УНИВЕРЗИТЕТ У БЕОГРАДУ
МАТЕМАТИЧКИ ФАКУЛТЕТ

Дамјан Ж. Петровић

МОДЕЛОВАЊЕ СТРУКТУРНОГ ЛОМА У ВРЕМЕНСКОЈ СЕРИЈИ

мастер рад

Београд, 2023.

Ментор:

др Јелена Јоцковић, доцент
Универзитет у Београду, Математички факултет

Чланови комисије:

др Павле Младеновић, редовни професор
Универзитет у Београду, Математички факултет

др Марко Обрадовић, доцент
Универзитет у Београду, Математички факултет

Датум одбране: 14.06.2023.

Хвала Јовани, Живораду и Власти на несебичној љубави и подршци. Хвала професорки Јоцковић на бескрајном стрпљењу.

Наслов мастер рада: Моделовање структурног лома у временској серији

Резиме: У раду се разматрају начини којима се моделира присуство структурног лома у кретању временске серије у контексту различитих врста модела. Значај овог проблема лежи у томе што је структурни лом честа појава код економских временских серија, као једна од њихових специфичности. Полазећи од чињенице да је стуктурни лом заправо скуп нестандардних опсервација у временској серији, можемо закључити да занемаривање његовог постојања приликом моделирања доводи до проблема у оцењивању параметара, као и прогнозирању будућег кретања серије. У раду се разматрају и методе тестирања којима се детектује постојање структурног лома, са посебним освртом на однос између присуства структурног лома и јединичног корена у временској серији.

Кључне речи: временска серија, интервенција, нестандрадна опсервација, структурни лом, ARMA модел, регресиони модел, јединични корен

Садржај

1 Увод **1**

2 Појам временске серије **3**

2.1 Временска серија 3

2.2 Стационарност 4

2.3 Аутоковаријациона и аутокорелациона функција 5

3 Моделовање стационарних временских серија **6**

3.1 Линеарни процес 6

3.2 Ауторегресиони (AR) модели 8

AR(p) модел 10

3.3 Модели покретних просека (MA) 11

MA(q) модел 11

3.4 Ауторегресиони модели покретних просека (ARMA) 12

3.5 Ауторегресиони интегрисани модели покретних просека (ARIMA) 13

4 Моделовање структурног лома **15**

4.1 Интервенција 15

4.2 Моделовање структурног лома - познат тренутак интервенције 17

4.3 Моделовање структурног лома - непознат тренутак интервенције 28

Адитивни и иновативни модели 28

Општи регресиони модел 32

5 Детектовање структурног лома **37**

5.1 Анализа графика 37

5.2 Доменско знање 38

5.3 Тестови за утврђивање присуства структурног лома 39

Жарк-Бера тест нормалности 39

Чоу тест 40
CUSUM тест 42
Секвенцијални тест 43
Тестирање присуства структурног лома у случају хетероскедастичне серије 44
5.4 Структурни лом код серија са јединичним кореном 44

6 Практични примери **50**
6.1 Стопа незапослености 50
6.2 Цена акција на берзи 55

7 Закључак **63**

Библиографија **65**

Глава 1

Увод

Временске серије често су под утицајем спољашњих, односно егзогених догађаја. Термин егзогени односи се на појаву да утицај који овакви догађаји врше на кретање серије, долази изван простора покривеног моделом којим је дата серија описана, те самим тим тај утицај није објашњен нити предвиђен приликом моделовања. Овакве догађаје називамо интервенцијама. Неки од најчешћих примера интервенција су промене економске политике, као што су промена каматне стопе, промена режима девизног курса, увођење нове таксе на увоз робе, реформа пореског система и слично. То могу бити и непланирани догађаји као што су штрајкови, прекиди у снабдевању електричном енергијом, природне непогоде итд. Временске серије чије понашање је од интереса при наведеним догађајима су најчешће економске природе, као на пример стопа запослености, бруто друштвени производ, индекс индустријске производње, цена неке сировине и слично. Иако су економске серије најзначајније за проучавање, интересантне примере можемо пронаћи и у другим областима. Приликом испитивања новог лека, значајно је анализирати како његово увођење у терапију утиче на неки од показатеља здравственог стања пацијента, нпр. ниво шећера у крви. Такође, након доношења одређеног законског прописа, потребно је утврдити да ли су постигнути жељени резултати, те се на пример може анализирати утицај закона о обавезном везивању појаса у аутомобилу на број саобраћајних несрећа са смртним исходом.
У зависности од дужине трајања дејства све интервенције делимо у две групе. Разликујемо једнократне интервенције (енгл. pulse intervention), чије дејство траје само одређен временски период, и трајне интервенције (енгл. step intervention), чије дејство траје до даљњег. Још једна подела обрађена у овом

раду односи се на тренутак у коме се јавља сама интервенција. Наиме, у зависности од тога да ли нам је тај тренутак априори познат или не, примењују се различити поступци анализе и моделовања.

Последица деловања интервенције на временску серију је структурни лом, односно скуп опсервација које нису у складу са дотадашњим кретањем временске серије. Последично, уколико би присуство структурног лома било било занемарено приликом моделовања серије, то би се негативно одразило на оцене параметара модела, као што су оцена варијансе, аутокорелационе функције и парцијалне аутокорелационе функције као и на поузданост прогнозирања будућих вредности серије.

Други део рада посвећен је појму временске серије и њеним најважнијим карактеристикама, као што су стационарност и корелациона структура. Треће поглавље односи се на ARMA моделе којима се описује стационарно кретање серије. Четврто поглавље посвећено је моделовању структурног лома. У првом делу овог поглавља обрађене су модификације ARMA модела, којима се моделује појава структурног лома, при чему се претпоставља да је тренутак интервенције која је узроковала структурни лом познат. Други део је посвећен моделима структурног лома у случају непознатог тренутка интервенције, при чему је посебна пажња посвећена општем регресионом моделу. У петом делу рада бавимо се техникама за детекцију промена у уобичајеном кретању временске серије у ситуацији када тренутак интервенције која је те промене изазвала није познат, односно тестовима за детектовање структурног лома у кретању серије. Посебно је обрађена тема међусобног односа структурног лома и јединичног корена, односно метода за тестирање присуства јединичног корена у ситуацији када је серија изложена структурном лому. У последњем, шестом делу рада дата су два практична примера, као илустрација примене неких од изложених теоријских резултата.

Глава 2

Појам временске серије

У овом поглављу објаснићемо основне појмове везане за анализу временских серија. Дефинисаћемо сам појам временске серије, својство стационарности, аутоковаријациону и аутокорелациону функцију.

2.1 Временска серија

Низ опсервација $\{X_t, t \in T\}$, где је T дискретан скуп временских тачака које су на једнаком растојању једна од друге, назива се **временска серија**. Дакле, временска серија је низ опсервација које су временски уређене. Другим речима, временска серија је дискретан случајни процес дефинисан на скупу еквидистантних временских тачака. Ово није једина дефиниција временске серије, која се такође може посматрати као једна конкретна реализација оваквог случајног процеса [9, p. 15].
Примери временске серије су цене акција неке компаније на берзи бележене сваког дана у исто време, температура ваздуха мерена сваког дана у исто време и на истом месту, број позива хитној помоћи на дневном нивоу, број путника на летовима неке авиокомпаније на месечном нивоу, потрошња електричне енергије у једној држави на месечном нивоу итд.

2.2 Стационарност

Једно од најважнијих својстава временских серија је **својство стационарности**[1]. Разликујемо концепте строге и слабе стационарности.
Строга стационарност (стационарност у ужем смислу) представља непромењивост својстава серије (очекивање, варијанса) током времена. Другим речима, временска серија је строго стационарна ако за произвољне $n \in \mathbb{N}$ и $k \in \mathbb{N}$ и произвољну n-торку $(t_1, t_2, ..., t_n)$ важи да n-димензионе случајне променљиве $(X_{t_1}, X_{t_2}, ..., X_{t_n})$ и $(X_{t_1+k}, X_{t_2+k}, ..., X_{t_n+k})$ имају исту расподелу вероватноћа.
Слаба стационарност (стационарност у ширем смислу) је својство које серија поседује уколико су њено математичко очекивање и дисперзија константни током времена, а коваријација између свака два члана серије зависи само од њиховог међусобног временског растојања. Дакле, серија је слабо стационарна ако важи:

$$
\begin{aligned}
&E(X_t) = \mu = const, \forall t \in T \\
&E(X_t^2) < \infty, \forall t \in T \\
&cov(X_t, X_{t-k}) = \gamma(k), \forall t \in T, \forall k \in \mathbb{N}
\end{aligned}
$$

Вреди истаћи да дефиниције строге и слабе стационарности нису еквивалентне. Наиме, стргого стационарна временска серија је истовремено слабостационарна само под условом да поседује коначан други моменат. Са друге стране слабо стационарна серија не мора бити и строго стационарна, уколико моменти реда већег од два нису стабилни.
Из претходних разматрања јасно је да својство стационарности представља гаранцију предвидивог кретања серије, што је кључан предуслов за креирање модела, помоћу којих се врши прогнозирање будуће вредности. Уколико временска серија која је предмет анализе није стационарна, она се, приликом моделовања, различитим поступцима (нпр. диференцирањем) трансформише у стационарну. Због тога ће се теоријска разматрања у овом раду ослањати на важење претпоставке о стационарности серије.

[1]Строго говорећи, стационарност је својство случајног процеса, па се у наставку ослањамо на дефиницију по којој временску серију идентификујемо са случајним процесом.

2.3 Аутоковаријациона и аутокорелациона функција

Основна карактеристика временских серија, која их разликује од простог случајног узорка у класичној статистичкој анализи, представља постојање корелације између чланова. Аутоковаријациона и аутокорелациона функција представљају фундаментално средство за анализу временских серија, којима се описује корелациона структура серије. **Аутоковаријациона функција** је низ аутоковаријационих коефицијената $\gamma_k, k \in \mathbb{N}_0$, где је:

$$\gamma_k = cov(X_t, X_{t-k}) = E(X_t - \mu)(X_{t-k} - \mu) \tag{2.1}$$

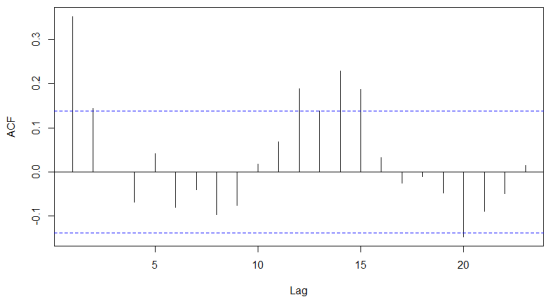
Аутокорелациона функција је низ аутокорелационих коефицијената $\rho_k, k \in \mathbb{N}_0$, где је:

$$\begin{aligned} \rho_k =& \frac{cov(X_t, X_{t-k})}{\sqrt{var(X_t)var(X_{t-k})}} \\ =& \frac{cov(X_t, X_{t-k})}{var(X_t)} \\ =& \frac{E(X_t - \mu)(X_{t-k} - \mu)}{E(X_t - \mu)^2} \end{aligned} \tag{2.2}$$

Из 2.1 и 2.2 следи да је:

$$\rho_k = \frac{\gamma_k}{\gamma_0}, k \in \mathbb{N}_0 \tag{2.3}$$

Графички приказ аутокорелационе функције назива се корелограм, и приказан је на слици 2.1.



Слика 2.1: Пример корелограма

Глава 3

Моделовање стационарних временских серија

У овом поглављу анализираћемо најчешће коришћене моделе за описивање стационарних временских серија. Излагање ћемо почети разматрањем линеарног процеса као полазног модела. Затим ћемо се бавити једноставнијим моделима, као што су ауторегресиони модели и модели покретних просека, а након тога ћемо описати ауторегресионе моделе покретних просека, као њихову комбинацију. На крају, даћемо преглед интегрисаних ауторегресионих модела покретних просека.

3.1 Линеарни процес

Према Волдовој теореми разлагања[1], свака слабо стационарна временска серија може се разложити на збир два међусобно некорелисана процеса, при чему је један детерминистички, а други стохастички. Додатно, стохастичка компонента може се представити у следећем облику:

$$X_t - \mu = e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \psi_3 e_{t-3} + ... \tag{3.1}$$

или краће:

$$X_t - \mu = \sum_{k=0}^{\infty} \psi_k e_{t-k}, \text{где је } \psi_0 = 1 \tag{3.2}$$

[1]Herman Ole Andreas Wold

У претходним изразима X_t је вредност серије у тренутку t, а μ означава детерминистичку компоненту серије, коју ћемо у даљим разматрањима посматрати као константну средњу вредност. Са e_t означен је процес белог шума, који се одликује нултом средњом вредношћу, константном варијансом и некорелисаношћу чланова у различитим временским тренуцима:

$$E(e_t) = 0, \quad var(e_t) = \sigma^2, \quad cov(e_t, e_{t-k}) = \begin{cases} \sigma^2, k = 0 \\ 0, k \neq 0 \end{cases} \tag{3.3}$$

Линеарна комбинација чланова процеса белог шума са ψ-пондерима назива се **линеарни процес**. Смисао датог израза је у томе да се стохастичка компонента временске серије, односно онај део чије се понашање не може предвидети на основу кретања саме серије у прошлости (одступање серије од средње вредности), представља као пондерисани збир случајних шокова (описаних члановима белог шума) који су резултат непредвиђених утицаја на серију током времена. Користећи оператор доцње L, за који важи $LX_t = X_{t-1}$, линеарни процес можемо записати на следећи начин:

$$X_t - \mu = e_t + \psi_1 L e_t + \psi_2 L^2 e_t + ... = (1 + \psi_1 L + \psi_2 L^2 + ...)e_t = \Psi(L)e_t \tag{3.4}$$

где је $\Psi(L)$ полином по оператору L. Значај линеарног процеса лежи у чињеници да се погодним избором ψ-пондера могу извести сви модели временских серија којима се бавимо у овом поглављу. Због тога ћемо извести изразе за очекивање, варијансу, аутоковаријациону и аутокорелациону функцију.

1. Варијанса:

$$\begin{aligned} var(X_t - \mu) = var(X_t) = E(X_t - \mu)^2 &= E(\sum_{k=0}^{\infty} \psi_k e_{t-k})^2 \\ &= E(\sum_{k=0}^{\infty} \psi_k^2 e_{t-k}^2 + 2\sum_{i<j} \psi_i e_{t-i} \psi_j e_{t-j}) \\ &= \sum_{k=0}^{\infty} E(\psi_k^2 e_{t-k}^2) + 2\sum_{i<j} \psi_i \psi_j E(e_{t-i} e_{t-j}) \qquad (3.5) \\ &= \sigma^2 \sum_{k=0}^{\infty} \psi_k^2 \qquad (3.6) \end{aligned}$$

Израз 3.6 следи из 3.5 јер на основу особина белог шума важи:

$$E(e_t^2) = var(e_t) + (E(e_t))^2 = \sigma^2$$
$$E(e_{t-i} e_{t-j}) = 0, \quad i \neq j$$

Закључујемо да је варијанса линеарног процеса коначна акко важи $\sum_{k=0}^{\infty} \psi_k^2 < +\infty$.

2. Аутоковаријација:

$$\begin{aligned}
\gamma_k &= cov(X_t, X_{t-k}) = E(X_t - \mu)(X_{t-k} - \mu) \\
&= E(e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + ... + \psi_k e_{t-k} + \psi_{k+1} e_{t-k-1} + ...) \\
&(e_{t-k} + \psi_1 e_{t-k-1} + \psi_2 e_{t-k-2} + ...) \\
&= E(\psi_k e_{t-k}^2 + \psi_1 \psi_{k+1} e_{t-k-1}^2 + \psi_2 \psi_{k+2} e_{t-k-2}^2 + ...) \\
&= \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{k+i}
\end{aligned}$$

3. Аутокорелација:

$$\begin{aligned}
\rho_k &= \frac{\gamma_k}{\gamma_0} \\
&= \frac{\sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{k+i}}{\sigma^2 \sum_{k=0}^{\infty} \psi_k^2} \\
&= \frac{\sum_{i=0}^{\infty} \psi_i \psi_{k+i}}{\sum_{k=0}^{\infty} \psi_k^2}
\end{aligned}$$

Примећујемо да у изразу за линеарни процес фигурише бесконачно много ψ-коефицијената, те се, како би њиме била описана стационарна временска серија, природно намеће ограничење у виду апсолутне конвергенције реда ових коефицијената, односно услов $\sum_{k=0}^{\infty} |\psi_k| < +\infty$. Са друге стране, ако је варијанса линеарног процеса коначна, следи да је услов апсолутне конвергенције поменутог реда испуњен.

3.2 Ауторегресиони (AR) модели

Прва класа модела изведених из линеарног процеса којима се бавимо су ауторегресиони модели. Суштина ових модела је да се садашња вредност серије регресира на вредности серије из прошлости, те отуда потиче и сам назив. Како бисмо извели израз за AR модел, потребна нам је алтернативна ауторегресиона репрезентација линеарног процеса:

$$X_t = \pi_1 X_{t-1} + \pi_2 X_{t-2} + \pi_3 X_{t-3} + ... + e_t \tag{3.7}$$

при чему претпостављамо да је $E(X_t) = 0$. Из овакве репрезентације модела видимо да је вредности серије у тренутку t линеарна комбинација вредности серије у претходним тренуцима, уз додатак члана белог шума e_t, који укључује све нове информације у тренутку t које нису објашњене кретањем серије у прошлости. Извешћемо везу између Ψ-пондера у стандардној репрезентацији линеарног процеса 3.1 и π-пондера из ауторегресионе репрезентације 3.7. Почињемо од полинима по оператору доцње $\Psi(L)$ из једнакости 3.4:

$$X_t = \Psi(L)e_t \tag{3.8}$$

На сличан начин користимо полином $\Pi(L)$ за AR репрезентацију:

$$\Pi(L)X_t = e_t \tag{3.9}$$

Одатле и из 3.7 следи:

$$\Pi(L)X_t = X_t - \pi_1 X_{t-1} - \pi_2 X_{t-2} - \pi_3 X_{t-3} - ... = (1 - \sum_{k=1}^{\infty} \pi_k L^k)X_t$$

Када једнакост 3.9 помножимо полиномом $\Psi(L)$ добијамо $\Psi(L)\Pi(L)X_t = \Psi(L)e_t = X_t$, одакле следи

$$\Psi(L)\Pi(L) = 1$$
$$\Pi(L) = \Psi(L)^{-1}$$

Последња једнакост означава везу између ауторегресионе модела и стандардне репрезентације линеарног процеса.
У практичним применама, наведени ауторегресиона репрезентација са бесконмачним бројем коефицијената није употребљива, те се за моделовање временских серија користе AR модели са коначним бројем коефицијената. Дакле, ако претпоставимо да у једнакости 3.7 важи $\pi_1 = \phi_1, \pi_2 = \phi_2, ..., \pi_p = \phi_p, \pi_k = 0, k > p$, где је $p > 0$, добијамо једначину ауторегресионог модела реда p:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + ... + \phi_p X_{t-p} + e_t \tag{3.10}$$

У наставку ћемо описати конкретни AR модел реда p, и извести везу са линеарним процесом.

