



HNB

EUROSUSTAV

ODABRANE TEME
PRIMIJEJENE
EKONOMETRIJE

Uvod u analizu vremenskih nizova

Tihana Škrinjarčić

Zagreb, 2023.

HRVATSKA NARODNA BANKA
EUROSUSTAV



HNB

EUROSUSTAV

ODABRANE TEME
PRIMIJEJENE
EKONOMETRIJE

Uvod u analizu vremenskih nizova

Tihana Škrinjaric

Zagreb, 2023.

HRVATSKA NARODNA BANKA

EUROSUSTAV

Sadržaj ne odražava nužno stavove Hrvatske narodne banke.
Hrvatska narodna banka nije odgovorna za eventualne greške
u tekstu.



Komentare, prijedloge i uočene greške
u kodovima molimo javiti na
publikacije@hnb.hr

Hvala!

Predgovor izdavača

“Odabrane teme primijenjene ekonometrije: Uvod u analizu vremenskih nizova” dio je serije edukativnih publikacija Hrvatske narodne banke. Riječ je o praktičnom pregledu osnovnih metoda ekonometrijske analize koji kombinira istraživačke probleme s ekonometrijskom teorijom, a napisan je tako da čitatelja vodi korak po korak do rješenja. Prvi dio "Odabrane teme primijenjene ekonometrije: Linearni regresijski modeli" objavljen je u ožujku 2023.

“Odabrane teme” praktičan su uvod za početnike koji se tek spremaju otisnuti u empirijska istraživanja, kao što su polaznici ekonomskih ili drugih, uglavnom društvenih studija, koji primjenjuju metode ekonometrijske analize. Osim toj primarnoj ciljnoj skupini, knjiga će dobro poslužiti kao konzultativni priručnik iskusnijim analitičarima u rješavanju novih konkretnih problema s kojima se susreću, kao i zainteresiranima koji će publikaciju samo prelistati kako bi dobili opći pregled izabranih metoda ekonometrijske analize. Nadamo se da će knjiga doprijeti i do akademskih korisnika koji nisu ekonomisti, primjerice iz STEM područja, i zaintrigirati ih za teme koje istražuju ekonomisti.

Edukativne publikacije na hrvatskom jeziku koje obrađuju ekonometrijsku analizu, uključujući konkretne primjere analize domaćih makroekonomskih ili mikroekonomskih podataka kao i popratnu podršku u obliku računalnih kôdova i riješenih problema, nisu široko dostupne. Nadamo se da će stavljanje ove publikacije na slobodno raspolaganje istraživačima, ekonomistima, studentima i drugoj zainteresiranoj javnosti unaprijediti razumijevanje metoda kojima se služe ekonomisti i potaknuti buduće zanimanje za ekonomska istraživanja.

Predgovor recenzenta

Autorica i u drugoj knjizi zadržava metodološku rigoroznost, spretno obrađujući kompleksne metodološke pristupe, donoseći istodobno i konkretnu praktičnu primjenu razmatranih metoda i modela. Na sustavan način uvodi čitatelja u sve složenije metode analize vremenskih nizova, objašnjavajući i primjerima ilustrirajući složene koncepte poput modela vektorske autoregresije, kointegracije ili ARIMA modela.

Knjiga može biti korisna ponajprije studentima ekonomskih ili financijskih usmjerenja, ali i metodološki relativno neiskusnim članovima akademske zajednice, istraživačima i središnjim bankarima. Zbog sve dostupnijih i sve frekventnijih podataka (*big data*) u svim segmentima gospodarstva i društva općenito, analiza vremenskih nizova postaje sve važnijom i u različitim sferama poslovne ekonomije, gdje se tim tipom analize sve donedavno nije ni koristilo.

Knjiga može biti posebno korisna zbog činjenice da autorica primjenu metoda analize vremenskih nizova ilustrira upotrebom programske podrške slobodnog dohvata R. Riječ je o besplatnom programu koji se razvija i unapređuje u realnom vremenu, što ga čini zanimljivim, a često i neizostavnim alatom i u poslovnom i u središnjem bankarstvu kao i u akademskoj zajednici.

izv. prof. dr. sc. Petar Sorić

Predgovor autorice

Knjiga "Odabrane teme primijenjene ekonometrije: Uvod u analizu vremenskih nizova" nastavak je niza odabranih tema koje omogućuju samostalno svladavanje osnova ekonometrijske analize. Kao i u prvome dijelu, u kojemu se pojašnjava linearni regresijski model, i ovdje su postavljeni temelji ekonometrijske analize i primjene u makroekonomiji.

