA модели са праговима r_1, r_2, \dots, r_{m-1} .

Дефиниција 6. $TMA(m; q_1, q_2, \dots, q_m)$ модел је стохастичка диференцна једначина

$$X_t = \left(b_{j,0} + \sum_{i=1}^{q_j} b_{j,i} e_{j,t-i} + e_{j,t} \right) I(Z_{t-d} \in R_j),$$

где је $e_{j,t} = \sigma_j e_t$, $0 < \sigma_j < \infty$.

Процес $\{X_t\}$ је $TMA(m; q_1, q_2, \dots, q_m)$ процес ако је генерисан претходним моделом.

У овом случају, X_t има један од MA подмодела у зависности од тога у ком интервалу се налази Z_{t-d} . За $Z_t = X_t$ добију се $SETMA$ модели.

Оно што је интересантно је да се $SETMA$ и MA модели не могу разликовати по аутокорелативној функцији, али се могу разликовати по условној аутокорелативној функцији дефинисаној са:

$$\rho_{\pm}(k|d) = \begin{cases} \frac{\text{cov}(X_t, X_{t-k} | X_{t-d} \geq c)}{D(X_{t-d} | X_{t-d} \geq c)} \\ \frac{\text{cov}(X_t, X_{t-k} | X_{t-d} < c)}{D(X_{t-d} | X_{t-d} < c)} \end{cases},$$

где + одговара случају када је $X_{t-d} \geq c$, а – када је $X_{t-d} < c$.

ARMA модели са праговима

Нека су интервали R_j дефинисани са $R_j \in (r_{j-1}, r_j]$, а $r_0 = -\infty, r_t = +\infty$, где су r_1, r_2, \dots, r_{m-1} прагови, m, p, q и d природни бројеви, а $a_{j,i}, i = 1, 2, \dots, p, j = 1, 2, \dots, m$ и $b_{j,i}, i = 1, 2, \dots, q, j = 1, 2, \dots, m$ реални бројеви.

Дефиниција 7. Стохастичка диференцна једначина

$$X_t = \sum_{j=1}^m \left(a_{j,0} + \sum_{i=1}^p a_{j,i} X_{t-i} + \sum_{i=1}^q b_{j,i} e_{t-i} \right) I(Z_{t-d} \in R_j) + e_t,$$

где су $X_0, X_{t-1}, \dots, X_{t-p+1}$ случајне променљиве, а $\{e_t\}$ низ иноваријација, је ARMA модел реда (m, p, q) са праговима r_1, r_2, \dots, r_{m-1} и кашњењем d .

Овако дефинисан модел се састоји од m ARMA подмодела и за њега се користи ознака $TARMA(m, p, q)$.

Процес $\{X_t\}$ је $TARMA(m, p, q)$ процес ако је генерисан претходним моделом. За $m = 1$ добија се ARMA(p, q) процес. Уколико је $Z_t = X_t$ ради се о SETARMA(m, p, q) процесу.

TAR модели са случајним коефицијентима

Нека је J_t случајна променљива дефинисана са:

$$J_t = \begin{cases} 1, & Z_{t-d} \in R_1 \\ 2, & Z_{t-d} \in R_2 \\ \dots & \dots \dots \dots \\ m, & Z_{t-d} \in R_m \end{cases}.$$

Модел $TAR(m; p_1, p_2, \dots, p_l)$ може да се запише у облику модела са случајним коефицијентима:

$$X_t = a_0(J_t) + \sum_{i=1}^{p(J_t)} a_i(J_t) X_{t-i} + e_t(J_t),$$

где је $a_i(J_t) = a_{j,i}$ за $J_t = j (j = 1, \dots, l; i = 1, \dots, p_j)$.

4. Закључак

Финансијске временске серије условљене разним утицајима на тржишту, поседују особине које их разликују од других временских серија, као што су асиметричност и присуство нестандартних опсервација које се појављују у групама. Кључни значај при доношењу финансијских одлука има предвиђање будућих вредности серија. Како линеарни модели не могу на адекватан начин да опишу финансијске временске серије, јавила се потреба за увођењем нелинеарних модела, а последњих неколико деценија интересовање за нелинеарним моделима је значајно порасло.