AR(p) модел

Као што се може видети из једнакости 3.10 садашња вредност серије је једнака линеарној комбинацији претходних p чланова серије, уз додатак члана белог шума e_t, који представља део необјашњен вредностима временске серије из прошлости. Модел се може представити коришћењем оператора доцње L:

$$(1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - ... - \phi_p L^p) X_t = e_t \tag{3.11}$$

Услов стационарности AR(p) модела задатог једначином 3.11 односи се на постојање оператора инверзног оператору $1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - ... - \phi_p L^p$. Овом моделу се придружује карактеристична једначина облика:

$$g^p - \phi_1 g^{p-1} - \phi_2 g^{p-2} - \phi_3 g^{p-3} - ... - \phi_p = 0. \tag{3.12}$$

Решења ове једначине одређују услов стационарности, и то на следећи начин [6, p. 55]:

1. 1. Ако за свако решење карактеристичне једначине $g_1, g_2, ..., g_p$ важи $|g_k| < 1$, тада је серија описана AR(p) моделом стационарна.
2. 2. Ако бар за једно од решења важи $|g_k| = 1$, док су сва остала решења по модулу мањи од 1, тада је серија нестационарна, и назива се **временска серија са јединичним кореном**.
3. 3. Ако бар за једно решење важи $|g_k| > 1$, док су сва остала решења по модулу мањи од 1, тада је серија нестационарна и назива се **експлозивна временска серија**.

Аутоковаријациона и аутокорелациона функција AR(p) модела добијају се множењем израза 3.10 са X_{t-k} и рачунањем математичког очекивања:

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \phi_3 \gamma_3 + ... + \phi_p \gamma_p + \sigma^2 \tag{3.13}$$

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \phi_3 \gamma_{k-3} + ... + \phi_p \gamma_{k-p}, \quad \forall k > 0 \tag{3.14}$$

Дељењем једнакости 3.13 и 3.14 са γ_0 добијамо аутокорелационе коефицијенте:

$$\rho_0 = 1$$

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \phi_3 \rho_{k-3} + ... + \phi_p \rho_{k-p}, \quad \forall k > 0$$

3.3 Модели покретних просека (MA)

Следећу класу модела који за своју основу имају линеарни процес чине модели покретних просека. Ови модели директно се изводе из стандардне репрезентације линеарног процеса. Ако претпоставимо да је коначан број коефицијената датих у 3.1 различит од нула, односно да важи

$$\psi_1 = -\theta_1, \psi_2 = -\theta_2, ..., \psi_q = -\theta_q, \psi_k = 0, k > q \tag{3.15}$$

где је $q > 0$, и да је $E(X_t) = 0$, добијамо израз за модел покретних просека реда q:

$$X_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} - ... - \theta_q e_{t-q} \tag{3.16}$$

Видимо да је вредност серије у тренутку t линеарна комбинација чланова белог шума у тренуцима $t, t-1, ..., t-q$. Израз 3.16 другачије записујемо помоћу оператора L:

$$X_t = e_t - \theta_1 L e_t - \theta_2 L^2 e_t - \theta_3 L^3 e_t - ... - \theta_q L^q e_t \tag{3.17}$$

или краће:

$$X_t = (1 - \Theta(L))e_t \tag{3.18}$$

где је $\Theta(L)$ полином реда q, по оператору L.

MA(q) модел

Модел покретних просека реда q представља се једнакошћу 3.16. Аутоковаријациона функција рачуна се на следећи начин:

$$\begin{aligned}
\gamma_k &= cov(X_t, X_{t-k}) = E(X_t X_{t-k}) - E(X_t)E(X_{t-k}) = E(X_t X_{t-k}) \\
&= E((e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - ... - \theta_q e_{t-q}) \\
&\qquad (e_{t-k} - \theta_1 e_{t-k-1} - \theta_2 e_{t-k-2} - ... - \theta_q e_{t-k-q})) \\
&= \begin{cases} (1 + \theta_1^2 + \theta_2^2 + ... + \theta_q^2)\sigma^2, k = 0 \\ (-\theta_1 + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + ... + \theta_k\theta_{q-k})\sigma^2, 1 \leq k \leq q \\ 0, k \geq q \end{cases}
\end{aligned}$$

Аутокорелациона функција добија се коришћењем формуле 2.3:

$$\begin{aligned}\rho_k &= \frac{\gamma_k}{\gamma_0} \\ &= \begin{cases} 1, k = 0 \\ \frac{-\theta_1+\theta_1\theta_{k+1}+\theta_2\theta_{k+2}+...+\theta_k\theta_{q-k}}{1+\theta_1^2+\theta_2^2+...+\theta_q^2}, 1 \leq k \leq q \\ 0, k \geq q \end{cases}\end{aligned}$$

Дакле, аутокорелациони коефицијенти за доцње веће од реда модела једнаки су нули, што значи да су чланови серије на растојању већем од реда MA модела некорелисани. Ово се може видети и из једнакости 3.16, јер су чланови белог шума међусобно некорелисани. Моделом 3.16 описује се стационарна временска серија ако постоји оператор инверзан оператору $\Theta(L)$ у једнакости 3.18. Услов под којим овај оператор постоји назива се услов инвертибилности.

3.4 Ауторегресиони модели покретних просека (ARMA)

Следећи модел који се користи за описивање стационарних временских серије је комбинација AR и MA модела. То је **ауторегресиони модел покретних просека (ARMA)**. Овај модел дат је у форми:

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - ... - \phi_p X_{t-p} = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - ... - \theta_q e_{t-q} \quad (3.19)$$

при чему је $E(X_t) = 0$. Једнакост 3.19 може се записати коришћењем оператора L:

$$(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p) X_t = (1 - \theta_1 L - \theta_2 L^2 - ... - \theta_q L^q) e_t \quad (3.20)$$

односно краће:

$$\Phi(L) X_t = \Theta(L) e_t \quad (3.21)$$

при чему важи претпоставка да полимоми $\Phi(L)$ и $\Theta(L)$ немају заједничке факторе. Претходна репрезенрација односи се на ARMA модел реда (p, q). Ако једнакост 3.19 напишемо као:

$$\begin{aligned}X_t =& \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \\ & e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - ... - \theta_q e_{t-q}\end{aligned}$$

а затим добијени израз помножимо са X_{t-k} и израчунамо математичко очекивање, добићемо аутоковаријациону функцију:

$$\begin{aligned}\gamma_k =& \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + ... + \phi_p\gamma_{k-p} + \\ & E(e_tX_{t-k}) - \theta_1E(e_{t-1}X_{t-k}) - ... - \theta_qE(e_{t-q}X_{t-k})\end{aligned} \tag{3.22}$$

Како због некорелисаности чланова белог шума важи $(X_{t-k}e_{t-q}) = 0$ за $k > q$, добијамо аутоковаријациону функција за доцњу већу од q:

$$\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + ... + \phi_p\gamma_{k-p} \tag{3.23}$$

На основу израза 3.22 и 3.23 можемо закључити да аутоковаријациона функција на доцњама мањим од реда компоненте покретних просека q зависи и од AR и од MA коефицијената, док се на доцњама већим од q губи утицај MA коефицијената, те се ова функција понаша као у случају AR модела.
Показаћемо да се ARMA модели изводе из линеарног процеса. Како се произвољни полином бесконачног реда $\Psi(L)$ може записати као количник два полинома коначних редова, добијамо следеће:

$$\Psi(L) = \frac{\Theta(L)}{\Phi(L)}$$

где су $\Theta(L)$ и $\Phi(L)$ полиноми реда p i q, респективно. Узимајући у обзир репрезентацију линеарног процеса (2.2) и једнакост (2.34) важи:

$$\begin{aligned}&X_t = \Psi(L)e_t = \frac{\Theta(L)}{\Phi(L)}e_t \\ &\Phi(L)X_t = \Theta(L)e_t \\ &X_t - \phi_1X_{t-1} - \phi_2X_{t-2} - ... - \phi_pX_{t-p} = e_t - \theta_1e_{t-1} - \theta_2e_{t-2} - ... - \theta_qe_{t-q}\end{aligned}$$

што је форма ARMA модела.
Када је реч о стационарности ARMA модела, она је одређена условом стационарности AR компоненте, док је инвертибилност ARMA модела одређена условом инвертибилности MA компоненте.

3.5 Ауторегресиони интегрисани модели покретних просека (ARIMA)

У наставку вреди споменути надградњу ARMA класе модела, којом се осим стационарних могу моделовати и нестационарне временске серије. У

даљој анализи полазимо од тога да је временска серија са којом радимо припада класи диференцно стационарних процеса, што значи да ћемо, њеним диференциарањем довољан број пута, добити стационарну серију. Овакви процеси називају се још и интегрисани процеси. Уколико почетну серију диференцирамо d пута, а након тога нову серију моделујемо ARMA(p, q) моделом, добијамо ауторегресиони интегрисани модел покретних просека реда d (ARIMA(p,d,q) модел) који можемо представити изразом:

$$(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p)(1 - L)^d X_t = \theta_0 + (1 - \theta_1 L - \theta_2 L^2 - ... - \theta_q L^q)e_t \tag{3.24}$$

или краће:

$$\Phi(L)(1 - L)^d X_t = \theta_0 + \Theta(L)e_t \tag{3.25}$$

где је параметар d ниво интегрисаности серије, а θ_0 параметар уз линеарни тренд у серији. При томе претпостављамо да полиноми $\Phi(L)$ и $\Theta(L)$ немају заједничке факторе.

Напоменимо да су све претходно разматране AR, MA и ARMA спецификације, специјални случајеви ARIMA модела.

Глава 4

Моделовање структурног лома

У овом поглављу, дефинисаћемо појмове интервенције и структурног лома. Затим ћемо описати технике моделовања структурног лома[1], у контексту надоградње ARMA модела који су детаљно анализирани у претходном поглављу. Примена поменутих модела изискује претпоставку да процес који је генерисао серију не трпи промене. Циљ је изградити модел који ће укључити могућност промене у кретању серије. Поступак изградње траженог модела зависи од тога да ли нам је познат тренутак у коме је дошло до интервенције, стога ћемо анализирати обе ситуације.

4.1 Интервенција

Временске серије често су под утицајем посебних догађаја који су егзогене природе, што значи да настају под утицајем спољашњих фактора. Овакав догађај назива се **интервенција**. Приликом анализе изложене у овом поглављу претпостављаћемо да је у питању познати догађај. Појава интервенције посебно је карактеристична за економске временске серије[9]. Осим неких неочекиваних догађаја са претежно негативном конотацијом, као што су природне катастрофе, штрајкови, економске кризе и сл. може се радити и о планираним интервенцијама као што су промене економске политике, доношење нових законских прописа, покретање маркетиншке кампање итд. Неки примери интервенција су:

[1] С обзиром на једнозначну кореспонденцију између интервенције и структурног лома који она изазива, еквивалентно се може користити израз моделовање интервенције.

- Нафтни шок из 1973. када су чланице Организације земаља извозница нафте донеле одлуку о ембаргу на извоз нафте у одређене државе, што се одразило на кретање серије цена нафте[10, p. 295].

- Санкције уведене СР Југославији 1992. од стране међународне заједнице утицале су на временску серију индекса индустијске проиводње[10, p. 296].

- Доношење новог закона којим је смањен дозвољени удео високо реактивних угљоводоника у гориву огледало се у промени кретања просечне концентрације кисеоника у центру Лос Анђелеса на месечном нивоу у периоду од 1955. до 1972. године[4].

- Одлука председника Сједињених Држава да покрене војне операције на Блиском истоку, као одговор на терористички напад од 11. септембра 2001. године драстично је утицала на његов рејтинг међу гласачима[1].

Последица дејства интервенције на кретање временске серије је појава нестандардних опсервација. То су опсервације које нису у складу са кретањем временске серије до тренутка у коме се десила интервенција. Овакав скуп опсервација назива се **структурни лом**.
Све инетрвенције се могу поделити у две групе на основу трајања њиховог дејства на кретање временске серије. Прву групу чине једнократне интервенције (енгл. pulse intervention). Њихов ефекат огледа се у појави једне или неколико нестандардних опсервација. Пример такве интервенције је наведена популарност председника САД, која је након одлуке о војној интервенцији на Блиском истоку нагло скочила, да би у наредном периоду полако опадала. Друга група су трајне интервенције (енгл. step intervention), чији је ефекат трајна промена кретања серије. Пример интервенције овог типа је промена законске регулативе којом је трајно промењена концентрација кисеоника у ваздуху у Лос Анђелесу.

4.2 Моделовање структурног лома - познат тренутак интервенције

У наставку ћемо, на једном једноставном моделу, приказати како се моделује структурни лом у случају када је познат тренутак интервенције. Поступак започињемо изградњом ARMA модела користећи опсервације серије које претходе периоду интервенције. Наведени поступак заснивамо на Бокс-Ценкинсовој стратегији моделовања[9, p. 189]. Након што дођемо до адекватног модела, формирамо модел интервенције, како бисмо на основу њега могли анализирати њене ефекте.

Нека је са X_t означена вредност серије моделиране на основу опсервација пре појаве интервенције. Модификација овог модела која обухвата постојање структурног лома у општем облику се представља на следећи начин[15]:

$$Y_t = f(t) + X_t \tag{4.1}$$

при чему је $f(t)$ параметарска функција времена која представља спољашњи утицај који трпи серија X_t. Конкретан облик ове функције у пракси се одређује на основу претпоставки истраживача о природи лома изазваног интервенцијом. У контексту ARMA модела, једнакост 4.1 записујемо као[4]:

$$X_t = \nu(L)I_t + \frac{\Theta(L)}{\Phi(L)}e_t \tag{4.2}$$

У једнакости 4.2 сабирак $\nu(L) = \frac{\omega(L)}{\delta(L)}$ означава полином по оператору доцње L, којим се моделује присуство лома, односно утицај интервенције на кретање серије током времена. Овај полином може се представити као:

$$\nu(L) = \frac{\omega(L)}{\delta(L)}L^k \tag{4.3}$$

где полином $\omega(L) = \omega_0 - \omega_1 L - ... - \omega_s L^s$ описује очекивани почетни ефекат интервенције, а полином $\delta(L) = \delta_0 - \delta_1 L - ... - \delta_r L^r$ описује трајни ефекат интервенције. Степен k, оператора L означава период времена између интервенције и појаве ефеката њеног дејства.

Променљива I_t из једнакости 4.2 је индикатор присуства структурног лома који се придружује полиному $\nu(L)$. Може се дефинисати на два начина у зависности од типа структурног лома. Нека је са T означен тренутак у коме

је дошло до интервенције. Првим начином се описује једнократни структурни лом:

$$I_t = IM_t = \begin{cases} 1, t = T \\ 0, t \neq T \end{cases} \tag{4.4}$$

Променљива IM_t назива се импулсна променљива (енгл. *pulse variable*). Другим начином се описује структурни лом који је трајног карактера:

$$I_t = S_t = \begin{cases} 1, t \geq T \\ 0, t < T \end{cases} \tag{4.5}$$

Променљива S_t зове се степеник променљива (енгл. *step variable*). Могућа је модификација претходне променљиве у ситуацији када се интервенција јавља у тренутку T_1, и траје до тренутка T_2:

$$I_t = S_t = \begin{cases} 1, T_1 \leq t \leq T_2 \\ 0, t < T_1, t > T_2 \end{cases} \tag{4.6}$$

Из табеле 4.1 [9, p. 221] можемо уочити везу између променљивих IM_t и S_t. Из последње колоне видимо да је разлика променљивих S_t и S_{t-1} једнака

Табела 4.1: Веза променљивих IM_t и S_t

t	1	2	...	T-1	T	T+1	T+2	...
S_t	0	0	...	0	1	1	1	...
S_{t-1}	0	0	...	0	0	1	1	...
$S_t - S_{t-1}$	0	0	...	0	1	0	0	...

1 у тренутку T, док је у свим осталим тренуцима једнака 0, што је управо дефиниција импулсне променљиве 4.4. Одатле имамо да важи $S_t - S_{t-1} = IM_t$, што значи да трајном структурном лому у нивоу временске серије, одговара једнократни лом у нивоу прве диференце те серије.

У наставку ћемо дати приказ неколико основних конкретизација модела 4.2, у зависности од избора параметара, односно коефицијената полинома $\omega(L)$ и $\delta(L)$, и експонента k. При том ћемо са X_t означавати вредности серије из периода пре тренутка интервенције T, док ћемо серију након интервенције обележавати са Y_t:

1. $\omega(L) = \omega_0$, $\delta(L) = 1$, $k = 1$, $I_t = IM_t$:

$$Y_t = \omega_0 LIM_t + X_t \tag{4.7}$$

За t = T + 1, из једнакости 4.7 имамо:

$$\begin{aligned} Y_{T+1} &= \omega_0 LIM_{T+1} + X_{T+1} \\ Y_{T+1} &= \omega_0 IM_T + X_{T+1} \\ Y_{T+1} &= \omega_0 + X_{T+1} \end{aligned} \tag{4.8}$$

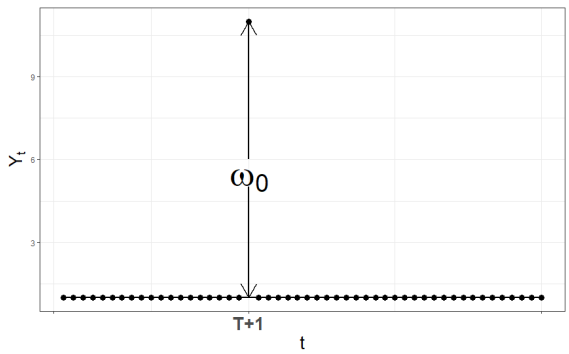
За t = T + 2, из једнакости 4.7 имамо:

$$\begin{aligned} Y_{T+2} &= \omega_0 LIM_{T+2} + X_{T+2} \\ Y_{T+2} &= \omega_0 IM_{T+1} + X_{T+2} \\ Y_{T+2} &= X_{T+2} \end{aligned}$$

Аналогно извођење важи за свако $t = T + j$, где је $j \geq 2$. На основу те чињенице и једнакости 4.8 можемо закључити да за $t = T + j$, $j \geq 1$ важи:

$$Y_{T+j} = \begin{cases} X_{T+j} + \omega_0, j = 1 \\ X_{T+j}, j > 1 \end{cases} \tag{4.9}$$

На слици 4.1 је дат графички приказ структурног лома описаног моделом 4.7:



Слика 4.1: Графички приказ модела 4.7

Видимо да се вредност серије мења за вредност коефицијента ω_0 у тренутку $T + 1$, што је у складу са задатом вредношћу степена оператора

доцње ($k = 1$), која описује ситуацију у којој се структурни лом појављује један тренутак након дејства интервенције. У свим осталим временским тачкама, вредност серије је непромењена, па имамо само једну нестандардну опсервацију.