Knjigu sam podijelila u tri cjeline. Prvi dio **Uvod u univarijatnu analizu vremenskih nizova** donosi osnove desezoniranja vremenskih nizova te odnos stacionarnih i nestacionarnih nizova, njihovih karakteristika i načina modeliranja. Iduća dva poglavlja razmatraju više nizova odjednom. U prvome, **Vektorski autoregresijski modeli**, pojašnjavaju se stacionarni nizovi i procjenjuju modeli u kojima je više endogenih varijabli, a u posljednjem, **Kointegracija u vektorskim modelima**, nestacionarni nizovi i modeli s više takvih varijabli. Uz materijale pripremila sam podatke i R-skriptu dostupnu na internetskim stranicama HNB-a, kako bi čitatelj lakše pratio materiju i usporedno sâm krenuo u modeliranje.

Ono što ovaj materijal donosi kao dodatnu vrijednost jest to što se uz detaljna teorijska pojašnjenja razmatraju brojni primjeri kako bi čitatelj mogao lakše povezati teoriju s onime što to nosi za posljedicu u empirijskom istraživanju i primjeni. Primjeri su pisani i predočeni u programskoj podršci R, a materija je pripremljena tako da uvelike olakšava samostalno učenje.

Materijal se preporučuje svima onima koji žele započeti vlastite ekonometrijske analize s posebnim osvrtom na vremenske nizove, neovisno o tome koliko poznaju ekonometriju ili primjenu ekonometrije u programskoj podršci R.



Kodovi dostupni na:

<https://www.hnb.hr/analize-i-publikacije/ostale-publikacije>



O autorici

Tihana Škrinjarić savjetnica je u Direkciji za makrobonitetnu politiku i financijsku stabilnost u Hrvatskoj narodnoj banci. Bavi se temama povezanim s kvantitativnom podrškom u makrobonitetnoj politici, analizom cikličkih rizika, ocjenom karaktera makrobonitetne politike i kalibracijom njezinih instrumenata. Prije dolaska u HNB 2021. godine, Tihana je bila docentica na Ekonomskom fakultetu u Zagrebu na Katedri za matematiku, gdje je držala kvantitativne kolegije poput ekonometrije i matematike. Doktorat iz ekonomije obranila je 2018. godine, na Ekonomskom fakultetu – Zagreb, s temom ekonometrijskih modela promjene režima u menadžmentu portfelja. U razdoblju rada na fakultetu Tihanini interesi istraživanja bili su financijska tržišta i financijska ekonometrija te primjena kvantitativnih modela i metoda u području financija. Tihana ima više od 150 objavljenih radova, poglavlja u knjigama i istraživanja iz područja financija i makrobonitetne politike, bila je recenzentica nekoliko desetaka časopisa te je dobitnica više nagrada za znanstvene članke iz tih područja. Krajem 2022. godine, Tihana je rangirana u top 12,5% europskih autora iz područja ekonomije u bazi Ideas RePec.

Link na Google Scholar:

<https://scholar.google.com/citations?user=C5yapgkAAAAJ&hl=hr&oi=ao>



SADRŽAJ

SADRŽAJ	i
OZNAKE	iv
1. UVOD U UNIVARIJATNU ANALIZU VREMENSKIH NIZOVA.....	1
1.1. Temeljni pojmovi	1
1.2. Priprema vremenskih nizova za regresijske analize	3
1.3. Binarne varijable za sezonalnost vremenskih nizova	4
1.4. Primjer	14
1.5. Pitanja za ponavljanje	29
1.6. Odabrane metode desezoniranja vremenskih nizova.....	33
1.6.1. Metoda pomičnih prosjeka	33
1.6.2. Klasično desezoniranje (dekomponiranje) vremenskog niza.....	34
1.6.3. Primjer.....	31
1.7. Pitanja za ponavljanje	34
1.8. Model višestruke linearne regresije s vremenskim nizovima.....	42
1.9. Stacionarni vremenski nizovi	43
1.9.1. Definiranje stacionarnosti	43
1.9.1.1. Striktne stacionarnosti.....	43
1.9.1.2. Slaba stacionarnost.....	43
1.9.1.3. Autokorelacijska funkcija	44
1.9.1.4. Parcijalna autokorelacijska funkcija	47
1.9.1.5. Testiranje značajnosti koeficijenta autokorelacije	49
1.9.2. Čisti slučajni proces	51
1.9.3. Autoregresijski proces, AR	54
1.9.3.1. Definiranje autoregresijskog procesa.....	54
1.9.3.2. Svojstva autoregresijskih procesa	55
1.9.4. Proces pomičnih prosjeka, MA	67
1.9.4.1. Definiranje procesa pomičnih prosjeka	67
1.9.4.2. Svojstva procesa pomičnih prosjeka	68
1.9.5. Woldov teorem dekompozicije	75
1.9.6. Invertibilnost stohastičkog procesa	76
1.9.7. Funkcija impulsnog odziva ARMA procesa	77