У овом раду су представљене најзаступљеније класе нелинеарних модела, а то су билинеарни модели, модели са случајним коефицијентима и модели са праговима (*TAR*, *SETAR* и *STAR*). Пре увођења ових модела описани су тестови којима може да се тестира присуство нелинеарности. Осим тога, представљене су и особине ових модела, као и бројне примене на реалним подацима.

Иако су линеарни модели због једноставне примене веома заступљени у истраживањима и анализама временских серија, нелинеарни модели због својих својстава добијају све више на значају и веома су корисни и применљиви у моделирању и прогнозирању финансијских временских серија.

5. Литература

- [1] R.S.Tsay, *Analysis of Financial Time Series (Second edition)*, New York: John Wiley & Sons, 2005.
- [2] Ruey S. Tsay, Rong Chen, *Nonlinear Time Series Analysis*, New York: John Wiley & Sons, 2019.
- [3] Walter Enders, *Applied econometric time series (Fourth edition)*, New York: John Wiley & Sons, 2015.
- [4] Jianqing Fan, Qiwei Yao, *Nonlinear time series: nonparametric and parametric methods*, Springer-Verlag New York, 2003.
- [5] Eric Ghysels, Massimiliano Marcellino, *Applied economic forecasting using time series methods*, United States of America by Oxford University Press, 2018.
- [6] Слободанка Јанковић, Бојана Милошевић, *Елементи финансијске математике*, Математички факултет, 2017.
- [7] T. Subba Rao, *On the Theory of Bilinear Time Series Models*, Journal of the Royal Statistical Society. Series B (Methodological), Vol. 43, No. 2, 1981.
- [8] T.D.Pham, L.T.Tran, *On the First Order Bilinear Time Series Model*, J. Applied Probability, vol 18, 1981.
- [9] Des F. Nicholls, Barry G. Quinn, *Random Coefficient Autoregressive Models: An Introduction*, Springer-Verlag New York, 1982.
- [10] Jiří Anděl, *Autoregressive series with random parameters*, Mathematische Operationsforschung und Statistik: A Journal of Theoretical and Applied Statistics, 2007.
- [11] Chan, Wong, Tong, *Some Nonlinear Threshold Autoregressive Time Series Models for Actuarial Use*, North American Actuarial Journal, volume 8, number 4
- [12] Randal Douc, Eric Moulines, David Stoffer, *Nonlinear Time Series: Theory Methods and Applications with R Examples*, CRS Press, Taylor&Francis Group, 2014.
- [13] Y.Liang, A.Thavaneswaren, N.Ravishanker, *RCA models: Joint prediction of mean and volatility*, Statistics and Probability Letters 83. 2013.
- [14] K.S. Chan, *Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model*, The Annals of Statistics, Vol. 21, No. 1, 1993.

- [15] Rong Chen, Ruey Tsay, *On the Ergodicity of TAR(1) Processes*, The Annals of Statistics, Vol. 1, No. 4, 1991.
- [16] <https://www.stat.tamu.edu/~suhasini/teaching673/chapter4.pdf>
- [17] <https://fred.stlouisfed.org/series/IRLTLT01USM156N>
- [18] <https://fred.stlouisfed.org/series/INDPRO>
- [19] <https://finance.yahoo.com/quote/%5EGSPC/>
- [20] <https://faculty.chicagobooth.edu/ruey.tsay/teaching/nts/>

Биографија

Јована Недељковић рођена је 6.1.1992. године у Београду. Завршила је Земунску гимназију, природно-математички смер са одличним успехом. Основне студије на Математичком факултету Универзитета у Београду, смер Статистика, актуарска и финансијска математика завршила је 2016. године, а академске 2016/2017. године уписала је мастер студије на истом смеру.

Радно искуство започела је 2017. године у Републичком заводу за статистику након обављене стручне праксе. Од новембра 2017. године запослена је у Народној банци Србије у Сектору за економска истраживања и статистику. Током радне каријере усавршавала се у земљи и иностранству и била учесник бројних семинара у организацији Међународног монетарног фонда, *Eurostat*-а и европских централних банака из области десезонирања, макроекономије, платног биланса и страних директних инвестиција.