2. $\omega(L) = \omega_0$, $\delta(L) = 1$, $k = 1$, $I_t = S_t$:

$$Y_t = \omega_0 L S_t + X_t \tag{4.10}$$

За t = T + 1, из једнакости 4.10 имамо:

$$\begin{aligned} Y_{T+1} &= \omega_0 L S_{T+1} + X_{T+1} \\ Y_{T+1} &= \omega_0 S_T + X_{T+1} \\ Y_{T+1} &= \omega_0 + X_{T+1} \end{aligned} \tag{4.11}$$

За t = T + 2, из једнакости 4.10 имамо:

$$\begin{aligned} Y_{T+2} &= \omega_0 L S_{T+2} + X_{T+2} \\ Y_{T+2} &= \omega_0 S_{T+1} + X_{T+2} \\ Y_{T+2} &= \omega_0 + X_{T+2} \end{aligned}$$

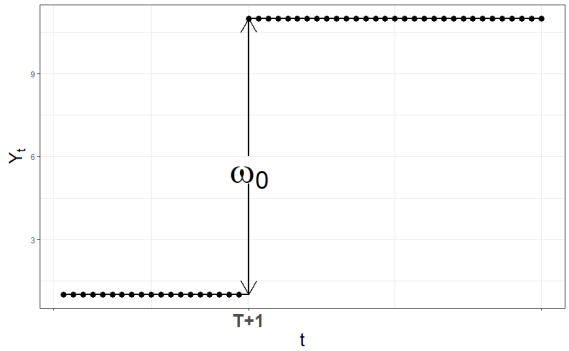
Аналогним извођењем за свако $t = T + j$, где је $j \geq 1$, долазимо до закључка да је вредност серије један тренутак након појаве интервенције увећана за коефицијент ω_0, и да је ова промена трајно присутна од тог тренутка, односно да важи:

$$Y_{T+j} = X_{T+j} + \omega_0, j \geq 1 \tag{4.12}$$

На слици 4.2 је дат графички приказ структурног лома описаног моделом 4.10. На приказаном графику можемо видети да се ниво серије мења за коефицијент ω_0 у тренутку $T + 1$, и та промена је трајног карактера.

3. $\omega(L) = \omega_0$, $\delta(L) = 1 - \delta_1 L$, $k = 1$, $I_t = IM_t$:

$$Y_t = \frac{\omega_0}{1 - \delta_1 L} L IM_t + X_t \tag{4.13}$$



Слика 4.2: Графички приказ модела 4.10

За t = T + 1, из једнакости 4.13 имамо:

$$
\begin{aligned}
Y_{T+1} &= \frac{\omega_0}{1-\delta_1 L} LIM_{T+1} + X_{T+1} \\
Y_{T+1} - \delta_1 L Y_{T+1} &= \omega_0 LIM_{T+1} + X_{T+1} - \delta_1 L X_{T+1} \\
Y_{T+1} - \delta_1 Y_T &= \omega_0 IM_T + X_{T+1} - \delta_1 X_T \\
Y_{T+1} &= \omega_0 + X_{T+1} + \delta_1 (Y_T - X_T)
\end{aligned}
$$

Из једнакости 4.13 следи да је $Y_T - X_T = \frac{\omega_0}{1-\delta_1 L} IM_{T-1} = 0$, па имамо:

$$
Y_{T+1} = X_{T+1} + \omega_0 \tag{4.14}
$$

За t = T + 2, из једнакости 4.13 имамо:

$$
\begin{aligned}
Y_{T+2} &= \frac{\omega_0}{1-\delta_1 L} LIM_{T+2} + X_{T+2} \\
Y_{T+2} - \delta_1 L Y_{T+2} &= \omega_0 LIM_{T+2} + X_{T+2} - \delta_1 L X_{T+2} \\
Y_{T+2} - \delta_1 Y_{T+1} &= \omega_0 IM_{T+1} + X_{T+2} - \delta_1 X_{T+1} \\
Y_{T+2} &= X_{T+2} + \delta_1 (Y_{T+1} - X_{T+1})
\end{aligned}
$$

Из једнакости 4.14 следи да је $Y_{T+1} - X_{T+1} = \omega_0$, па имамо:

$$
Y_{T+2} = X_{T+2} + \delta_1 \omega_0 \tag{4.15}
$$

За t = T + 3, из једнакости 4.13 имамо:

$$
\begin{aligned}
Y_{T+3} &= \frac{\omega_0}{1-\delta_1 L} LIM_{T+3} + X_{T+3} \\
Y_{T+3} - \delta_1 L Y_{T+3} &= \omega_0 LIM_{T+3} + X_{T+3} - \delta_1 L X_{T+3} \\
Y_{T+3} - \delta_1 Y_{T+3} &= \omega_0 IM_{T+2} + X_{T+3} - \delta_1 X_{T+2} \\
Y_{T+3} &= X_{T+3} + \delta_1 (Y_{T+2} - X_{T+2})
\end{aligned}
$$

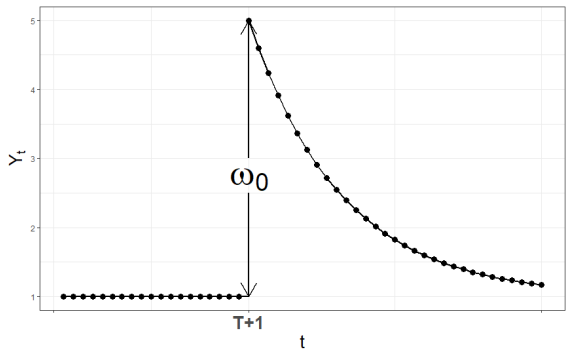
Из једнакости 4.15 следи да је $Y_{T+2} - X_{T+2} = \delta_1\omega_0$, па имамо:

$$Y_{T+3} = X_{T+3} + \delta_1^2\omega_0 \tag{4.16}$$

На основу једнакости 4.14, 4.15 и 4.16, можемо закључити да за $t = T+j$, $j \geq 1$ важи:

$$Y_{T+j} = X_{T+j} + \delta_1^{j-1}\omega_0 \tag{4.17}$$

На слици 4.3 је дат графички приказ структурног лома описаног моделом 4.13 у случају када је $0 < \delta_1 < 1$. Приметимо да се у случају $\delta_1 = 0$ овај модел своди на случај 4.9, а ако је $\delta_1 = 1$ ради се о моделу 4.17.



Слика 4.3: Графички приказ модела 4.13

4. $\omega(L) = \omega_0$, $\delta(L) = 1 - \delta_1 L$, $k = 1$, $I_t = S_t$:

$$Y_t = \frac{\omega_0}{1-\delta_1 L} L S_t + X_t \tag{4.18}$$

За t = T + 1, из једнакости 4.18 имамо:

$$\begin{aligned}
&Y_{T+1} = \frac{\omega_0}{1-\delta_1 L} L S_{T+1} + X_{T+1} \\
&Y_{T+1} - \delta_1 L Y_{T+1} = \omega_0 L S_{T+1} + X_{T+1} - \delta_1 L X_{T+1} \\
&Y_{T+1} - \delta_1 Y_T = \omega_0 S_T + X_{T+1} - \delta_1 X_T \\
&Y_{T+1} = \omega_0 + X_{T+1} + \delta_1(Y_T - X_T)
\end{aligned}$$

Из једнакости 4.18 следи да је $Y_T - X_T = \frac{\omega_0}{1-\delta_1 L} S_{T-1} = 0$, па имамо:

$$Y_{T+1} = X_{T+1} + \omega_0 \tag{4.19}$$

За t = T + 2, из једнакости 4.18 имамо:

$$
\begin{aligned}
Y_{T+2} &= \frac{\omega_0}{1-\delta_1 L} L S_{T+2} + X_{T+2} \\
Y_{T+2} - \delta_1 L Y_{T+2} &= \omega_0 L S_{T+2} + X_{T+2} - \delta_1 L X_{T+2} \\
Y_{T+2} - \delta_1 Y_{T+1} &= \omega_0 S_{T+1} + X_{T+2} - \delta_1 X_{T+1} \\
Y_{T+2} &= X_{T+2} + \omega_0 + \delta_1 (Y_{T+1} - X_{T+1})
\end{aligned}
$$

Из једнакости 4.19 следи да је $Y_{T+1} - X_{T+1} = \omega_0$, па имамо:

$$Y_{T+2} = X_{T+2} + \omega_0(1+\delta_1) \tag{4.20}$$

За t = T + 3, из једнакости 4.18 имамо:

$$
\begin{aligned}
Y_{T+3} &= \frac{\omega_0}{1-\delta_1 L} L S_{T+3} + X_{T+3} \\
Y_{T+3} - \delta_1 L Y_{T+3} &= \omega_0 L S_{T+3} + X_{T+3} - \delta_1 L X_{T+3} \\
Y_{T+3} - \delta_1 Y_{T+3} &= \omega_0 S_{T+2} + X_{T+3} - \delta_1 X_{T+2} \\
Y_{T+3} &= X_{T+3} + \omega_0 + \delta_1 (Y_{T+2} - X_{T+2})
\end{aligned}
$$

Из једнакости 4.20 следи да је $Y_{T+2} - X_{T+2} = \omega_0(1+\delta_1)$, па имамо:

$$Y_{T+3} = X_{T+3} + \omega_0 + \delta_1\omega_0(1+\delta_1) = X_{T+3} + \omega_0(1+\delta_1+\delta_1^2) \tag{4.21}$$

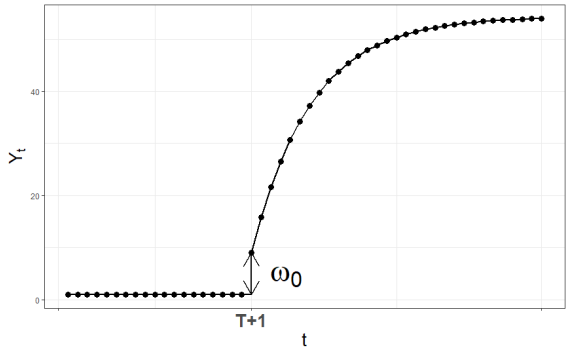
На основу једнакости 4.19, 4.20 и 4.21, можемо закључити да за $t = T+j$, $j \geq 1$ важи:

$$Y_{T+j} = X_{T+j} + \omega_0(1+\delta_1+\delta_1^2+...+\delta_1^{j-1}) \tag{4.22}$$

На слици 4.4 је дат графички приказ структурног лома описаног моделом 4.18 у случајевима $0 < \delta_1 < 1$ и $\delta_1 = 1$. У случају $\delta_1 = 0$, ради се о моделу 4.17.

5. Могуће је користити и сложеније форме функције $\nu(L)$, као на пример $\nu(L) = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L})$, $I_t = IM_t$:

$$Y_t = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L}) IM_t + X_t \tag{4.23}$$



Слика 4.4: Графички приказ модела 4.18

За t = T + 1, из једнакости 4.23 имамо:

$$
\begin{aligned}
&Y_{T+1} = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L})IM_{T+1} + X_{T+1} \\
&Y_{T+1} = \frac{\omega_1 L - \omega_1 L^2 + \omega_2 L - \omega_2 \delta_1 L^2}{1-(1+\delta_1)L + \delta_1 L^2} IM_{T+1} + X_{T+1} \\
&Y_{T+1} - X_{T+1} - (1+\delta_1)(Y_T - X_T) + \delta_1(X_{T-1} - Y_{T-1}) \\
&= \omega_1 IM_T - \omega_1 IM_{T-1} + \omega_2 IM_T - \omega_2 \delta_1 IM_{T-1} \\
&Y_{T+1} - X_{T+1} - (1+\delta_1)(Y_T - X_T) + \delta_1(X_{T-1} - Y_{T-1}) \\
&= \omega_1 + \omega_2
\end{aligned}
$$

На основу једнакости 4.23 важи да је $Y_T - X_T = Y_{T-1} - X_{T-1} = 0$, па следи:

$$
Y_{T+1} = \omega_1 + \omega_2 + X_{T+1} \tag{4.24}
$$

За t = T + 2, из једнакости 4.23 имамо:

$$
\begin{aligned}
&Y_{T+2} = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L})IM_{T+2} + X_{T+2} \\
&Y_{T+2} = \frac{\omega_1 L - \omega_1 L^2 + \omega_2 L - \omega_2 \delta_1 L^2}{1-(1+\delta_1)L + \delta_1 L^2} IM_{T+2} + X_{T+2} \\
&Y_{T+2} - X_{T+2} - (1+\delta_1)(Y_{T+1} - X_{T+1}) + \delta_1(X_T - Y_T) \\
&= \omega_1 IM_{T+1} - \omega_1 IM_T + \omega_2 IM_{T+1} - \omega_2 \delta_1 IM_T \\
&Y_{T+2} - X_{T+2} - (1+\delta_1)(Y_{T+1} - X_{T+1}) + \delta_1(X_T - Y_T) \\
&= -\omega_1 - \omega_2 \delta_1
\end{aligned}
$$

Из једнакости 4.23 следи да је $Y_T - X_T = 0$, док из 4.24 следи $Y_{T+1} - X_{T+1} = \omega_1 + \omega_2$ па следи:

$$\begin{aligned}
&Y_{T+2} - X_{T+2} = (1+\delta_1)(\omega_1+\omega_2) - \omega_1 - \omega_2\delta_1 \\
&Y_{T+2} - X_{T+2} = \omega_1 + \omega_2 + \delta_1\omega_1 + \delta_1\omega_2 - \omega_1 - \omega_2\delta_1 \\
&Y_{T+2} = \delta_1\omega_1 + \omega_2 + X_{T+2}
\end{aligned} \tag{4.25}$$

За t = T + 3, на основу једнакости 4.23 важи:

$$\begin{aligned}
&Y_{T+3} = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L})IM_{T+3} + X_{T+3} \\
&Y_{T+3} = \frac{\omega_1 L - \omega_1 L^2 + \omega_2 L - \omega_2\delta_1 L^2}{1-(1+\delta_1)L+\delta_1 L^2} IM_{T+3} + X_{T+3} \\
&Y_{T+3} - X_{T+3} - (1+\delta_1)(Y_{T+2} - X_{T+2}) + \delta_1(X_{T+1} - Y_{T+1}) \\
&= \omega_1 IM_{T+2} - \omega_1 IM_{T+1} + \omega_2 IM_{T+2} - \omega_2\delta_1 IM_{T+1} \\
&Y_{T+3} - X_{T+3} = (1+\delta_1)(Y_{T+2} - X_{T+2}) + \delta_1(X_{T+1} - Y_{T+1})
\end{aligned}$$

Из једнакости 4.23 следи да је $Y_T - X_T = 0$, из 4.24 следи $Y_{T+1} - X_{T+1} = \omega_1 + \omega_2$, а из 4.25 следи $Y_{T+2} - X_{T+2} = \delta_1\omega_1 + \omega_2$, па имамо:

$$\begin{aligned}
&Y_{T+3} - X_{T+3} = (1+\delta_1)(\delta_1\omega_1 + \omega_2) - \delta_1(\omega_1+\omega_2) \\
&Y_{T+3} - X_{T+3} = \delta_1\omega_1 + \omega_2 + \delta_1^2\omega_1 + \delta_1\omega_2 - \delta_1\omega_1 - \delta_1\omega_2 \\
&Y_{T+3} = \delta_1^2\omega_1 + \omega_2 + X_{T+3}
\end{aligned} \tag{4.26}$$

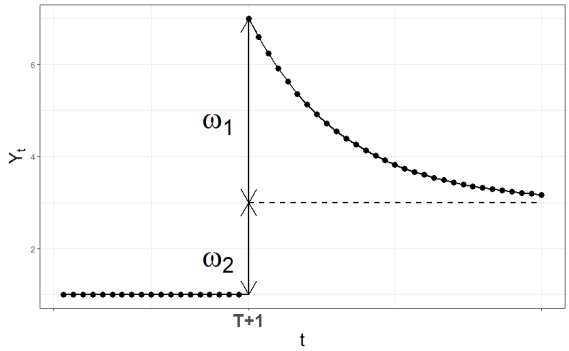
Из једнакости 4.24, 4.25 и 4.26, следи да за $t = T + j$, $j \geq 1$ важи:

$$Y_{T+j} = X_{T+j} + \delta_1^{j-1}\omega_1 + \omega_2 \tag{4.27}$$

На слици 4.5 је дат графички приказ структурног лома описаног моделом 4.23 у случају $0 < \delta_1 < 1$.

6. Ако у претходном примеру, уместо импулсне, искористимо степеник вештачку променљиву, добијамо $\nu(L) = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L})$, $I_t = S_t$:

$$Y_t = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L})S_t + X_t \tag{4.28}$$



Слика 4.5: Графички приказ модела 4.23

За t = T + 1, из једнакости 4.28 имамо:

$$Y_{T+1} = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L})S_{T+1} + X_{T+1}$$
$$Y_{T+1} = \frac{\omega_1 L - \omega_1 L^2 + \omega_2 L - \omega_2 \delta_1 L^2}{1-(1+\delta_1)L + \delta_1 L^2} S_{T+1} + X_{T+1}$$
$$Y_{T+1} - X_{T+1} - (1+\delta_1)(Y_T - X_T) + \delta_1(X_{T-1} - Y_{T-1})$$
$$= \omega_1 S_T - \omega_1 S_{T-1} + \omega_2 S_T - \omega_2 \delta_1 S_{T-1}$$
$$Y_{T+1} - X_{T+1} - (1+\delta_1)(Y_T - X_T) + \delta_1(X_{T-1} - Y_{T-1})$$
$$= \omega_1 + \omega_2$$

Из 4.28 следи да је $Y_T - X_T = Y_{T-1} - X_{T-1} = 0$, па закључујемо да важи:

$$Y_{T+1} = \omega_1 + \omega_2 + X_{T+1} \tag{4.29}$$

За t = T + 2, на основу једнакости 4.28 важи:

$$Y_{T+2} = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L})S_{T+2} + X_{T+2}$$
$$Y_{T+2} = \frac{\omega_1 L - \omega_1 L^2 + \omega_2 L - \omega_2 \delta_1 L^2}{1-(1+\delta_1)L + \delta_1 L^2} S_{T+2} + X_{T+2}$$
$$Y_{T+2} - X_{T+2} - (1+\delta_1)(Y_{T+1} - X_{T+1}) + \delta_1(X_T - Y_T)$$
$$= \omega_1 S_{T+1} - \omega_1 S_T + \omega_2 S_{T+1} - \omega_2 \delta_1 S_T$$
$$Y_{T+2} - X_{T+2} - (1+\delta_1)(Y_{T+1} - X_{T+1}) + \delta_1(X_T - Y_T)$$
$$= \omega_1 - \omega_1 + \omega_2 - \omega_2 \delta_1$$

Из 4.29 следи да је $Y_{T+1} - X_{T+1} = \omega_1 + \omega_2$, те важи:

$$Y_{T+2} - X_{T+2} = \omega_2 - \omega_2 \delta_1 + (1+\delta_1)(\omega 1 + \omega_2)$$
$$Y_{T+2} = X_{T+2} + (1+\delta_1)\omega_1 + 2\omega_2 \tag{4.30}$$

За t = T + 3, из једнакости 4.28 добијамо:

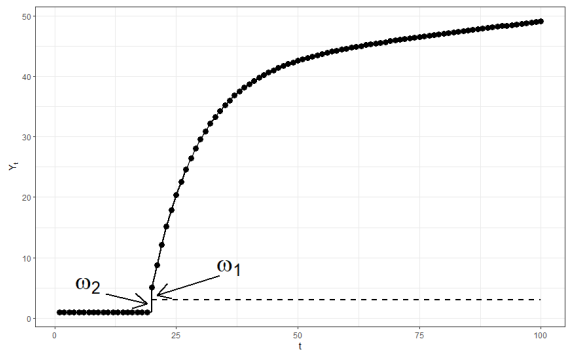
$$
\begin{aligned}
&Y_{T+3} = (\frac{\omega_1 L}{1-\delta_1 L} + \frac{\omega_2 L}{1-L})S_{T+3} + X_{T+3} \\
&Y_{T+3} = \frac{\omega_1 L - \omega_1 L^2 + \omega_2 L - \omega_2 \delta_1 L^2}{1-(1+\delta_1)L + \delta_1 L^2} S_{T+3} + X_{T+3} \\
&Y_{T+3} - X_{T+3} - (1+\delta_1)(Y_{T+2} - X_{T+2}) + \delta_1(X_{T+1} - Y_{T+1}) \\
&= \omega_1 S_{T+2} - \omega_1 S_{T+1} + \omega_2 S_{T+2} - \omega_2 \delta_1 S_{T+1} \\
&Y_{T+3} - X_{T+3} - (1+\delta_1)(Y_{T+2} - X_{T+2}) + \delta_1(X_{T+1} - Y_{T+1}) \\
&= \omega_1 - \omega_1 + \omega_2 - \omega_2 \delta_1
\end{aligned}
$$

На основу једнакости 4.30 важи да је $Y_{T+2} - X_{T+2} = (1+\delta_1)\omega_1 + 2\omega_2$, те имамо:

$$
\begin{aligned}
&Y_{T+3} - X_{T+3} = \omega_2 - \omega_2\delta_1 + (1+\delta_1)((1+\delta_1)\omega_1 + 2\omega_2) - \delta_1(\omega_1 + \omega_2) \\
&Y_{T+2} = X_{T+2} + (1+\delta_1+\delta_1^2)\omega_1 + 3\omega_2 \qquad (4.31)
\end{aligned}
$$

Сада, на основу једнакости 4.29, 4.30 и 4.31, можемо закључити да важи:

$$Y_{T+j} = X_{T+j} + (1+\delta_1+...+\delta_1^{j-1})\omega_1 + j\omega_2 \qquad (4.32)$$



Слика 4.6: Графички приказ модела 4.28

На слици 4.6 је дат графички приказ структурног лома описаног моделом 4.28 у случају $0 < \delta_1 < 1$.