1.9.8.	Autoregresijski proces pomičnih prosjeka, ARMA	79
1.9.9.	Primjer	83
1.9.10.	Pitanja za ponavljanje	98
1.10.	Nestacionarni vremenski nizovi	105
1.10.1.	Deterministički trend	106
1.10.2.	Stohastički trend	109
1.10.3.	Pogrešno uklanjanje trenda	112
1.10.3.1.	Detrendiranje stohastičkog trenda	112
1.10.3.2.	Diferenciranje determinističkog trenda	112
1.10.4.	Model slučajnog pomaka	114
1.10.4.1.	Model slučajnog pomaka bez konstante	114
1.10.4.2.	Model slučajnog pomaka s konstantom	115
1.10.4.3.	Hipoteza efikasnog tržišta	117
1.10.5.	Primjer	119
1.10.6.	Napomena o problemu prividne regresije	123
1.10.7.	Testovi jediničnog korijena	125
1.10.7.1.	Dickey-Fullerov test i prošireni Dickey-Fullerov test	125
1.10.8.	Pitanja za ponavljanje	143
1.11.	Autoregresijski model distribuiranih pomaka, ARDL	144
1.12.	Analiza uzročnosti	160
1.13.	Kointegracija	164
1.14.	Model korekcije pogreške	170
1.14.1.	Definiranje modela korekcije pogreške	170
1.14.2.	Prednosti i nedostaci pristupa modela korekcije pogreške	173
1.14.3.	Napomene vezane uz model korekcije pogreške	173
1.15.	Zaključno o (ne)stacionarnosti varijabli i odgovarajućim modelima	174
1.16.	Primjer	174
1.17.	Informacijski kriteriji	183
1.18.	Pitanja za ponavljanje	185
2.	VEKTORSKI AUTOREGRESIJSKI MODELI	197
2.1.	Temeljni pojmovi	197
2.1.1.	Oznake	197
2.1.2.	Strukturni i reducirani oblik modela	198

2.2.	Stabilnost i stacionarnost modela	201
2.2.1.	Stabilnost i stacionarnost VAR(1) modela.....	201
2.2.2.	Stacionarnost multivarijatnog procesa y_t	203
2.2.3.	Momenti VAR(1) modela	204
2.3.	Nekoliko napomena o VAR modelima	208
2.4.	Analiza uzročnosti	209
2.4.1.	Grangerov test uzročnosti.....	209
2.4.2.	Testovi blok-egzogenosti	212
2.5.	Inovacijska analiza	212
2.5.1.	Funkcija impulsnog odziva	213
2.5.1.1.	Definiranje funkcije impulsnog odziva.....	213
2.5.1.2.	Cholesky dekompozicija.....	214
2.5.2.	Dekompozicija varijance.....	222
2.6.	Kriteriji odabira adekvatnog pomaka p	225
2.6.1.	Test omjera vjerodostojnosti	226
2.6.2.	Informacijski kriteriji	227
2.7.	Multivarijatni testovi vezani uz inovacijske procese.....	229
2.7.1.	Multivarijatni testovi autokorelacije	229
2.7.2.	Multivarijatni test heteroskedastičnosti.....	232
2.7.3.	Multivarijatni test normalnosti	233
2.8.	Primjer	235
2.9.	Pitanja za ponavljanje	243
3.	KOINTEGRACIJA U VEKTORSKIM MODELIMA	255
3.1.	Temeljni pojmovi	255
3.2.	Testiranje kointegriranosti u VEC modelu.....	256
3.3.	Deterministički regresori u VAR i VEC modelu.....	266
3.4.	Primjer	269
3.5.	Pitanja za ponavljanje	276
	LITERATURA.....	281
	POPIS POJMOVA	283

OZNAKE

\mathbf{x} – vektor stupac

\mathbf{x}' – vektor redak, radi se o transponatu vektora stupca \mathbf{x}

x – varijabla

\mathbf{X} – matrica

$E(\cdot)$ – operator očekivanja. Čitam: „očekivanje od...“

$E(y | x)$ – operator uvjetnog očekivanja. Čitam: „očekivana vrijednost y -a s obzirom na dani x “

$P(\cdot)$ – vjerojatnost nastupa nekog događaja. Čitam: „vjerojatnost nastupa događaja ... je „

$P(y | x)$ – uvjetna vjerojatnost nastupa nekog događaja. Čitam: „vjerojatnost da nastupi y uz dane vrijednosti x iznosi... “

$\frac{dy(x)}{dx}$ – prva derivacija funkcije y , slučaj jedne nezavisne varijable

$\frac{\partial y(x_1, \dots, x_i, \dots, x_n)}{\partial x_i}$ – parcijalna derivacija prvog reda funkcije y po varijabli x_i , slučaj više

nezavisnih varijabli

0 – broj nula

$\mathbf{0}$ – nul-vektor

I – jedinična matrica

y_i – i -ta opservacija varijable Y , empirijska vrijednost

\hat{y}_i – i -ta procijenjena vrijednost varijable Y

\exists – postoji, npr. $\exists x$ – čitam „postoji (barem jedna) varijabla x “

\neq – različito je

\sim – slijedi, npr. $x \sim N$ – čitam „varijabla x slijedi normalnu distribuciju“

\rightarrow – teži prema, npr. $x \rightarrow y$ – čitam „vrijednost x teži prema vrijednosti y “

Δ – delta, prva diferencija, promjena varijable, $\Delta y_t = y_t - y_{t-1}$

Δ^2 – druga diferencija, promjena promjene varijable

$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = y_t - y_{t-1} - (y_{t-1} - y_{t-2})$



1.

**UVOD U UNIVARIJATNU
ANALIZU VREMENSKIH
NIZOVA**



AWN

1. UVOD U UNIVARIJATNU ANALIZU VREMENSKIH NIZOVA

1.1. Temeljni pojmovi

U ekonomskim analizama najčešće se susrećemo s vremenskim nizovima, gdje se temeljem podataka iz prošlosti nastoji opisati jedna ili više pojava istovremeno, te prognozirati buduće vrijednosti razmatrane pojave. Većina ekonomskih podataka prikuplja se kroz vrijeme i tako dobivamo **vremenske nizove**. Vremenski niz je skup kronološki uređenih vrijednosti varijable koja predočuje neku pojavu u vremenu. Ovdje se radi o istome entitetu (čovjek, država, dionica, ...) čija se opažanja bilježe na određene datume i odnose se na više vremenskih trenutaka ili razdoblja. Primjerice, ako se razmatra BDP Hrvatske u razdoblju od 2000. do 2020. godine, pri čemu se prikupljaju tromjesečni podaci, radi se o vremenskom nizu.

Vremenski nizovi se mogu razmatrati temeljem jedne jednadžbe ili pak temeljem više jednadžbi – vektorskih modela. Ako se najprije fokus vrši na jednu jednadžbu, osnovni model vremenskih nizova koji se može razmatrati je onaj u kojemu se sam niz rastavlja (dekomponira) na **trend** (T_t), **sezonsku** (S_t), **cikličku** (C_t) i **slučajnu** (ε_t) komponentu:

$$y_t = T_t + S_t + C_t + \varepsilon_t, \quad (1.1)$$

pri čemu je model u (1.1) **aditivni** oblik, koji se koristi u analizi kada sezonska komponenta ima ujednačene amplitude tijekom vremena. S druge strane, ako sezonska komponenta ima neujednačene amplitude, koje se proporcionalno povećavaju ili smanjuju s vremenom, razmatra se **multiplikativni** oblik modela:

$$y_t = T_t \cdot I_{S_t} \cdot I_{C_t} \cdot \varepsilon_t, \quad (1.2)$$

gdje su I_{S_t} i I_{C_t} indeksi koji se odnose na sezonsku i cikličku komponentu (nepomnoženi sa 100, tj. osnovica je 1). Trend komponenta upućuje na osnovni tijek razvoja neke pojave u vremenu, stoga se pojava može predočiti pomoću nekog funkcionalnog oblika, gdje je trend nezavisna varijabla. Sezonska komponenta, primjerice, predočuje periodični utjecaj klimatskog faktora i drugih učinaka na samu pojavu, pri čemu se približno jednako obnavlja pojava unutar jedne godine. Ciklička komponenta predočava utjecaj različitih čimbenika na neku pojavu, pri čemu se pojava obnavlja približno podjednako unutar dvije ili više godina. Slučajna (ili stohastička, iregularna, rezidualna) komponenta predočava slučajnu varijablu, jer na vremenske nizove utječu i nesistematski faktori koje nije moguće predviđati. Trend i ciklička komponenta se često zajedno promatraju, pa se model (1.1), odnosno (1.2) može zapisati na sljedeći način:

$$y_t = (TC)_t + S_t + \varepsilon_t, \quad (1.3)$$

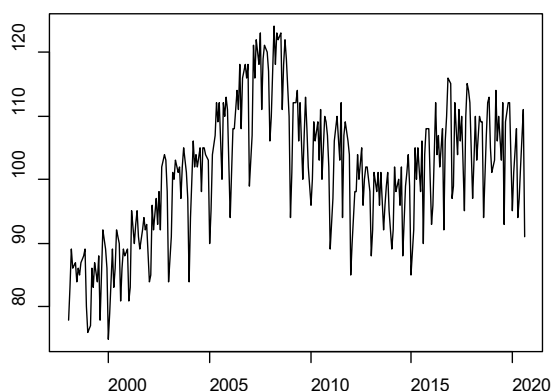
odnosno

$$y_t = TI_{C_t} \cdot I_{S_t} \cdot \varepsilon_t. \quad (1.4)$$

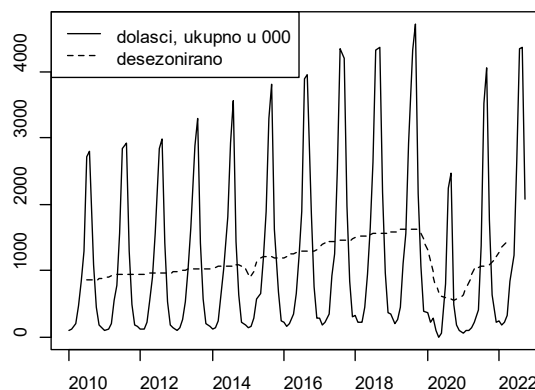
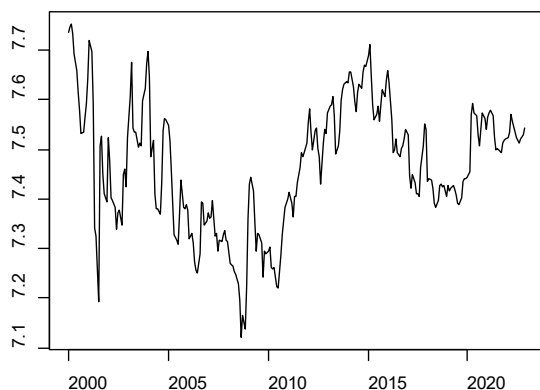
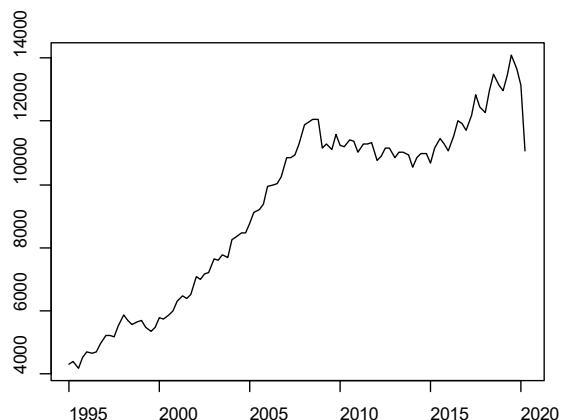
Slika 1.1. predočava četiri vremenska niza: panel (a) se odnosi na kretanje indeksa industrijske proizvodnje u Hrvatskoj, od siječnja 1998. do kolovoza 2020. godine (2015=100); panel (b) odnosi se na desezoniranu vrijednost BDP-a Hrvatske, od 1. tromjesečja 1995. do 2. tromjesečja 2020. godine (tekuće cijene, milijuni eura). Ono što se može uočiti na panelu (a) jest da postoji učinak sezone u kretanju indeksa industrijske proizvodnje, dok je taj učinak sezone isključen u kretanju BDP-a na panelu (b). Nadalje, panel (c) prikazuje kretanje prosječnog mjesečnog tečaja kune naspram eura, gdje je također vidljiv učinak sezone, s obzirom na pritiske na tečaj

tijekom ljetnih mjeseci. Iako je sezonski učinak nešto slabiji u odnosu na onaj na panelu (a), za dodatne analize u kojima se koristi ovaj tečaj, potrebno ga je izbaciti. Na panelu (d) je prikazana dinamika dolazaka turista u Hrvatsku, u originalnom obliku u kojem je izuzetno prisutan učinak sezone, kao i njegova desezonirana serija. Uočava se velika razlika između vrhova i dna originalnog i desezoniranog niza.