4.3 Моделовање структурног лома - непознат тренутак интервенције

У овом одељку дискутоваћемо о моделовању структурног лома у ситуацији када немамо информацију о тренутку у ком се десила интервенција која га је изазвала. Овакви типови интервенција најчешће подразумевају неке неочекиване догађаје, али нестандардне опсервације које чине структурни лом могу бити последица грешака у мерењу или бележењу података.

Адитивни и иновативни модели

Размотрићемо два модела који се могу користити у овом случају. То су **адитивни** и **иновативни** модел[8].
Адитивни модел нестандардне опсервације (additive outlier) дефинишемо на следећи начин:

$$Y_t = X_t + \omega I_t \tag{4.33}$$

где је са X_t означена серија из које смо искључили нестандардне опсервације, а са Y_t посматрана серија у чијем кретању је присутан структурни лом, односно нестандардне опсервације. Променљива I_t је импулсна вештачка променљива уведена у 4.4. Коефицијент ω представља квантификацију адитивног утицаја структурног лома на ниво серије. Ако претпоставимо да се серија X_t може описати ARMA(p, q) моделом једнакост 4.33 можемо записати као:

$$Y_t = \frac{\Theta(L)}{\Phi(L)} e_t + \omega I_t \tag{4.34}$$

На основу једнакости 4.33 можемо закључити да је у адитивном моделу серија захваћена утицајем нестандардне опсервације само у тренутку у коме се та опсервација десила.
Иновативни модел нестандардне опсервације (inovative outlier) дефинишемо следећом једнакошћу:

$$Y_t = X_t + \frac{\Theta(L)}{\Phi(L)} \omega I_t = \frac{\Theta(L)}{\Phi(L)} (e_t + \omega I_t) \tag{4.35}$$

при чему је X_t серија ослобођена нестандардних опсервација, за коју претпостављамо да се може моделовати ARMA(p, q) моделом, I_t индикаторска

променљива као у случају адитивног модела, а ω квантификатор иновативног утицаја нестандардних опсервација на кретање серије. Ако посматрамо серију у тренутку појаве структурног лома T, из једнакости 4.35 можемо закључити да је ефекат случајног шока описаног чланом процеса белог шума e_T појачан за коефицијент ω:

$$Y_T = \frac{\Theta(L)}{\Phi(L)}(e_T + \omega) \tag{4.36}$$

За разлику од адитивног модела, овде се наведени ефекат преноси кроз наредне опсервације захваљујући меморији ARMA(p, q) процеса. Ако количник $\frac{\Theta(L)}{\Phi(L)}$ запишемо као полином $\Psi(L)$ (репрезентација линеарног процеса), за t = T+i, i = 0, 1, ... важи:

$$Y_{T+i} = (\psi_0 + \psi_1 L + \psi_2 L^2 + ...)(e_{T+i} + \omega I_{T+i}) = X_{T+i} + \omega\psi_i \tag{4.37}$$

јер из својстава импулсне вештачке променљиве следи да је $\psi_j L^j I_{T+i} = \psi_j$, за $j = i$, односно 0, за $j \neq i$. Приметимо да у случају да је X_t стационарна серија, следи да бесконачна сума коефицијената линеарног процеса конвергира, те сами коефицијенти конвергирају ка нули. Стога, на основу 4.37 можемо закључити да утицај нестандардних опсервација опада ка нули.

Треба напоменути да серија коју анализирамо може бити под утицајем структурних ломова различитог типа, у различитим временским тренуцима. На основу тога, ако се структурни ломови јављају у тренуцима $T_1, T_2, ..., T_k$ можемо разматрати модел облика:

$$Y_t = \sum_{j=1}^{k} \omega_j \nu_j(L) I_t^{(T_j)} + X_t \tag{4.38}$$

У претходној једнакости ω_j означава коефицијент утицаја нестандардне опсервације у тренутку T_j. Функција $\nu_j(L)$ узима вредност 1 у случају да користимо адитивни модел, док код иновативног модела важи $\nu_j(L) = \frac{\Theta(L)}{\Phi(L)}$. Са $I_t^{(T_j)}$ означена је индикаторска променљива која узима вредност 1 за $t = T_j$, а нулу иначе.

Претходно изложени модели почивају на претпоставци да можемо доћи до одговарајућег модела серије X_t, која представља почетну серију коју проучавамо, ослобођену нестандардних опсервација. Међутим, услед недостатка информације о постојању и тренуцима појаве нестандардних опсервација нисмо у могућности да оценимо коефицијенте модела за серију X_t. Занемарујући

постојање нестандардних опсервација приликом изградње модела, добићемо пристрасне оцене коефицијената[10, p. 299]. У циљу превазилажења наведеног проблема приказаћемо итеративну процедуру [5, pp. 470–472], којом утврђујемо присуство нестандардних опсервација, као и њихов тип, на основу чега се одлучујемо за примену неког од описаних модела. Основни кораци су следећи:

1. Изграђујемо ARMA модел на основу посматране серије Y_t, без икаквих претпоставки о постојању нестандардних опсервација. Након тога оцењујемо резидуале добијеног модела са $\hat{\epsilon_t} = \hat{\theta}^{-1}(L)\hat{\phi}(L)X_t = \hat{\pi}(L)X_t$, и њихову дисперзију $\hat{\sigma_a}^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{\epsilon_t}^2$.

2. За свако $t = 1, 2, ..., n$, дефинишемо тест статистике за детектовање нестандардних опсервација. Разликујемо тест статистике за проналажење адитивних, односно иновативних нестандардних опсервација.
   У наставку излажемо мотивацију за долажење до адекватних тест статистика, користећи претпоставку да нам је познат тренутак појаве нестандардне опсервације, што у пракси није случај. Полазимо од израза 4.34 и 4.35, и дефинишемо $\pi(L) = \theta^{-1}(L)\phi(L) = 1 - \sum_{i=1}^{\infty}\pi_i L^i$ и $\epsilon_t = \pi(L)Y_t$, за t = 1, 2, ..., n. Тада једнакост 4.34 можемо записати као:

$$\epsilon_t = \omega\pi(L)I_t + e_t = \omega x_t + e_t$$

   где је $x_t = \pi(L)I_t$. Приметимо да за $t = T + i$, $i = 1, 2, 3, ...$ важи $x_t = (1 - \sum_{i=1}^{\infty}\pi_i L^i)I_t = -\pi_i$, док је за $t = T$, $x_T = 1$, па можемо писати $x_t = -\pi_i$, $t = T + i, i = 0, 1, 2, ...$ при чему је $\pi_0 = -1$. На основу методе најмањих квадрата добијамо оцену $\hat{\omega}_{A,T}$ за параметар ω:

$$S = \sum_{t=1}^{n} e_t^2 = \sum_{t=1}^{n}(\epsilon_t - \omega x_t)^2$$

$$\frac{\partial S}{\partial \omega} = 2\sum_{t=1}^{n}(\epsilon_t - \omega x_t)(-x_t) = 0$$

$$\sum_{t=1}^{n}\epsilon_t x_t - \omega\sum_{t=1}^{n}x_t^2 = 0$$

$$\hat{\omega}_{A,T} = \frac{\sum_{t=1}^{n}\epsilon_t x_t}{\sum_{t=1}^{n}x_t^2} = \frac{\sum_{t=T}^{n}\epsilon_t x_t}{\sum_{t=T}^{n}x_t^2}$$

$$\hat{\omega}_{A,T} = \frac{\epsilon_T - \sum_{t=T+1}^{n}\pi_{t-T}\epsilon_t}{\sum_{t=T}^{n}\pi_{t-T}^2}$$

Изводимо варијансу добијене оцене:

$$\begin{aligned} var(\hat{\omega}_{A,T}) &= var\left(\frac{\sum_{t=T}^{n}\epsilon_t x_t}{\sum_{t=T}^{n} x_t^2}\right) \\ &= (\sum_{t=T}^{n} x_t^2)^{-2}\sum_{t=T}^{n} x_t^2 var(\epsilon_t) \\ &= (\sum_{t=T}^{n} \pi_{t-T}^2)^{-2}\sum_{t=T}^{n} \pi_{t-T}^2 var(\epsilon_t) \\ &= \frac{\sigma_e^2}{\tau^2} \end{aligned}$$

где је $\tau^2 = \sum_{t=T}^{n} \pi_{t-T}^2$.

Тестирамо нулту хипотезу $H_0 : \omega = 0$, против алтернативе $H_1 : \omega \neq 0$. У том циљу формирамо тест статистику:

$$\lambda_{A,T} = \frac{\tau\hat{\omega}_{A,T}}{\sigma_e} \tag{4.39}$$

У случају иновативног модела, полазећи од једнакости 4.35 добијамо:

$$\epsilon_t = \omega I_t + e_t$$

Применом методе најмањих квадрата добијамо оцену $\hat{\omega}_{I,T}$ параметра ω:

$$\begin{aligned} S &= \sum_{t=1}^{n} e_t^2 = \sum_{t=1}^{n}(\epsilon_t - \omega I_t)^2 \\ \frac{\partial S}{\partial \omega} &= 2\sum_{t=1}^{n}(\epsilon_t - \omega I_t)(-I_t) = 0 \\ \hat{\omega}_{I,T} &= \epsilon_T \end{aligned}$$

Варијанса добијене оцене је $var(\hat{\omega}_{I,T}) = \sigma_e^2$. И у овом случају нулта хипотеза гласи $H_0 : \omega = 0$, а алтернативна $H_1 : \omega \neq 0$. Формирамо следећу тест статистику:

$$\lambda_{I,T} = \frac{\hat{\omega}_{I,T}}{\sigma_e} \tag{4.40}$$

При услову да важи нулта хипотеза статистике 4.39 и 4.40 имају нормалну расподелу[5, p. 471].

На основу претходног излагања, оцене наведених тест статистика у произвољном тренутку t дефинишемо на следећи начин:

$$\hat{\lambda}_{A,t} = \frac{\hat{\tau}\hat{\omega}_{A,t}}{\hat{\sigma}_e} \qquad \hat{\lambda}_{I,t} = \frac{\hat{\omega}_{I,t}}{\hat{\sigma}_e} \tag{4.41}$$

и на основу њих дефинишемо статистику:

$$\hat{\lambda}_T = \max_t(\max((|\hat{\lambda}_{A,t}|, |\hat{\lambda}_{I,t}|))) \tag{4.42}$$

где је са T обележен тренутак у коме се постиже максимум. Уколико је задовољен услов $\hat{\lambda}_T = |\hat{\lambda}_{A,T}| > c$, детектујемо присуство адитивне нестандардне опсервације у тренутку T. Аналогно, у случају да важи $\hat{\lambda}_T = |\hat{\lambda}_{I,T}| > c$ закључујемо да се ради о нестандардној опсервацији иновативног типа. У пракси се за вредност константе c узимају вредности 3.0, 3.5 или 4.0.

У зависности од типа нестандардне опсервације модификујемо оцене резидуала. У случају адитивног модела, нова оцена резидуала је $\tilde{\epsilon_t} = \hat{\epsilon_t} - \hat{\omega}_{A,T}\hat{\pi}(L)I_t$, док је код иновативног модела модификована оцена $\tilde{\epsilon_t} = \hat{\epsilon_t} - \hat{\omega}_{I,T}$. На основу модификованих оцена рачунамо нову варијансу резидуала $\tilde{\sigma_e}^2$.

3. На основу нових резидуала $\tilde{\epsilon}_t, t = 1, 2, ..., n$ понављамо поступак из претходног корака, све док не пронађемо све нестандардне опсервације.

Претпоставимо да смо користећи изложену процедуру детектовали присуство нестандардних опсервација у тренуцима $T_1, T_2, ..., T_k$. Коначни модел серије Y_t која садржи структурни лом биће облика 4.43. Серију ослобођену нестандардних опсервација добијамо користећи израз:

$$X_t = Y_t - \sum_{j=1}^{k} \hat{\omega}_j \hat{\nu}_j(L) I_t^{(T_j)} \tag{4.43}$$

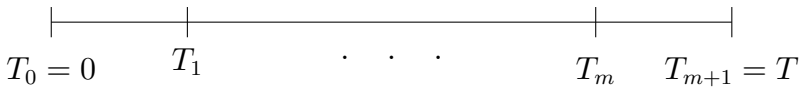
Процедуру понављамо док не добијемо задовољавајући модел.

Општи регресиони модел

До сада смо се бавили моделовањем структурног ломна у оквиру ARMA модела. У овом одељку бавићемо се анализом структурног лома у непознатим тренуцима, у контексту општег регресионог модела која је изложена у [12]. Модел који анализирамо дат је у следећем облику:

$$y_t = x_t'\beta + z_t'\delta_j + u_t \tag{4.44}$$

где је $t = T_{j-1} + 1, ..., T_j$, за $j = 1, ..., m + 1$. У претходном изразу, са m је означен број структурних ломова, а са $T_1, T_2, ..., T_m$ тренуци у којима се ти ломови дешавају. Ово значи да у кретању серије можемо разликовати $m + 1$ интервала облика $[T_{j-1} + 1, T_j]$, при чему је $T_0 = 0$ и $T_{m+1} = T$.



Вектори $x_t(p \times 1)$ и $z_t(q \times 1)$ представљају независне променљиве (са x_t' и z_t' означени су транспоновани вектори), док је са u_t означен шум. Са β и δ_j означени су вектори регресионих коефицијената уз x_t и z_t, респективно. Приметимо да коефицијенти δ_j зависе од структурног лома који се десио у тренутку T_j, док су коефицијенти β константни, односно нису под утицајем структурних промена. Другим речима, регресиони коефицијенти δ_j различити су у сваком од $m + 1$ интервала. На овај начин описан је парцијални модел структурног лома. Уколико важи $p = 0$, у питању је чист модел структурног лома.

Систем једначина 4.44 за $t = 1, ..., T$ можемо записати у матричном облику:

$$Y = X\beta + \overline{Z}\delta + U \tag{4.45}$$

У наставку су представљени чиниоци који фигуришу у једнакости 4.45:

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix} \qquad X = \begin{pmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,p} \\ x_{2,1} & x_{2,2} & \dots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{T,1} & x_{T,2} & \dots & x_{T,p} \end{pmatrix} \qquad U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix}$$

Матрица $\overline{Z}$ је дијагонална блок матрица облика $\overline{Z} = diag(Z_1, Z_2, ..., Z_{m+1})$, при чему важи:

$$Z_i = \begin{pmatrix} z_{T_{i-1}+1} \\ z_{T_{i-1}+2} \\ \vdots \\ z_{T_i} \end{pmatrix} = \begin{pmatrix} z_{T_{i-1}+1,1} & z_{T_{i-1}+1,2} & \dots & z_{T_{i-1}+1,q} \\ z_{T_{i-1}+2,1} & z_{T_{i-1}+2,2} & \dots & z_{T_{i-1}+2,q} \\ \vdots & \vdots & \ddots & \vdots \\ z_{T_i,1} & x_{T_i,2} & \dots & x_{T_i,q} \end{pmatrix}$$

$$\delta = \begin{pmatrix} \delta_1' \\ \delta_2' \\ \vdots \\ \delta_{m+1}' \end{pmatrix} = (\delta_{1,1}, \delta_{1,2}, ..., \delta_{1,q}, ..., \delta_{m+1,1}, \delta_{m+1,2}, ..., \delta_{m+1,q})'$$

Циљ је оценити непознате векторе коефицијената β и δ_j, као и непознате тренутке ломова $T_1, T_2, ..., T_m$, на основу доступних опсервација (y_t, x_t, z_t). За оцењивање користимо методу најмањих квадрата. Узмимо произвољну поделу интервала $[0, T]$ на $m+1$ делова, и означимо је са $\{T_j\} = (T_1, ..., T_m)$. На основу наведене поделе и модела 4.45, добијамо следећу суму квадрата резидуала:

$$(Y - X\beta - \overline{Z}\delta)'(Y - X\beta - \overline{Z}\delta) = \sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} (y_t - x_t'\beta - z_t'\delta_i)^2 \tag{4.46}$$

Минимизовањем наведене суме добијамо оцене регресионих коефицијената β и δ за поделу $\{T_j\}$, означене са $\hat{\beta}(\{T_j\})$ и $\hat{\delta}(\{T_j\})$. Замењујући добијене оцене у 4.45 поново добијамо израз за суму квадрата резидуала, као функцију поделе $\{T_j\} = (T_1, ..., T_m)$:

$$S(T_1, ..., T_m) = (Y - X\hat{\beta}(\{T_j\}) - \overline{Z}\hat{\delta}(\{T_j\}))'(Y - X\hat{\beta}(\{T_j\}) - \overline{Z}\hat{\delta}(\{T_j\})) \tag{4.47}$$

Оцена тренутака структурних ломова $(\hat{T}_1, ..., \hat{T}_m)$ добија се минимизовањем суме $S(T_1, ..., T_m)$ по свим поделама $(T_1, ..., T_m)$:

$$(\hat{T}_1, ..., \hat{T}_m) = argmin_{(T_1,...,T_m)} S(T_1, ..., T_m) \tag{4.48}$$

Коначне оцене регресионих коефицијената представљају функције оцене $\{\hat{T}_j\} = (\hat{T}_1, ...\hat{T}_m)$, односно важи $\hat{\beta} = \hat{\beta}(\{\hat{T}_j\})$, $\hat{\delta} = \hat{\delta}(\{\hat{T}_j\})$.
Како бисмо могли да опишемо граничну расподелу оцена $(\hat{T}_1, ...\hat{T}_m)$, потребно је увести неке претпоставке, које се тичу скупа дозвољивих параметара, тренутака ломова, регресора и грешака. Пре навођења претпоставки, дефинисаћемо ω_t као вектор свих регресора, $\omega_t = (x_{t_1}, ..., x_{t_p}, z_{t_1}, ..., z_{t_q})$.

1. **Претпоставка о регресорима**: За свако $j = 1, 2, ..., m+1$ и $l_j \leq T_j - T_{j-1}$ важи:

$$\frac{1}{l_j} \sum_{t=T_{j-1}+1}^{T_{j-1}+l_j} \omega_t \omega_t' \xrightarrow{c.c.} Q_j, l_j \to \infty \tag{4.49}$$

   где је Q_j неслучајна позитивно дефинитна матрица. Суштина ове претпоставке је у томе да су регресори, односно вектори w_t, реализација стационарног случајног низа, односно временске серије, чиме се ограничава ниво нестационарности серије чије ломове анализирамо.

2. **Претпоставка о грешкама**: Претпоставке о грешкама наведеног регресионог модела формулишу се преко разних облика слабе зависности случајних низова (L^r-mixingale низови) и омогућавају извођење граничних резултата и асимптотског понашања оцена за тренутке структурног лома.

3. **Претпоставка о ограничењу за минимизацију**: Минимизација из једнакости 4.48 врши се на скупу свих подела интервала $[0, T]$ таквим да важи $T_i - T_{i-1} \geq \epsilon T$, за неко $\epsilon > 0$. Као што је наведено у [3] ово ограничење односи се на ситуацију када су као регресори дозвољене прошле вредности зависне променљиве. У том случају сваки од сегмената који се разматрају приликом рачунања мимимума мора садржати позитиван део укупног узорка. Како ϵ може бити произвољно мала величина, ова претпоставка не представља практично ограничење.

4. **Претпоставка о тренуцима ломова**: Нека су са T_j^0 означени стварни непознати тренуци структурних ломова. Тада важи $T_j^0 = [T\lambda_j^0]$, где је $0 < \lambda_1^0 < ... < \lambda_m^0 < 1$. Ова претпоставка осигурава да су тренуци ломова асимптотски различити. Када расте обим узорка T, дужина свих интервала $[T_{j-1}, T_j]$ сразмерно расте. Уколико ово не би важило, тада би са повећањем узорка, односно додавањем нових података (под условом да нема нових ломова) дужина последњег интервала расла, док би дужине преосталих интервала постале занемарљиве, чиме би се пореметио оквир анализе коју разматрамо у овом поглављу.

На крају ћемо споменути један актуелан проблем у анализи стркутурног лома, у контексту разматраног општег регресионог модела. Наиме, како је наведено у раду [11] изложена процедура оцењивања тренутака лома додатно се компликује у случају када се лом дешава на самом крају узорка. Посматрајмо модел облика:

$$y_t = x_t'\beta + x_t'\delta I(t > T + k_0) + u_t \tag{4.50}$$

где је $t = 1, ..., T, T + 1, ..., T + m$. Са x_t означен је p-димензиони регресор, β представља p-димензиони вектор коефицијената, док је $\delta \neq 0$. Ово значи да се у тренутку $T + k_0$ дешава промена нивоа серије, док је пре тог тренутка нема структурних промена. Дефинишимо удео опсервација које претходе

структурном лому, у односу на цео узорак:

$$r = \frac{T + k_0}{T + m} \tag{4.51}$$

Јасно је да када $T \to \infty$, $r \to 1$, што значи да је у асимптотској анализи удео узорка након појаве структурног лома занемарљив.

Укратко ћемо изложити идеју изнесену у раду [11]. Посматрајмо резултате статистичког теста чија нулта хипотеза гласи H_0: $k_0 = k_1$, док је алтернативна хипотеза H_1: $k_0 \neq k_1$, са нивоом значајности α. Наведено тестирање понавља се за $k_1 = 0, 1, ..., m$. Интервал поверења са нивоом поверења $1 - \alpha$ гради се од свих тренутака $T + k_1$ таквих да је за k_1 прихваћена нулта хипотеза.

Глава 5

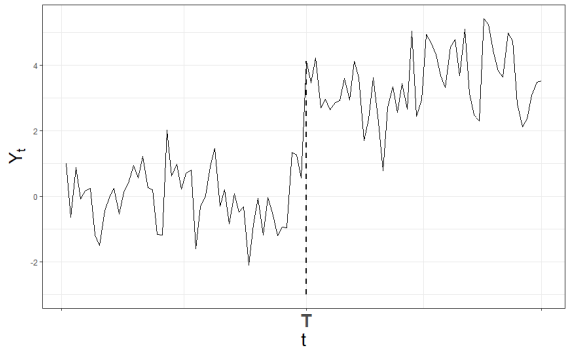
Детектовање структурног лома

Постоји више метода и тестова за детектовање присуства структурног лома у временској серији. То је широко поље истраживања и не постоји универзални тест који би био примењив у свакој ситуацији. Како је наведено у [4], стандардне процедуре као што је Студентов тест(Student's T-test), којим би се утврдило постојање промене у средњој вредности, не могу се користити у случају аналаизе временских серија, јер није задовољена претпоставка независности опсервација. Избор оптималног теста зависи од фактора као што је број структурних ломова, и информација о тренуцима интервенција. Додатно, осим формалних тестова, корисно је анализирати саме податке, и појаву која их је генерисала.

5.1 Анализа графика

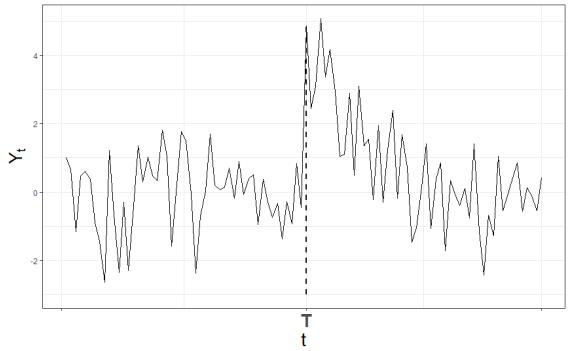
Прелиминарни метод за проналажење структурног лома је анализа графика временске серије. Уколико на графику можемо уочити промене у кретању временске серије, односно одударање од претходног понашања серије, можемо посумњати на присуство структурног лома. Иако не можемо говорити о егзактном доказу присуства структурног лома, визуелна анализа графика је добра полазна тачка, након које се може наставити са применом неког од тестова које ћемо описати у наставку.

На слици 5.1 приказан је график временске серије са структурним ломом у тренутку T који је трајног карактера. До тог тренутка, модел процеса који генерише серију је облика $Y_t = 0.1Y_{t-1} + e_t$, док је након тог тренутка у питању модел $Y_t = 3 + 0.1Y_{t-1} + e_t$, из чега следи да појаву структурног лома



Слика 5.1: График временске серије са трајним структурним ломом

карактерише промена нивоа серије, што се јасно види са графика.



Слика 5.2: График временске серије са привременим структурним ломом

На слици 5.2 дат је приказ графика у случају када серија поседује структурни лом привременог карактера. У овом случају, ниво серије након почетног скока у тренутку T експоненцијално опада ка првобитним вредностима.

5.2 Доменско знање

Закључивање на основу знања о временској серији коју анализирамо, односно о феномену који та серија описује (доменско знање), представља значајну методу која, иако се не може формализовати, може играти значајну улогу у утврђивању присуства структурног лома. Примера ради, уколико посматрамо временску серију индекса индустријске производње у СР Југославији[10, p. 296], знајући за санкције уведене 30. маја 1992. године од стране међународне заједнице и њихов утицај на наведени економски параметар, можемо

закључити да је у поменутој временској тачки дошло до појаве структурнг лома. Слично можемо резоновати и у случају серије мерења концентрације кисеоника у Лос Анђелесу, када имамо сазнање о променама законске регулативе која директно утиче на поменуту појаву. Оваква разматрања долазе у обзир и у случајевима познатих интервенција као што су промена фискалне политике, увођење технолошке иновације и слично. И у овом случају, као и код анализе графика серије, потврду постојања структурног лома добићемо применом у ту сврху дизајнираних тестова.

5.3 Тестови за утврђивање присуства структурног лома

Тестови за утврђивање присуства структурног лома могу се поделити у више врста на основу тога да ли откривају тачно један, или више структурних ломова у серији, да ли се могу примењивати секвенцијално (за откривање нових тренутака ломова између два постојећа), да ли претпостављају нека додатна ограничења за параметре или грешку модела итд. Овај проблем детаљно је разматран у раду [12]. У наставку ћемо навести неколико познатијих тестова.

Жарк-Бера тест нормалности

Један од статистичких тестова који може указати на присуство екстремних догађаја, а самим тим и структурног лома јесте Жарк-Бера тест нормалности (Jarque-Bera nornality test)[1][9, p. 224]. Наиме, на основу почетне спецификације модела којим описујемо посматрану серију, анализирамо резидуале оцењеног модела. Уколико утврдимо да расподела резидуала одудара од нормалне, можемо говорити о присуству екстремних догађаја, и измени почетног модела како би се у њега уврстила информација о таквим догађајима.
Нулта хипотеза гласи да резидуали имају нормалну расподелу. Тест статистика која се користи је облика:

$$JB = \frac{n}{6}(S^2 + \frac{1}{4}(K-3)^3) \tag{5.1}$$

[1]Carlos M. Jarque Uribe, Anil K. Bera

где је са n означен број опсервација, док су S и K ознаке за коефицијент асиметрије (skewness) и спљоштености (kurtosis), редом. Оцене ових коефицијената дате су са:

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^{\frac{3}{2}}}$$
$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^{2}}$$

при чему су $\hat{\mu}_3$ и $\hat{\mu}_4$ оцене трећег и четвртог момента расподеле, а $\hat{\sigma}^2$ оцена варијансе. При нултој хипотези статистика JB има χ^2_2 расподелу. Како је при нормалној расподели $S = 0$, $K = 3$, из 5.1 следи да са порастом одступања ове две оцене од датих вредности расте вредност статистике JB, те стога нулту хипотезу одбацујемо уколико је вредност те статистике већа од критичне вредности, која ће у случају нивоа значајности α бити једнака квантилу $\chi^2_{1-\alpha,2}$ расподеле.

Чоу тест

Чоу (Chow[2]) тест је један од традиционалних тестова присуства структурног лома у случају када је познат тренуткак интервенције[7, p. 622]. Додатно ограничење за употребу овог теста је немогућност његове примене у случају присуства више од једног лома у серији. Идеја која лежи у основи овог теста је поређење коефицијената две линеарне регресије, дефинисане над два скупа опсервација, пре и након тренутка интервенције. Нека је тренутак интервенције означен са T. Дефинишемо регресиони модел на целом скупу опсервација $1, 2, ..., n$:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + e_t$$

као и два регресиона модела на одвојеним скуповима опсервација пре и после тренутка T:

$$Y_t = \begin{cases} \phi_1^{(1)} Y_{t-1} + \phi_2^{(1)} Y_{t-2} + ... + \phi_p^{(1)} Y_{t-p} + e_t, & t = 1, 2, ..., T \\ \phi_1^{(2)} Y_{t-1} + \phi_2^{(2)} Y_{t-2} + ... + \phi_p^{(2)} Y_{t-p} + e_t, & t = T+1, T+2, ..., n \end{cases}$$

За нулту хипотезу узима се $H_0 : \phi_1^{(1)} = \phi_1^{(2)}, \phi_2^{(1)} = \phi_2^{(2)}, ..., \phi_p^{(1)} = \phi_p^{(2)}$, док је алтернативна хипотеза $H_1 : \phi_k^{(1)} \neq \phi_k^{(2)}$ за неко $k \in 1, 2, ..., n$. Дакле, нулта

[2]Gregory Chi-Chong Chow

хипотеза каже да се подаци пре и после тачке T могу се описати истом регресионом правом, те се не може говорити о структурном лому. Алтернативна хипотеза претпоставља да се опсервације пре и након тренутка T описују различитим регресионим правама, што указује на постојање структурног лома у тој тачки.

Рачунамо суму квадрата резидуала за све три регресије:

$$RSS_{1,n} = \sum_{i=1}^{n} e_{i,1}^2$$

$$RSS_{1,T} = \sum_{i=1}^{T} e_{i,2}^2$$

$$RSS_{T+1,n} = \sum_{i=T+1}^{n} e_{i,3}^2$$

Дефинишемо тест статистику:

$$F_T = \frac{(n-2p)}{p} \frac{RSS_{1,n} - (RSS_{1,T} + RSS_{T+1,n})}{(RSS_{1,T} + RSS_{T+1,n})} \tag{5.2}$$

При нултој хипотези статистика F_T има $F_{p,n}$ расподелу. Из 5.2 следи да ће, када важи H_0, односно уколико су регресиони параметри једнаки за опсервације пре и након тренутка T, важити $e_{i,1} = e_{i,2} = e_{i,3}$ за свако $i = 1, ..., n$, те ће вредност тест статистике бити једнака нули. Закључујемо да ћемо нулту хипотезу одбацити у корист алтернативне уколико је реализована вредност статистике F_T већа од критичне вредности, која је за ниво значајности α једнака квантилу реда $1 - \alpha$ за $F_{p,n}$ расподелу.

Надоградња Чоу теста која решава проблем његове непримењивости у случају непознатог тренутка интервенције је Квантов[3] тест количника веродостојности (The Quandt likelihood ratio test). Основна идеја овог теста састоји се у томе да се тест статистика 5.2 рачуна у свим временским тачкама $T_1, T_2, ..., T_k$ у којима постоји могућност присуства структурног лома. За Квантову тест статистику узима се $\max_k T_1, T_2, ..., T_k$. Овај тест је ушао у употребу када је развијена, до тада непозната, гранична расподела наведене статистике[2].

[3]Richard E. Quandt

CUSUM тест

Тест кумулативних сума[14] базира се на парцијалним сумама рекурзивних резидуала. Претпоставимо да посматрамо линеарни регресиони модел са k предиктора $y_t = \mathbf{x}_t^T\beta + \epsilon_t$, где је $\mathbf{x}_t$ вектор облика $(x_{t,1}, x_{t,2}, ..., x_{t,k})^T$. Рекурзивни резидуал дефинишемо као:

$$r_t = \frac{y_t - \mathbf{x}_t^T\hat{\beta}_{t-1}}{\sqrt{v_t}} \tag{5.3}$$

где је $\hat{\beta}_{t-1} = (\hat{\beta}_{t-1,1}, \hat{\beta}_{t-1,2}, ..., \hat{\beta}_{t-1,k})^T$ оцена вектора регресионих параметара $\beta = (\beta_1, \beta_2, ..., \beta_k)^T$ на основу првих $t-1$ опсервација, а $v_t = 1+\mathbf{x}_t^T(\sum_{k=1}^{t-1}\mathbf{x}_k\mathbf{x}_k^T)^{-1}\mathbf{x}_t$ Рекурзивни резидуал је заправо стандардизована грешка прогнозе за један корак унапред.

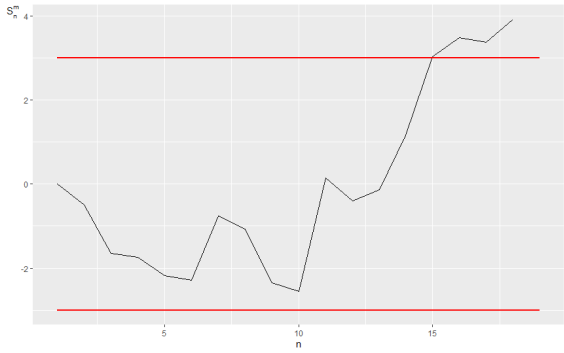
Нека је m тренутак од кога пратимо понашање вектора параметара β. Нулта хипотеза гласи да је вектор параметара константан кроз време ($H_0 : \beta_t = \beta, t = m+1, t = m+2, ...$), док алтернативна говори о нестабилности параметара. Посматрамо парцијалне суме рекурзивних резидуала 5.3:

$$S_n^m = \frac{1}{\hat{\sigma}}\sum_{t=m}^{m+n} r_t \tag{5.4}$$

где је $\hat{\sigma}^2$ конзистентна оцена варијансе процеса ϵ_t. Дефинишемо границу

$$L = \sqrt{n+m-k}\sqrt{c + ln(\frac{n+m-k}{m-k})} \tag{5.5}$$

где константа c зависи од величине теста[7, p. 625]. Уколико за неко $n \geq 1$ важи $|S_n^m| > L$ одбацујемо нулту хипотезу у корист алтернативне о нестабилности параметара. Валидност оваквог закључивања почива на чињеници да су, уколико важи H_0, рекурзивни резидуали $r_m, r_{m+1}, ..., r_T$ независне случајне величине са $N(0, \sigma^2)$ расподелом[14]. Одступање низа парцијалних сума 5.4, од нуле оспорава претпоставку о нултој вредности очекивања резидуала, те се нулта хипотеза одбацује у корист алтернативне. Понашање низа 5.4 може се анализирати помоћу графичке репрезентације дате на слици 5.3. Хоризонталне црвене линије представљају границе 5.5. Уколико крива вредности низа парцијалних сума прелази ове границе, одбацујемо нулту хипотезу у корист алтернативе.



Слика 5.3: Кретање низа парцијалних сума рекурзивних резидуала

Секвенцијални тест

Наведимо још једну врсту тестова (изложених у раду [12]) специфичних по томе што су дизајнирани за детектовање нових ломова између већ постојећих.

Претпоставимо да имамо модел са l ломова детектованих у тачкама ($\hat{T}_1$, ... , $\hat{T}_l$). На сваком од $l+1$ интервала одређених овим тачкама, врши се тестирање на присуство додатног структурног лома, при чему се на тим интервалима рачунају суме квадрата резидуала. Затим се врши поређење минумима добијених сума, са сумом квадрата резидуала за модел са l ломова. Уколико је добијени минимум довољно мали, одбацује се нулта хипотеза о одсуству додатног структурног лома. Тачка у којој је постигнута минимална сума квадртата резидуала управо је тачка у којој је детектован додатни лом. Формално тест дефинишемо на следећи начин:

$$W_T(l+1|l) = \{S_T(\hat{T}_1, ..., \hat{T}_l) - \min_{1 \leq i \leq l+1} \inf_{\tau \in \Lambda_{i,\epsilon}} S_T(\hat{T}_1, ..., \hat{T}_{i-1}, \tau, \hat{T}_i, ..., \hat{T}_l)\}/\sigma^2 \quad (5.6)$$

при чему је $S_T()$ сума квадрата резидуала, док је са $\Lambda_{i,\epsilon}$ означен скуп свих тачака из интервала $(\hat{T}_{i-1}, \hat{T}_i)$ које су на одређеној удаљености од граница интервала (односно тренутака постојећих структурних ломова):

$$\Lambda_{i,\epsilon} = \{\tau | \hat{T}_{i-1} + (\hat{T}_i - \hat{T}_{i-1})\epsilon \leq \tau \leq \hat{T}_i - (\hat{T}_i - \hat{T}_{i-1})\epsilon \quad (5.7)$$

Овакав тест формира темељ процедуре секвенцијалног тестирања, која подразумева узастопну примену наведеног теста, почевшти од $l = 0$, све до итерације у којој нулта хипотеза о одсуству лома не буде одбачена. Међутим како је наведено у [12], ову процедуру не би требало примењивати механички, због

могућности прераног заустављања. Препоручљиво је утврдити постојање било каквог лома у серији, а затим применити процедуру за неко $l > 0$.