Slika 1.1. (a) Kretanje indeksa industrijske proizvodnje u Hrvatskoj



(b) Kretanje desezoniranog BDP-a Hrvatske

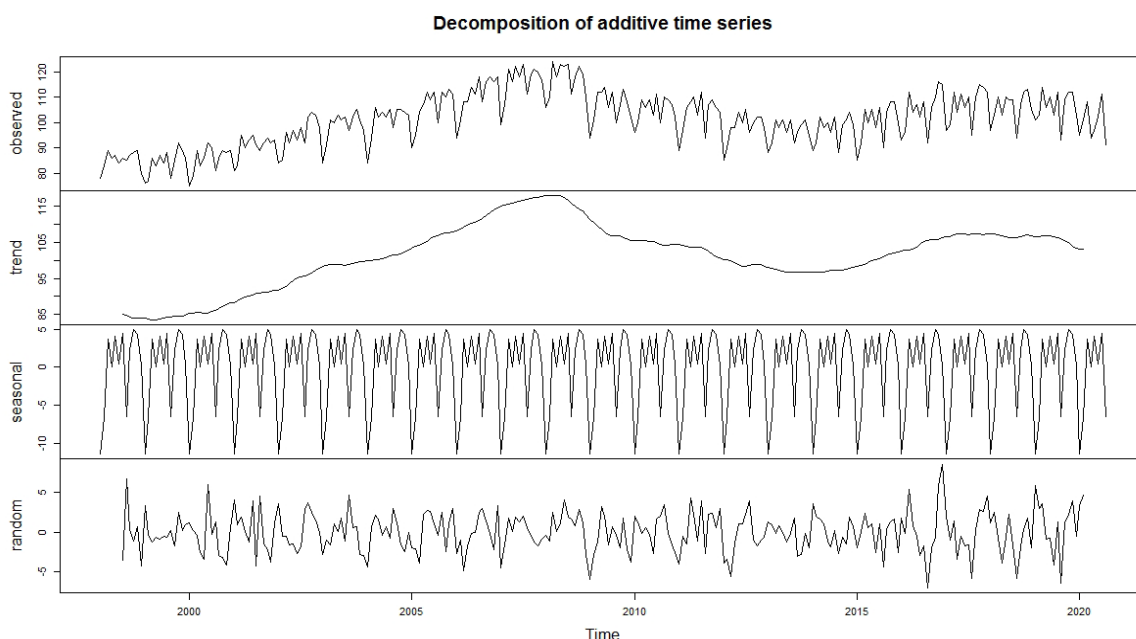


(c) Kretanje tečaja EUR-HRK

(d) Kretanje dolazaka turista u Hrvatsku

Ako bismo htjeli niz indeksa industrijske proizvodnje dekomponirati, tj. raščlaniti na komponente od kojih se sastoji, to se može predočiti kao što je učinjeno na slici 1.2. Uočava se da postoji sezonska komponenta s ujednačenim obnavljanjem svake godine.

Slika 1.2. Raščlanjivanje indeksa industrijske proizvodnje na trend i cikličku, sezonsku i slučajnu komponentu



Napomena: observed, trend, seasonal i random se odnose na originalni vremenski niz, trend i cikličku komponentu, sezonsku komponentu te slučajnu komponentu.

Vremenske nizove moguće je analizirati **metodama analize u domeni kutnih frekvencija i u domeni vremena**. Modeli koji se analiziraju u domeni kutnih frekvencija se nazivaju spektralni modeli, dok su modeli u domeni vremena stohastički modeli, koji opisuju (ne)stacionarne procese. Stacionarni procesi obrađuju se u odjeljku 1.9.

1.2. Priprema vremenskih nizova za regresijske analize

U praksi se često događa da je prikupljene podatke prije same analize potrebno urediti. Ovdje se misli na dodatne transformacije koje se mogu vršiti nad podacima poput logaritmiranja, diferenciranja, desezoniranja, deflacioniranja, itd. Ako ekonomska teorija pretpostavlja vezu između stopa rasta, a podaci su prikupljeni u razinama, potrebno je najprije izračunati stope rasta, te potom vršiti analizu. **Logaritmiranje** podataka je jedna od najčešćih transformacija nad podacima u empirijskim analizama, iz dva razloga. Prvi je što se stope rasta često računaju temeljem logaritmiranih vrijednosti varijabli, a drugi to što ova transformacija smanjuje varijancu novoga niza (važnost ovoga obrađena je u prvom dijelu ove serije u poglavlju 4.3.).

Kako se često za varijable razmatraju **stope rasta**, one se mogu računati na diskretni i kontinuirani način. U diskretnome slučaju stopa rasta računa se formulom:

$$y_t = \frac{Y_t - Y_{t-1}}{Y_{t-1}} \cdot 100\%, \quad (1.5)$$

dok se u kontinuiranome slučaju računa formulom¹:

¹ Pojednostosti o razlikama te kada se dobiju približno jednaki rezultati dani su u Aljinović, Marasović i Šego (2011:118-119).

$$y_t = \ln\left(\frac{Y_t}{Y_{t-1}}\right) \cdot 100\%, \quad (1.6)$$

gdje \ln predstavlja logaritam s bazom e (prirodni logaritam). Umjesto baze e može se koristiti bilo koja druga baza, no u stranoj literaturi najčešće se koristi \ln . Dodatno, često se razmatra i stopa rasta koja se računa u odnosu na isti mjesec ili tromjesečje iz prethodne godine:

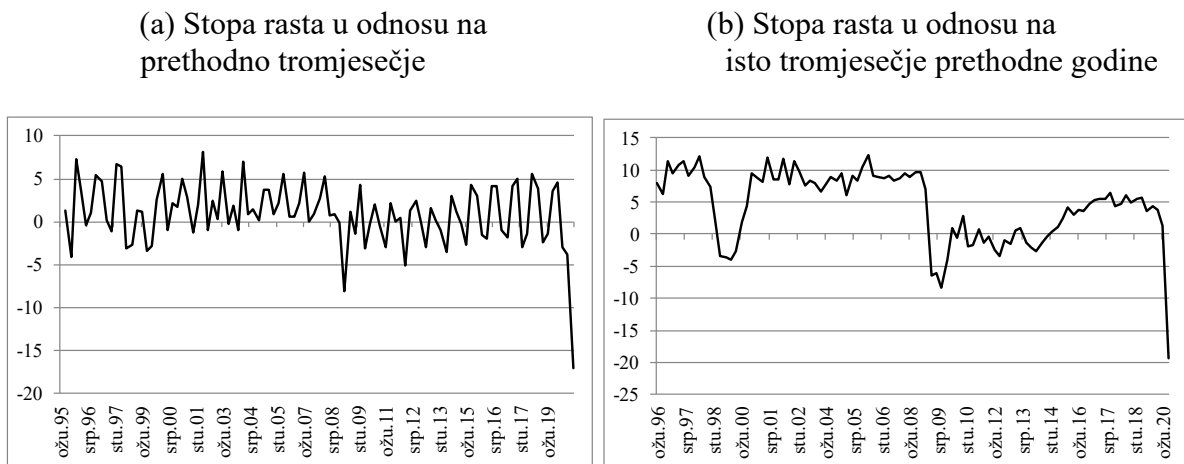
$$y_t = \ln\left(\frac{Y_t}{Y_{t-12}}\right) \cdot 100\%, \quad (1.7)$$

odnosno

$$y_t = \ln\left(\frac{Y_t}{Y_{t-4}}\right) \cdot 100\%. \quad (1.8)$$

Slika 1.3. uspoređuje stope rasta BDP-a izračunatu temeljem formule (1.6) i (1.8).

Ako se želi analizirati realni niz ili međuovisnost između realnih nizova umjesto nominalnih, potrebno je podatke **deflacionirati**. To je postupak uklanjanja utjecaja promjena cijena na promjenu vrijednosti neke pojave. U tome slučaju nominalne (originalne) vrijednosti pretvaraju se u realne vrijednosti, na način da se nominalne vrijednosti nekog vremenskog niza podijele s određenim indeksom - deflatorom. Ekonomske baze podataka poput Eurostata nude različite indekse za svrhe deflacioniranja, a često se danas mogu preuzeti gotovi podaci, nad kojima je već deflacioniranje izvršeno. **Desezoniranje** je postupak uklanjanja učinka sezonalnosti u vremenskim nizovima. Detaljnije će se obrađivati u odjeljku 1.6.



Slika 1.3. Stopa rasta BDP-a

1.3. Binarne varijable za sezonalnost vremenskih nizova

Binarne varijable mogu se koristiti u opisivanju sezonalnosti vremenskih nizova, kako bi se modelirao učinak sezonskih faktora. Mnoge ekonomske pojave su podložne utjecaju sezone, poput dolazaka i noćenja turista, indeks industrijske proizvodnje, BDP, itd. Ako se razmotri slika 1.1. (a), već je spomenuto kako je vidljiv učinak sezone u indeksu industrijske proizvodnje vezanim uz Hrvatsku.

Sezonski učinak se može u regresijski model uključiti pomoću binarnih varijabli. Naime, ako je sezonska komponenta konstantna tijekom cijelog razmatranog razdoblja, **učinci sezone** na određenu varijablu se mogu procijeniti **regresijskim pristupom**. Ako se radi o mjesečnim podacima, i želi se modelirati promjena prosječne razine pojave s obzirom na mjesec o kojem se radi, definiraju se sljedeće binarne varijable:

$$D1_t = \begin{cases} 1, & \text{za siječanj} \\ 0, & \text{inače} \end{cases}, D2_t = \begin{cases} 1, & \text{za veljaču} \\ 0, & \text{inače} \end{cases}, \dots, D11_t = \begin{cases} 1, & \text{za studeni} \\ 0, & \text{inače} \end{cases}, \quad (1.9)$$

te se procijeni sljedeći model:

$$y_t = \beta_0 + \beta_1 D1_t + \beta_2 D2_t + \dots + \beta_{11} D11_t + \varepsilon_t, \quad (1.10)$$

gdje se vrijednost procijenjene konstante interpretira kao prosječna razina pojave u prosincu tijekom promatranog razdoblja (jer se radi o slučaju kada su sve binarne varijable iz (1.9) jednake nuli i upravo je to slučaj prosinca), dok se interpretacije preostalih procijenjenih parametara vrše u odnosu na prosječnu razinu pojave u prosincu. Primjerice, **interpretacija** $\hat{\beta}_2$ glasila bi: u promatranom razdoblju, prosječna razina pojave y je u veljači za $\hat{\beta}_2$ jedinica veća/manja u odnosu na prosinac.