Тестирање присуства структурног лома у случају хетероскедастичне серије

На крају овог поглавља, споменућемо нову и актуелну тему истраживања на пољу анализе временских серија. У питању је тестирање на присуство структурног лома у случају када је серија хетероскедастична. Појам хетероскедастичности односи се на расподелу резидуала. Наиме, уколико је варијанса резидуала функција времена, односно ако постоји систематска промена у распршењу резидуала током времена, кажемо да је временска серија хетероскедастична.
Изложићемо разматрања аутора рада [13]. Како аутори наводе, промена варијансе мења претпоставке о асимптотској расподели конвенционалних тестова на присуство лома, као што је CUSUM. Предлажу се два начина за креирање тестова робустних на хетероскедастичност резидуала.
Први начин подразумева групу метода (two-steps methods) у чијој је основни процедура која се састоји из два корака. Први корак је оцењивање резидуала при нултој хипотези, односно уз претпоставку да у серији није присутан лом. У другом кораку оцењени резидуали се користе како би се направила корекција при којој је асимптотска дистрибуција тест статистике валидна при услову хетероскедастичности.
Други начин, у контексту регресионог модела, подразумева дизајнирање тестова којима се проверава постојање лома како у коефицијентима модела, тако и у варијанси резидуала, При томе, тренуци ломова у коефицијентима и резидуалима могу да буду различити или да се преклапају. Циљ је тестирање на присуство лома у коефицијентима, при дозвољеним променама у варијанси резидуала.

5.4 Структурни лом код серија са јединичним кореном

Захваљујући значају потенцијалног присуства структурног лома и јединичног корена у временској серији приликом практичних економских ана-

лиза, однос између ова два феномена тема је многобројних радова, као што су [12] и [16]. Суштина овог односа лежи у чињеници да ће већина тестова за утврђивање присуства јединичног корена прихватити модел са јединичним кореном, уколико је у серији присутан структурни лом, иако је серија стационарна или тренд стационарна у периодима између тренутака ломова. Са друге стране, тестови за детектовање структурног лома прихватиће хипотезу о постојању лома, уколико је у серији присутан јединични корен, и поред константних параметара модела [12]. Због тога се уобичајени тестови којима се утврђује присуство јединичног корена морају прилагодити у ситуацији када је у серији присутан структурни лом.

У овом контексу навешћемо поставку изложену у раду [12]. Претпоставимо да посматрамо модел временске серије облика:

$$y_t = \mu + \beta t + u_t \tag{5.8}$$

У питању је серија са присутним трендом. Члан u_t дат је једнакошћу $\Delta u_t = C(L)e_t$, при чему је e_t $i.i.d.(0, \sigma_e^2)$. За полином $C(L)$ важи $C(L) = \sum_{j=0}^{\infty} c_j L^j$, при чему је $\sum_{j=0}^{\infty} j|c_j| < \infty$, $c_0 = 1$. Када на серију применимо декомпозицију на тренд и цикличну компоненту, тренд компонента се може представити на следећи начин:

$$\tau_t = \mu + \beta t + C(1) \sum_{j=1}^{t} e_j \tag{5.9}$$

На основу једнакости 5.9 можемо закључити да се и сама компонента тренда састоји из два дела, детерминистичке компоненте линеарног тренда, и стохастичке компоненте представљене случајним ходом помноженим фактором $C(1)$. Другим речима, у сваком тренутку присутна је промена у тренду серије. Треба приметити да уколико серија нема јединични корен важи $C(1) = 0$, и тренд је детерминистичка функција времена. Наведена поставка наводи на закључак да се проблем разликовања тренд стационарне серије од серије са јединичним кореном своди на следеће питање: да ли се тренд мења у сваком тренутку или никада? Емпиријски је показано је да ће конкретни подаци фаворизовати тренд који се стално мења, у односу на тренд који се никада не мења. Због тога аутор рада [12] предлаже алтернативну формулацију: да ли се тренд мења у сваком тренутку или понекад? Кључно питање је шта се подразумева под термином понекад. Како је наведено у истом раду, у

многим практичним разматрањима број промена у кретању тренда је мали, а најчешће је у питању само једна промена повезана са одређеним догађајем важним за кретање серије. Статистички посматрано, ове промене потичу из другачије расподеле у односу на случајне шокове присутне у сваком тренутку. Међутим, ову расподелу је тешко одредити управо због мале фреквенције појава, те се овакве промене у тренду моделују као структурни ломови. У том случају се тестирање присуства јединичног корена разматра уз услов да је дозвољена могућност појаве ретких трајних промена (структурних ломова) у тренду временске серије.
У раду [12] наведена су четири типа промене у кретању серије.

1. Промена нивоа серије у чијем кретању није присутан тренд
2. Промена нивоа тренда
3. Промена нагиба тренда
4. Промена нивоа и нагиба тренда

За сваки од ова четири типа можемо посматрати модел код кога се промена дешава тренутно (модел адитивног аутлајера) и модел код кога се промена дешава постепено (модел иновативног аутлајера), што повлачи различите статистичке процедуре за тестирање присуства јединичног корена у серији [12], [9, p. 225]. У наставку су изложене спецификације наведених модела дате у раду [12]. У случају адитивних модела, које означавамо са AO (енг. aditive outlier), имамо:

1. Модел AO - 0
   $y_t = \mu_1 + (\mu_2 - \mu_1)S_t + u_t$
2. Модел AO - A
   $y_t = \mu_1 + \beta t + (\mu_2 - \mu_1)S_t + u_t$
3. Модел AO - B
   $y_t = \mu_1 + \beta_1 t + (\beta_2 - \beta_1)S_t^* + u_t$
4. Модел AO - C
   $y_t = \mu_1 + \beta_1 t + (\mu_2 - \mu_1)S_t + (\beta_2 - \beta_1)S_t^* + u_t$

Иновативне моделе, које означавамо са IO (енг. inovational outlier), представљамо засебно у случају важења нулте хипотезе (постојање јединичног корена), односно алтернативне хипотезе (тренд-стационарна серија). При нултој хипотези модели су облика:

1. Модел IO-0-UR
   $y_t = y_{t-1} + C(L)(e_t + \delta I_t)$
2. Модел IO-A-UR
   $y_t = y_{t-1} + b + C(L)(e_t + \delta I_t)$
3. Модел IO-C-UR
   $y_t = y_{t-1} + b + C(L)(e_t + \delta I_t + \eta S_t)$

При алтернативној хипотези модели су дефинисани на следећи начин:

1. Модел IO-0-TS
   $y_t = y_{t-1} + C^*(L)(e_t + \theta I_t)$
2. Модел IO-A-TS
   $y_t = y_{t-1} + b + C^*(L)(e_t + \theta I_t)$
3. Модел IO-C-TS
   $y_t = y_{t-1} + b + C^*(L)(e_t + \theta I_t + \gamma S_t)$

У претходним једнакостима важи

$$I_t = \begin{cases} 1, t = T \\ 0, t \neq T \end{cases} \tag{5.10}$$

$$S_t^* = \begin{cases} t - T, t = T \\ 0, t \neq T \end{cases} \tag{5.11}$$

$$S_t = \begin{cases} 1, t > T \\ 0, t \neq T \end{cases} \tag{5.12}$$

Додатно важи $C^*(L) = (1 - L)^{-1}C(L)$.
Стандардни тест за утврђивање присуства структурног лома у временској

серији је проширени Дики-Фулеров тест (Augmented Dicky-Fuller test - ADF). Овај тест базира се на тестирању хипотезе да је $\alpha = 1$, где је α параметар у следећој регресионој једначини (ADF статистика):

$$y_t = \mu + \beta t + \alpha y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-1} + e_t \tag{5.13}$$

Када је у серији присутан структурни лом, моћ наведеног теста је смањена [12]. Степен смањења моћи теста заиси од степена промене у кретању серије. С обзиром да мале промене изазивају незнатну редукцију моћи теста, приликом тестирања у обзир треба узети само промене довољно велике да изазову значајно опадање моћи теста.

Уколио претпоставимо да нам је познат тренутак појаве структурног лома T, аутор рада [12] предлаже модификацију једнакости 5.13 укључивањем претходно дефинисаних вештачких променљивих I_t, S_t, S_t^*. У случају иновативног модела, посматрамо једначину:

$$y_t = \mu + \theta S_t + \beta t + \gamma S_t^* + \delta I_t + \alpha y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-1} + e_t \tag{5.14}$$

У зависности од конкретног модела, изузимају се одређени регресори из опште једначине 5.14. Процедура се донекле разликује у случају адитивног модела. У првом кораку, из серије y_t изузимамо тренд, чиме добијамо серију без тренда $\tilde{y}_t$. У другом кораку, за моделе АО-0, АО-А и АО-С имамо једначину:

$$\tilde{y}_t = \tilde{y}_{t-1} + \sum_{j=0}^{k} d_j I_{t-j} + \sum_{i=1}^{k} a_i \Delta \tilde{y}_{t-i} + e_t \tag{5.15}$$

Како се наводи у раду [12], у случају модела АО-В, нема потребе за укључивањем вештачких променљивих, јер нема промене нивоа серије, а два сегмента тренда су спојена у тренутку лома (промена је присутна само у нагибу тренда).

Како се наведена процедура односи на ситуацију када нам је познат тренутак лома Т, природно се поставља питање тестирања присуства јединичног корена у присуству структурног лома у непознатом тренутку. Навешћемо једну алтернативу која се примењује у том случају. То је Зивот-Ендрјусов тест (Zivot-Andrews test). Идеја иза овог теста је оцењивање тренутка лома

у поступку тестирања. Регресионе једначине су идентичне као код познатог тренутка лома, осим што се искључује вештачка променљива I_t. Међутим, кључна разлика је у формулацији хипотеза. Нулта хипотеза подразумева присуство јединичног корена у серији у којој није присутан лом, док алтернативна укључује тренд стационарну серију са ломом. Посматрају се сви временски тренуци изузев одређеног процента на почетку и крају низа (најчешће по 15 %), и за сваки од њих рачуна се вредност ADF статистике. Вредност теста је минимум израчунатих ADF статистика. Уколико је та вредност мања од одређене критичне вредности, прихвата се алтернативна хипотеза о тренд стационарној серији са ломом у тренутку за који је добијена минимална ADF статистика.

Глава 6

Практични примери

У овој глави бавићемо се практичном применом до сада изложених теоријских резултата, кроз неколико примера. Сва израчунавања биће вршена у програмском пакету R.

6.1 Стопа незапослености

Предмет анализе је временска серија коју сачињавају вредности стопе незапослености на месечном нивоу, у Сједињеним Америчким Државама, у периоду од децембра 2009. године до априла 2021. године[1]. Према дефиницији Међународне организације рада[2], стопа незапослености представља проценат лица која не раде, али активно траже запослење и у могућности су да почну да раде, у односу на укупну радну снагу[3]. У овом истраживању обухваћена су лица старија од 15 година. Подаци су дати у табели 6.1

На слици 6.1 дат је графички приказ података. Са графика се може видети да након марта 2020. године долази до нагле промене у кретању временске серије, што је прва индиција за постојање структурног лома. Овакво разматрање је додатно поткрепљено чињеницом да у том периоду долази до ескалације епидемије вируса COVID-19, што је неспорно утицало на ситуацију

[1]Подаци су преузети са U.S. Bureau of Labor Statistics, Unemployment Rate [UNRATE], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/UNRATE, September 10, 2021.

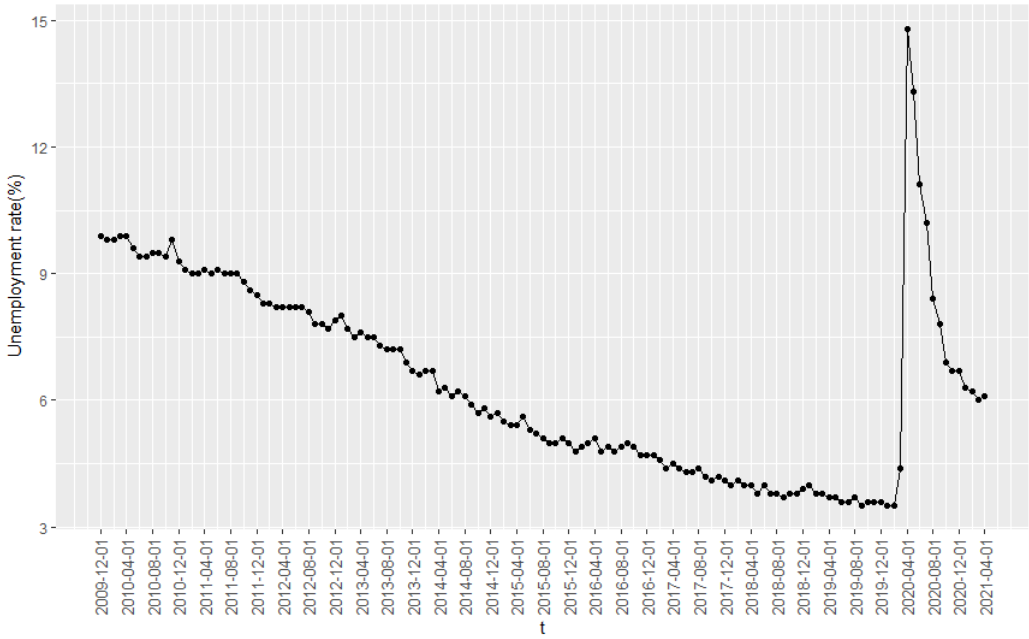
[2]Resolution concerning statistics of work, employment and labour underutilization, adopted by the 19th International Conference of Labour Statisticians, Geneva, October 2013

[3]Радну снагу чине сва запослена и незапослена лица.

Табела 6.1: Стопа незапослености 01.12.2009.-01.04.2021. (%)

Датум	Стопа	Датум	Стопа	Датум	Стопа	Датум	Стопа
2009-12-01	9.9	2012-11-01	7.7	2015-10-01	5.0	2018-09-01	3.7
2010-01-01	9.8	2012-12-01	7.9	2015-11-01	5.1	2018-10-01	3.8
2010-02-01	9.8	2013-01-01	8.0	2015-12-01	5.0	2018-11-01	3.8
2010-03-01	9.9	2013-02-01	7.7	2016-01-01	4.8	2018-12-01	3.9
2010-04-01	9.9	2013-03-01	7.5	2016-02-01	4.9	2019-01-01	4.0
2010-05-01	9.6	2013-04-01	7.6	2016-03-01	5.0	2019-02-01	3.8
2010-06-01	9.4	2013-05-01	7.5	2016-04-01	5.1	2019-03-01	3.8
2010-07-01	9.4	2013-06-01	7.5	2016-05-01	4.8	2019-04-01	3.7
2010-08-01	9.5	2013-07-01	7.3	2016-06-01	4.9	2019-05-01	3.7
2010-09-01	9.5	2013-08-01	7.2	2016-07-01	4.8	2019-06-01	3.6
2010-10-01	9.4	2013-09-01	7.2	2016-08-01	4.9	2019-07-01	3.6
2010-11-01	9.8	2013-10-01	7.2	2016-09-01	5.0	2019-08-01	3.7
2010-12-01	9.3	2013-11-01	6.9	2016-10-01	4.9	2019-09-01	3.5
2011-01-01	9.1	2013-12-01	6.7	2016-11-01	4.7	2019-10-01	3.6
2011-02-01	9.0	2014-01-01	6.6	2016-12-01	4.7	2019-11-01	3.6
2011-03-01	9.0	2014-02-01	6.7	2017-01-01	4.7	2019-12-01	3.6
2011-04-01	9.1	2014-03-01	6.7	2017-02-01	4.6	2020-01-01	3.5
2011-05-01	9.0	2014-04-01	6.2	2017-03-01	4.4	2020-02-01	3.5
2011-06-01	9.1	2014-05-01	6.3	2017-04-01	4.5	2020-03-01	4.4
2011-07-01	9.0	2014-06-01	6.1	2017-05-01	4.4	2020-04-01	14.8
2011-08-01	9.0	2014-07-01	6.2	2017-06-01	4.3	2020-05-01	13.3
2011-09-01	9.0	2014-08-01	6.1	2017-07-01	4.3	2020-06-01	11.1
2011-10-01	8.8	2014-09-01	5.9	2017-08-01	4.4	2020-07-01	10.2
2011-11-01	8.6	2014-10-01	5.7	2017-09-01	4.2	2020-08-01	8.4
2011-12-01	8.5	2014-11-01	5.8	2017-10-01	4.1	2020-09-01	7.8
2012-01-01	8.3	2014-12-01	5.6	2017-11-01	4.2	2020-10-01	6.9
2012-02-01	8.3	2015-01-01	5.7	2017-12-01	4.1	2020-11-01	6.7
2012-03-01	8.2	2015-02-01	5.5	2018-01-01	4.0	2020-12-01	6.7
2012-04-01	8.2	2015-03-01	5.4	2018-02-01	4.1	2021-01-01	6.3
2012-05-01	8.2	2015-04-01	5.4	2018-03-01	4.0	2021-02-01	6.2
2012-06-01	8.2	2015-05-01	5.6	2018-04-01	4.0	2021-03-01	6.0
2012-07-01	8.2	2015-06-01	5.3	2018-05-01	3.8	2021-04-01	6.1
2012-08-01	8.1	2015-07-01	5.2	2018-06-01	4.0		
2012-09-01	7.8	2015-08-01	5.1	2018-07-01	3.8		
2012-10-01	7.8	2015-09-01	5.0	2018-08-01	3.8		

на тржишту рада у САД. Да бисмо испитали присуство лома, применићемо Chow и CUSUM тест. Након имплементације Chow теста у програмском пакету R, за временску тачку март 2020. године, добијамо резултате сумиране у табели 6.2.



.

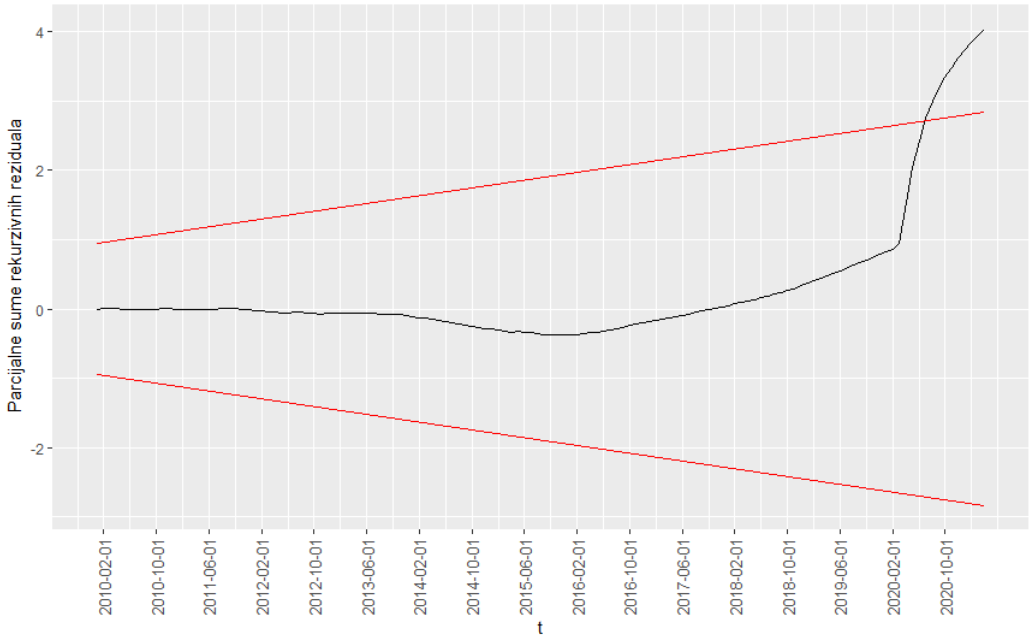
Слика 6.1: График временске серије стопе незапосленоси

Табела 6.2: Резултати Chow теста

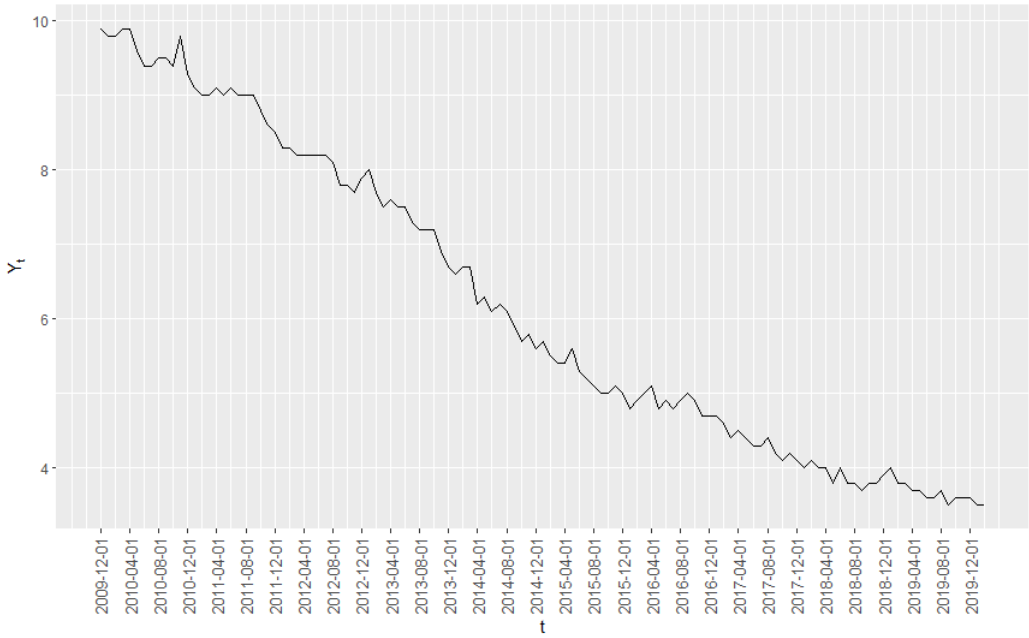
F статистика	p-вредност
695.41	$< 2.2e-16$

Велика вредност тест статистике, и мала p-вредност указују да треба одбацити нулту хипотезу, у корист алтернативне која говори о промени параметара кретања серије у тачки март 2020. Применом CUSUM теста у R-у, добијамо резултат приказан на слици 6.2, са које видимо да низ парцијалних сума рекурзивних резидуала прелази задате границе, што нас такође наводи на закључак о постојању структурног лома.

У следећем кораку, конструисаћемо ARIMA модел за серију стопа незапослености, која претходи тачки лома, користећи Бокс-Џенкинсову стратегију. График ове серије дат је на слици 6.3. Приметан је опадајући тренд стопе незапослености, што указује на опоравак економије након економске кризе из 2008. године. Поступком диференцирања трансформишемо серију у стационарну. Анализом корелограма и употребом Акаике информационог критеријума(AIC) за поређење различитих модела, долазимо до ARIMA(2,1,2) спецификације модела. На графику 6.4 дат је приказ прогнозираних вредности за период након марта месеца 2021. године (црвена боја) и стварних

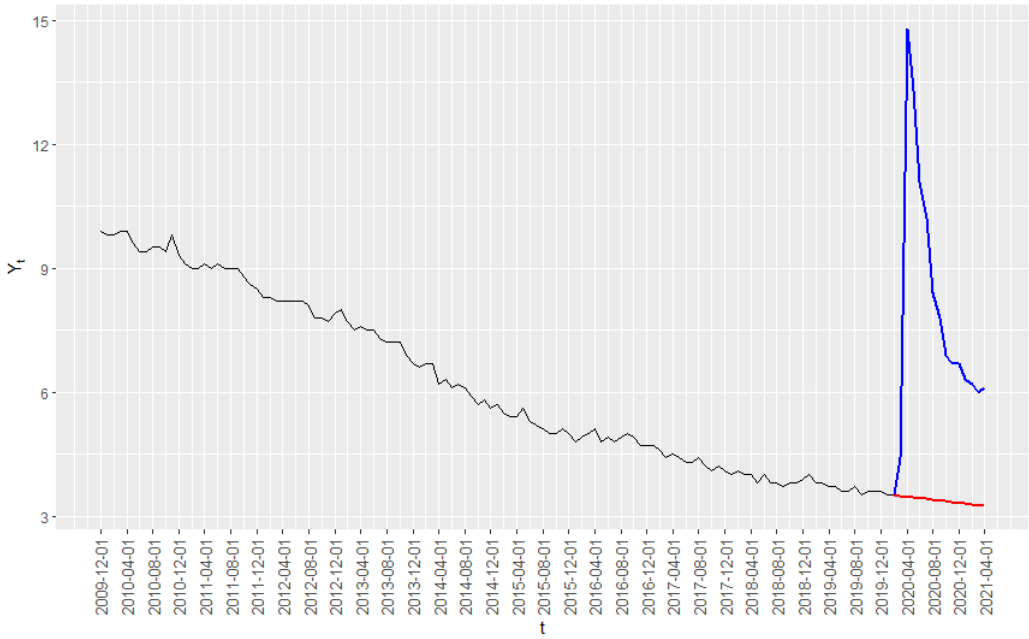


Слика 6.2: График временске серије са привременим структурним ломом



Слика 6.3: Стопа незапослености пре структурног лома

вредности серије које су последица структурног лома у марту 2020. Јасно је изражена потреба модификовања полазног модела како бисмо могли да прогнозирамо кретање стопе незапослености у наредном периоду. График



Слика 6.4: Поређење прогнозе на основу модела без структурног лома и стварних вредности серије

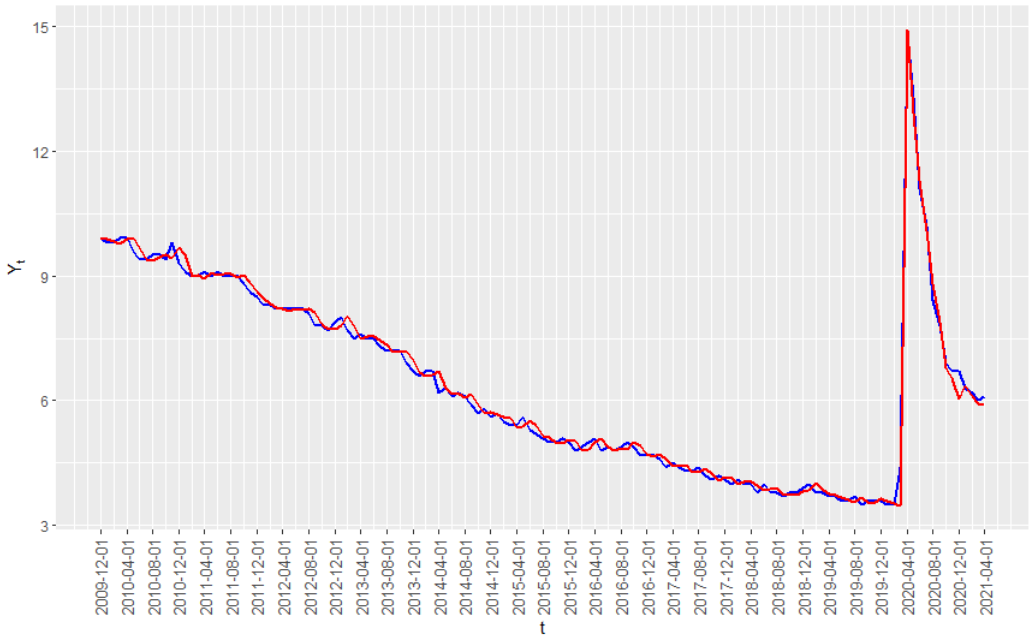
серије указује да је потребно применити модел 4.13. Битни је напоменути да се, иако се промена у кретању серије дешава у марту 2021, пун ефекат структурног лома очитава у априлу исте године. Наведена чињеница оправдава примену оператора доцње на импулсну променљиву IM_t у моделу 4.13. Поступак изложен у одељку 4.2 имплементирамо у програмском пакету R:

```
pulse = rep(0, length(t))
pulse[match("2020-04-01", t)] = 1
arimax(stopa nezaposlenosti,
       order = c(2,1,2), xtransf = data.frame(pulse),
       transfer = list(c(1,0)))
```

Варијабла pulse представља индикаторску променљиву 4.4 на коју је примењен оператор доцње. Ова варијабла се функцији arimax прослеђује као вредност аргумента xtransf. Аргумент transfer садржи вектор чије координате представљају редом степене полинома $\omega(L)$ и $\delta(L)$ из једнакости 4.3. Добијени модел описан је следећом једнакошћу:

$$X_t = \frac{10.7255}{1 + 0.7973L} LIM_t + \frac{1 + 0.0708L - 0.6155L^2}{1 + 0.1687L + 0.3257L^2} e_t \tag{6.1}$$

Анализом резидуала проверавамо адекватност модела. На графику 6.5 приказано је поређење стварних вредности серије стопе незапослености у периоду од децембар 2009. до априла 2021. године (плава линија), и вредности добијених моделом 6.1, у истом периоду (црвена линија).



Слика 6.5: Поређење модела 6.1 са стварним вредностима серије

6.2 Цена акција на берзи

У овом примеру, анализираћемо временску серију која представља цене акција компаније Infosys Limited, са седиштем у Бангалору, Индија. Ова мултинационална компанија послује у области информационих технологија, пружајући услуге развоја и одржавања софтвера за клијенте у области финансија, осигурања итд.

Серију чине цене акција на водећој берзи у Индији (National Stock Excange of India - NSE), која се налази у Мумбају, бележене на дневном нивоу, у периоду од 01.01.2015. до 31.12.2015[4]. Цене су бележене у индијским рупијама. Подаци су приказани у табели 6.3

[4]Подаци су преузези са www.kaggle.com/atulanandjha/national-stock-exchange-time-series

Табела 6.3: Цене акција 01.01.2015.-01.12.2015. (%)

Датум	Цена	Датум	Цена	Датум	Цена
2015-01-01	2545.55	2015-05-06	2463.05	2015-08-31	2564.05
2015-01-02	2579.45	2015-05-07	2545.25	2015-09-01	2537.40
2015-01-05	2540.25	2015-05-08	2527.25	2015-09-02	2591.80
2015-01-06	2446.60	2015-05-11	2556.00	2015-09-03	2605.60
2015-01-07	2417.70	2015-05-12	2522.90	2015-09-04	2547.45
2015-01-08	2443.80	2015-05-13	2515.40	2015-09-07	2539.75
2015-01-09	2512.30	2015-05-14	2499.25	2015-09-08	2540.30
2015-01-12	2509.70	2015-05-15	2511.15	2015-09-09	2549.75
2015-01-13	2497.90	2015-05-18	2520.40	2015-09-10	2545.40
2015-01-14	2521.95	2015-05-19	2512.70	2015-09-11	2552.00
2015-01-15	2539.10	2015-05-20	2557.05	2015-09-14	2551.65
2015-01-16	2532.15	2015-05-21	2555.50	2015-09-15	2547.60
2015-01-19	2511.00	2015-05-22	2617.90	2015-09-16	2550.35
2015-01-20	2500.45	2015-05-25	2601.00	2015-09-18	2551.25
2015-01-21	2513.80	2015-05-26	2608.85	2015-09-21	2551.05
2015-01-22	2513.55	2015-05-27	2615.25	2015-09-22	2526.80
2015-01-23	2503.60	2015-05-28	2609.50	2015-09-23	2532.65
2015-01-27	2502.05	2015-05-29	2610.30	2015-09-24	2580.40
2015-01-28	2539.60	2015-06-01	2617.75	2015-09-28	2593.70
2015-01-29	2544.65	2015-06-02	2599.90	2015-09-29	2588.25
2015-01-30	2482.05	2015-06-03	2610.90	2015-09-30	2588.05
2015-02-02	2514.20	2015-06-04	2605.65	2015-10-01	2644.60
2015-02-03	2558.25	2015-06-05	2575.85	2015-10-05	2711.65
2015-02-04	2514.30	2015-06-08	2577.45	2015-10-06	2700.00
2015-02-05	2552.40	2015-06-09	2561.65	2015-10-07	2653.95
2015-02-06	2575.75	2015-06-10	2600.75	2015-10-08	2633.80
2015-02-09	2512.90	2015-06-11	2571.30	2015-10-09	2629.65
2015-02-10	2441.15	2015-06-12	2505.80	2015-10-12	2592.60
2015-02-11	2459.90	2015-06-15	2504.80	2015-10-13	2598.55
2015-02-12	2462.15	2015-06-16	2512.00	2015-10-14	2483.70
2015-02-13	2538.75	2015-06-17	2513.80	2015-10-15	2467.75
2015-02-16	2584.80	2015-06-18	2530.05	2015-10-16	2473.75
2015-02-18	2635.65	2015-06-19	2549.55	2015-10-19	2488.75

Табела 6.3: Цене акција 01.01.2015.-01.12.2015. (%)

Датум	Цена	Датум	Цена	Датум	Цена
2015-02-19	2681.05	2015-06-22	2591.50	2015-10-20	2526.75
2015-02-20	2675.55	2015-06-23	2574.60	2015-10-21	2530.00
2015-02-23	2696.65	2015-06-24	2570.30	2015-10-23	2537.15
2015-02-24	2704.75	2015-06-25	2550.00	2015-10-26	2536.55
2015-02-25	2672.20	2015-06-26	2592.80	2015-10-27	2531.10
2015-02-26	2656.45	2015-06-29	2592.90	2015-10-28	2529.10
2015-02-27	2662.35	2015-06-30	2550.95	2015-10-29	2495.35
2015-02-28	2675.25	2015-07-01	2593.10	2015-10-30	2495.20
2015-03-02	2669.40	2015-07-02	2577.75	2015-11-02	2517.35
2015-03-03	2776.00	2015-07-03	2603.90	2015-11-03	2543.45
2015-03-04	2746.00	2015-07-06	2633.30	2015-11-04	2533.80
2015-03-05	2696.45	2015-07-07	2616.25	2015-11-05	2482.65
2015-03-09	2646.80	2015-07-08	2595.10	2015-11-06	2505.75
2015-03-10	2642.70	2015-07-09	2522.50	2015-11-09	2471.35
2015-03-11	2606.55	2015-07-10	2474.35	2015-11-10	2458.15
2015-03-12	2619.05	2015-07-13	2509.55	2015-11-11	2468.80
2015-03-13	2582.35	2015-07-14	2499.50	2015-11-13	2397.40
2015-03-16	2562.30	2015-07-15	2545.75	2015-11-16	2374.60
2015-03-17	2583.70	2015-07-16	2540.20	2015-11-17	2390.55
2015-03-18	2559.35	2015-07-17	2564.40	2015-11-18	2351.45
2015-03-19	2596.25	2015-07-20	2579.15	2015-11-19	2362.55
2015-03-20	2610.95	2015-07-21	2561.40	2015-11-20	2398.35
2015-03-23	2617.15	2015-07-22	2528.05	2015-11-23	2391.05
2015-03-24	2591.30	2015-07-23	2488.60	2015-11-24	2367.95
2015-03-25	2573.75	2015-07-24	2505.15	2015-11-26	2345.55
2015-03-26	2513.10	2015-07-27	2497.05	2015-11-27	2353.35
2015-03-27	2514.80	2015-07-28	2502.05	2015-11-30	2364.70
2015-03-30	2564.15	2015-07-29	2506.75	2015-12-01	2361.90
2015-03-31	2553.95	2015-07-30	2480.65	2015-12-02	2365.20
2015-04-01	2542.65	2015-07-31	2510.75	2015-12-03	2350.75
2015-04-06	2547.30	2015-08-03	2509.60	2015-12-04	2328.40
2015-04-07	2585.00	2015-08-04	2508.90	2015-12-07	2319.80
2015-04-08	2641.40	2015-08-05	2547.15	2015-12-08	2329.00

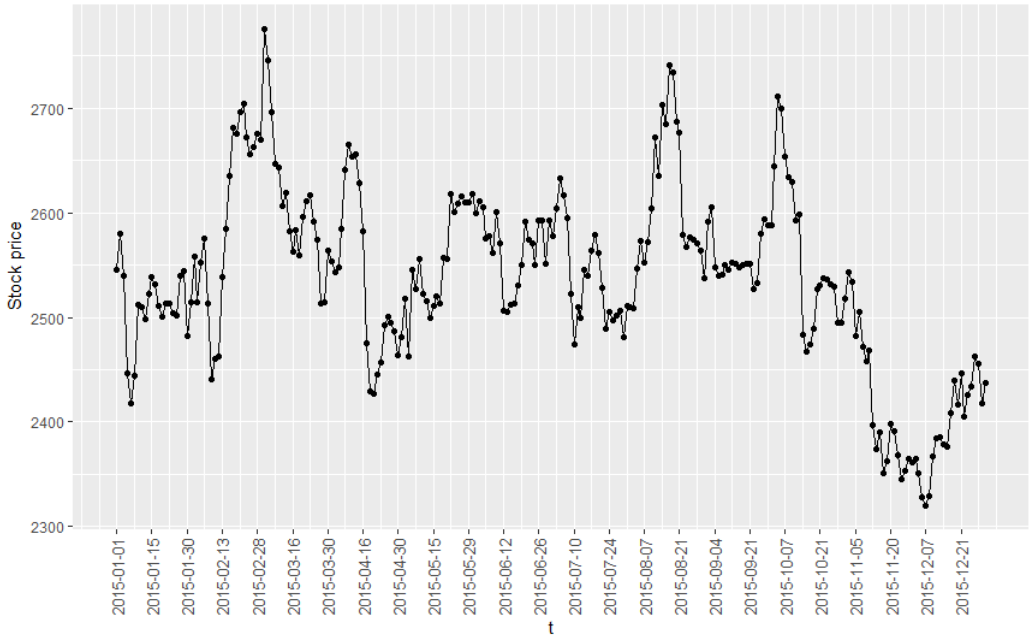
Табела 6.3: Цене акција 01.01.2015.-01.12.2015. (%)

Датум	Цена	Датум	Цена	Датум	Цена
2015-04-09	2664.55	2015-08-06	2573.30	2015-12-09	2367.65
2015-04-10	2653.75	2015-08-07	2552.05	2015-12-10	2384.55
2015-04-13	2655.50	2015-08-10	2572.10	2015-12-11	2386.00
2015-04-15	2627.75	2015-08-11	2604.20	2015-12-14	2378.45
2015-04-16	2582.80	2015-08-12	2671.70	2015-12-15	2375.85
2015-04-17	2474.85	2015-08-13	2635.60	2015-12-16	2408.40
2015-04-20	2428.80	2015-08-14	2703.20	2015-12-17	2439.75
2015-04-21	2427.05	2015-08-17	2684.75	2015-12-18	2416.80
2015-04-22	2445.10	2015-08-18	2740.90	2015-12-21	2446.40
2015-04-23	2457.25	2015-08-19	2733.60	2015-12-22	2405.05
2015-04-24	2493.05	2015-08-20	2686.85	2015-12-23	2425.80
2015-04-27	2500.20	2015-08-21	2676.15	2015-12-24	2434.25
2015-04-28	2495.00	2015-08-24	2578.65	2015-12-28	2462.70
2015-04-29	2487.30	2015-08-25	2567.15	2015-12-29	2455.80
2015-04-30	2463.70	2015-08-26	2576.75	2015-12-30	2418.30
2015-05-04	2481.55	2015-08-27	2574.65	2015-12-31	2436.85
2015-05-05	2517.85	2015-08-28	2570.85		

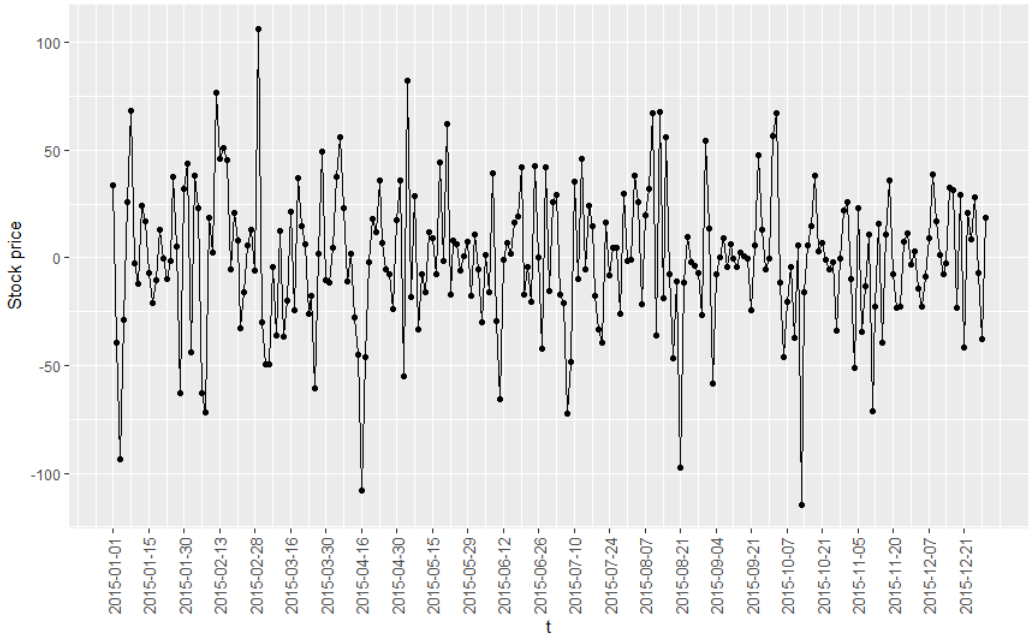
На слици 6.6 дат је графички приказ посматране временске серије. Може се закључити да кретање серије није стационарно. Стога процедуру изградње модела започињемо поступком диференцирања, како бисмо добили серију која је стационарна. На слици 6.7 приказана је диференцирана серија. Видимо да постоје опсервације чије вредности значајније одступају од кретања серије. Формално тестирање присуства нестандардних опсервација вршимо применом итеративног поступка описаног у одељку 4.3. У првом кораку, не узимајући у обзир нестандардне опсервације, користећи функцију auto.arima имплементирану у R-у, долазимо до следећег ARIMA(2,0,2) модела:

$$Y_t - 1.3085Y_{t-1} + 0.3942Y_{t-2} = e_t + 1.2647e_{t-1} - 0.2898e_{t-2} \tag{6.2}$$

Овај модел приказан је на слици 6.8. У другом кораку формирамо тест статистике 4.41. Поступак изложен у одељку 4.3 имплементиран је у R-у, у оквиру функција $detectAO()$ и $detectIO()$, којима се као аргумент прослеђује модел 6.2. Резултати су приказани у табели 6.4.

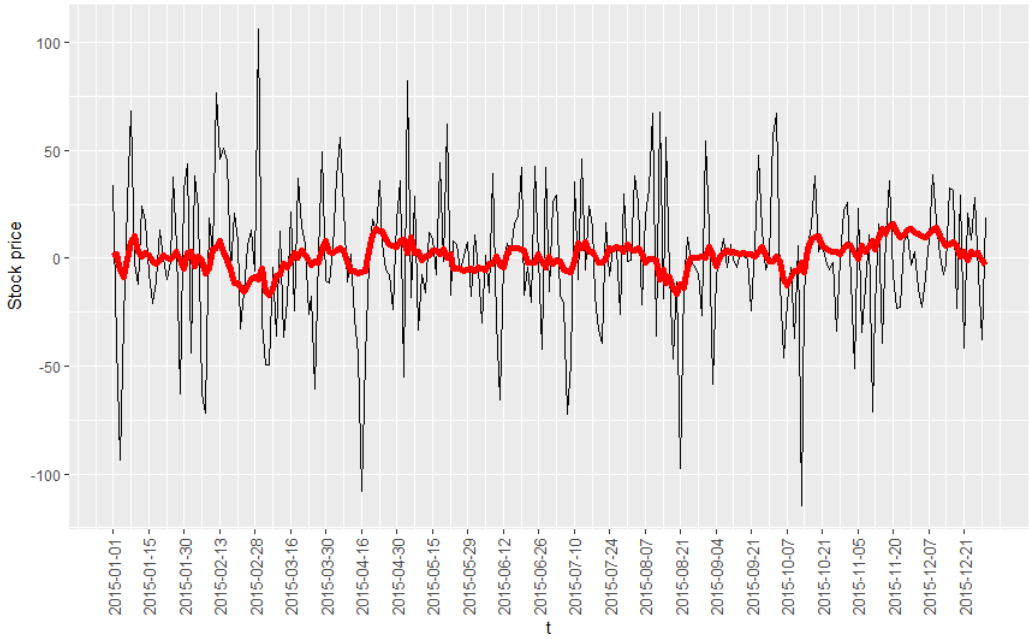


Слика 6.6: График временске серије вредности акција



Слика 6.7: График диференциране серије вредности акција

Видимо да је по апсолутној вредности највећа тест статистика $\hat{\lambda}_{A,195}$, те управо овај резултат узимамо за вредност тест статистике 4.42. Закључујемо да је опсервација под бројем 195, која одговара 13. октобру 2015. адитивни



Слика 6.8: График серије и модела 6.2

Табела 6.4: Тест статистике

Тип аутлајера	Опсервација	Тест статистика λ
IO	42	3.902825
IO	195	-3.785184
AO	195	-4.110865

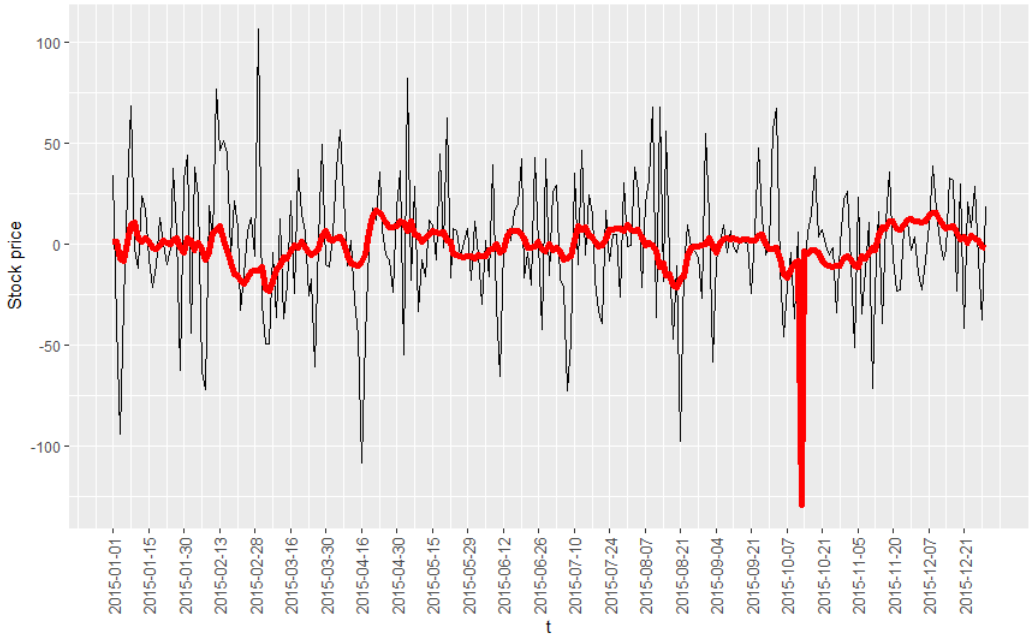
аутлајер. Имајући у виду ову чињеницу, конструишемо нови модел, надограђујући претходни додавањем импулсне променљиве у наведеној временској тачки, што се у R-у постиже коришћењем функције $arimax()$:

```
pulse = rep(0, length(t))
pulse[match("2015-10-13", t)] = 1
arimax(diferencirane_cene_akcija,
       order = c(2,0,2), xtransf = data.frame(pulse),
       transfer = list(c(0,0)))
```

Оваква модификација модела почива на природи адитивног аутлајера чије се дејство испољава само у тренутку у коме се тај аутлајер појављује. Добијени модел је облика:

$$Y_t - 1.3010Y_{t-1} + 0.3935Y_{t-2} = 0.0382 + e_t + 1.2961e_{t-1} - 0.2961e_{t-2} - 124.6114I_t \tag{6.3}$$

Графички приказ овог модела са укљученим адитивним аутлајером дат је на слици 6.9. У трећем кораку, понављамо поступак са новим моделом, који



.

Слика 6.9: График серије и модела 6.3

као аргумент прослеђујемо функцијама $detectAO()$ и $detectIO()$. У табели 6.5

Табела 6.5: Тест статистике

Тип аутлајера	Опсервација	Тест статистика λ
IO	42	4.113748
AO	No AO detected	

приказани су резултати позива наведених функција. Видимо да није детектован ни један адитивни аутлајер, те можемо закључити да смо их досадашњим поступком успешно моделовали. Поред тога, пронађен је један иновативни аутлајер који одговара опсервацији под редним бројем 42, што кореспондира са 2. мартом 2015. Како бисмо у наш модел укључили наведени аутлајер, поново користимо $arimax()$ функцију у R-у, која за ту сврху има аргумент io чија вредност се поставља на редни број временске тачке у којој се иновативни аутлајер јавља:

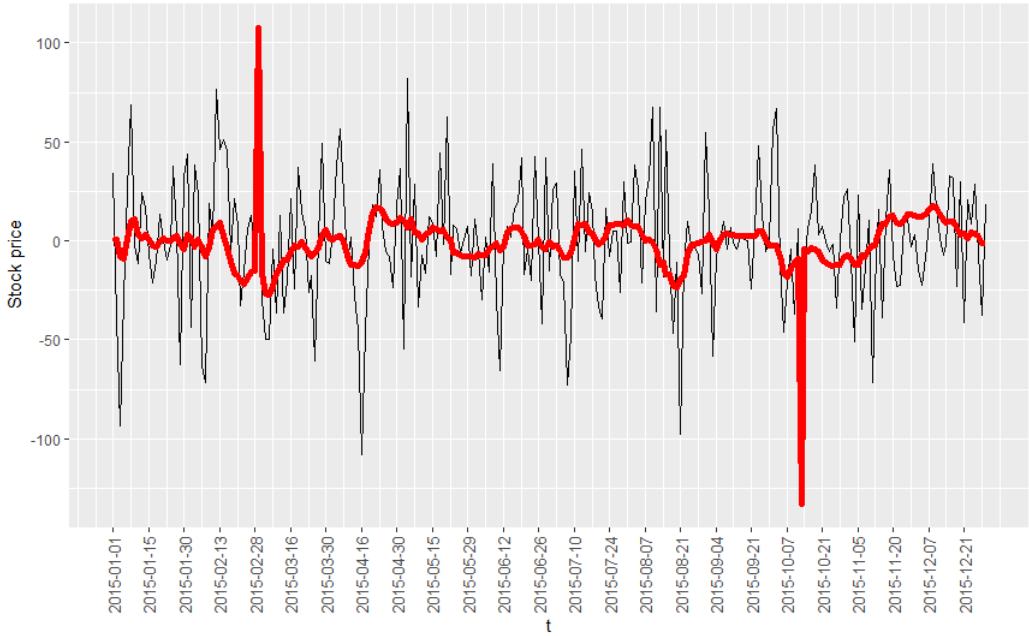
```
pulse = rep(0, length(t))
pulse[match("2015-10-13", t)] = 1
```

```
arimax(diferencirane_cene_akcija,
       order = c(2,0,2), xtransf = data.frame(pulse),
       transfer = list(c(0,0)), io = list(c(42)))
```

Модел до ког смо дошли описан је једначином:

$$Y_t - 1.3317Y_{t-1} + 0.43Y_{t-2} = -0.4038 + e_t + 1.3324e_{t-1} - 0.3331e_{t-2} \\ -128.088I_t + \nu(L)123.8706 \qquad (6.4)$$

Графички приказ модела са детектованим адитивним и иновативним аутлајерима дат је на слици 6.10. У складу са изложеном теоријом, како бисмо



Слика 6.10: График серије и модела 6.4

привели анализу крају, понављамо поступак детектовања нестандардних опсервација, користећи модел 6.4. Приказ резултата дајемо у табели 6.6. Ни једна нестандарнда опсервација није пронађена, те за коначан модел серије узимамо 6.4.

Табела 6.6: Тест статистике

Тип аутлајера	Опсервација	Тест статистика λ
IO	No IO detected	
AO	No AO detected	

Глава 7

Закључак

У овом раду, обрађене су методе за моделовање структурног лома у временској серији. Наведеној теми приступљено је у контексту ARMA модела стационарних временских серија, као и општег регресионог модела. У првом делу рада дат је преглед ARMA, од специјалних (AR и MA модели) ка општем случају. Дефинисани су појмови интервенције и нестандардне опсервације, као узрока појаве структурног лома. Описано је неколико примера појаве интервенције у реалном животу, који илуструју значај бављења овом темом приликом анализирања важних економских временских серија.
Посебна пажња посвећена је разлици у приступу у зависности од доступности информације о тренутку лома. Приказана је употреба вештачких променљивих у надоградњи ARIMA модела код познатог тренутка интервенције, као и адитивни и иновативни модел који се могу користити у ситуацији када је тренутак појаве структурног лома непознат. Додатно, изложена је итеративна процедура, којом се у пракси долази до наведених модела. На крају овог дела обрађено је моделовање лома у општијем регресионом моделу.
Поред самог моделовања, обрађена је и тема детекције присуства структурног лома. У оквиру ове теме, наведене су методе и статистички тестови којима се потврђује појава структурног лома у серији. Изложен је значај коришћења неформалних метода, као што је анализа графика серије, и примена доменског знања о подацима који су предмет истраживања. Применом ових метода, пре самих статистичких тестова, могуће је значајно олакшати поступак детектовања структурног лома. Тестови који су приказани у раду су Жак-Бера тест, Чоу тест, и CUSUM тест, који се заснивају на различитим приступима откривању одступања серије од њеног уобичајеног кретања.

Обрађен је и секвенцијални тест за детектовање нових структурних ломова између већ постојећих. Приказани су и актуелни правци истраживања који се тичу тестирања присуства лома у случају хетероскедастичне серије.
Посебан део рада посвећен је односу између структурног лома и јединичног корена у временској серији, односно тестирању присуства јединичног корена у случају серије са ломом. У том случају, потребно је модификовати стандардне тестове као што је проширени Дики-Фулеров тест.
На крају рада, дата су два практична примера анализе структурног лома, при чему први илуструје ситуацију познатог тренутка интервенције, док други описује поступак у слулају када је наведени тренутак непознат. Приказано је коришћење уграђених R функција које се, приликом анализе структурног лома, примењују за моделовање (*arimax*), односно детектовање (*detectAO*, *detectIO*). Из ових примера, може се сагледати значај укључивања структурног лома у модел временске серије, у контексту што тачнијег описа кретања серије. Тиме се побољшавају перформансе модела приликом прогнозирања даљег кретања серије, што је једна од најважнијих циљева моделовања података. Вреди нагласити да ситуација изазвана пандемијом корона вируса, односно њен утицај на кретање многих значајних временских серија, додатно подиже значај ове теме, и даје подстрек за даље истраживање на моделима који не припадају класи ARIMA.

Библиографија

[1] Public image of George W. Bush. on-line at: `https://en.wikipedia.org/wiki/Public_image_of_George_W._Bush`.

[2] Donald W.K. Andrews and Werner Ploberger. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62:1383–1414, 1994.

[3] Jushan Bai and Pierre Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–78, 1998.

[4] G. E. P. Box and G. C. Tiao. Intervention Analysis with Applications to Economic and Environmental Problems. *Journal of the American Statistical Association*, 70(349):70–79, 1975.

[5] George E. P. Box, Gwilym M. Jenkins, and Gregory C. Reinsel. *Time series analysis: forecasting and control, Third edition.* Prentice Hall, Inc,, 1994.

[6] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time series analysis: forecasting and control, Fifth edition.* Wiley, 2015.

[7] Michael P. Clements and David F. Hendry. *Forecasting with breaks, Handbook of Economic Forecasting, Volume 1.* North Holland, 2006.

[8] A. S. Fox. Outliers in time series. *Journal of Royal Statistical Society*, 43:350–363, 1972.

[9] Zorica Mladenović i Aleksandra Nojković. *Primenjena analiza vremenskih serija.* Ekonomski fakultet, Beograd, 2015.

[10] Zlatko J. Kovačić. *Analiza vremenskih serija.* Ekonomski fakultet, Beograd, 1995.

[11] Eiji Kurozumi. Confidence sets for the date of a structural change at the end of the sample. *Journal of Time Series Analysis*, (39):850–862, 2018.

[12] Pierre Perron. Dealing with structural breaks. *Palgrave handbook of econometrics, volume 1: econometric theory*, 05 2005.

[13] Pierre Perron and Yohei Yamamoto. Structural change tests under heteroskedasticity: Joint estimation versus two-steps methods. *Journal of Time Series Analysis*, (43):389–411, 2022.

[14] J.M. Evans R. L. Brown, J. Durbin. Techniques for Testing the Constancy of Regression Relationships over Time. *Journal of Royal Statistical Society*, 37(2):149–192, 1975.

[15] Ruey S. Tsay. Outliers, level shifts, and variance changes in time series. *Journal of Forecasting*, 7(1):1–20, 1988.

[16] Jurgen Wolters and Uwe Hassler. Unit root testing. *Modern Econometric Analysis*, pages 49–51, 2006.

Биографија аутора

Дамјан Петровић рођен је у Београду, 29. јула 1995. године. Основну школу „Ослободиоци Београда” завршио је 2010. године, а Пету београдску гимназију, природно-математички смер, 2014. године. Исте године, по завршетку средње школе, уписује основне академске студије на Математичком факултету Универзитета у Београду, на смеру Статистика, финансијска и актуарска математика, које завршава 2019. године. Након тога уписује мастер академске студије на истом смеру.