S obzirom da bi za prosinac prosječna razina pojave iznosila: $\hat{y}_t = \hat{\beta}_0 + 0 + 0 + \dots + 0 = \hat{\beta}_0$, dok bi za veljaču iznosila: $\hat{y}_t = \hat{\beta}_0 + 0 + \hat{\beta}_2 + \dots + 0 = \hat{\beta}_0 + \hat{\beta}_2$, upravo vrijednost $\hat{\beta}_2$ predstavlja razliku između prosječne razine pojave u prosincu i veljači.

Uočimo kako je u modelu (1.10) uključeno 11 binarnih varijabli **jer je konstanta već uključena u model**. Već je u prvome dijelu ove serije ("Odabrane teme primijenjene ekonometrije: Linearni regresijski model") obrađen problem multikolinearnosti² varijabli kada su uz konstantu bile uključene sve binarne varijable s obzirom na broj modaliteta koje mogu poprimiti. Također, ako bismo u slučaju modela (1.10) htjeli razmotriti sve binarne varijable, na način da se definira i binarna varijabla za prosinac, $D12_t = \begin{cases} 1, & \text{za prosinac} \\ 0, & \text{inače} \end{cases}$, tada bismo morali procijeniti model:

$$y_t = \beta_1 D1_t + \beta_2 D2_t + \dots + \beta_{11} D11_t + \beta_{12} D12_t + \varepsilon_t, \quad (1.11)$$

gdje uočavamo da je potrebno izostaviti konstantu. U tom slučaju će interpretacija procijenjenih parametara biti sljedeća: $\hat{\beta}_j$ bit će prosječna razina pojave \hat{y}_t u mjesecu j .

Primjer 1.1.

Učitajmo datoteku „iip.txt“ u RStudio koja sadrži podatke o kretanju mjesečnih vrijednosti indeksa industrijske proizvodnje (IIP) u Hrvatskoj, te potom definirajmo da se radi o vremenskom nizu, čija se početna vrijednost odnosi na siječanj 1998. godine. Ukupno je 272

² Vidjeti publikaciju Odabrane teme primijenjene ekonometrije: Linearni regresijski model.

mjesečna podatka prikupljeno do kolovoza 2020. godine. Definirajmo sezonske binarne varijable te potom procijenimo kretanje vrijednosti IIP-a pomoću tih varijabli i interpretirajmo rezultat.

Na slici 1.4. predočeno je učitavanje datoteke u RStudio (naredba "read table") te naredbe:

```
ts(iip$IIP,start=c(1998,1),frequency=12),
```

pri čemu je sada potrebno pomoću naredbe `ts(...)` definirati da se radi o vremenskom nizu (ts – engl. *time series*), čiji je prvi argument naziv niza koji se želi definirati kao vremenski, potom se pomoću `start=...` definira početni datum, a pomoću `frequency=` se definira frekvencija podataka unutar jedne godine. Kada se radi o mjesečnim podacima, frekvencija je 12, kao na slici 1.4., dok za tromjesečne podatke bismo pisali 4. Nadalje, za definiranje binarnih varijabli, može se koristiti paket `forecast` i naredba `seasonaldummy(...)` koja definira 11 binarnih varijabli za mjesečne podatke, odnosno 3 za tromjesečne. Ako se otvori tablica „binarne“, na slici 1.5., uočava se da kada se radi o siječnju, u prvome stupcu vrijednost odgovarajućeg elementa poprima vrijednost 1, a 0 inače. Slično se može uočiti za preostale stupce.

```
iip<-read.table("iip.txt",header=T,sep="\t")
IIP<-ts(iip$IIP,start=c(1998,1),frequency=12)
library(forecast)
binarne<-seasonaldummy(IIP)
```

Slika 1.4. Učitavanje podataka, definiranje vremenskih nizova i binarnih varijabli

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
1	1	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0
4	0	0	0	1	0	0	0	0
5	0	0	0	0	1	0	0	0
6	0	0	0	0	0	1	0	0
7	0	0	0	0	0	0	1	0
8	0	0	0	0	0	0	0	1
9	0	0	0	0	0	0	0	0

Slika 1.5. Tablica „binarne“

Ako se želi procijeniti model (1.10), dovoljno je pomoću naredbe `lm(...)` dobiti ispis na slici 1.6. Stoga se može zapisati sljedeći model:

$$\widehat{IIP}_t = 101,68 - 12,11 D1_t - 7,29 D2_t + \dots + 4 D11_t,$$

$$D1_t = \begin{cases} 1, & \text{za siječanj} \\ 0, & \text{inače} \end{cases}, \quad D2_t = \begin{cases} 1, & \text{za veljaču} \\ 0, & \text{inače} \end{cases}, \quad \dots, \quad D11_t = \begin{cases} 1, & \text{za studeni} \\ 0, & \text{inače} \end{cases},$$

gdje je oznaka `D1` za "binarneJan" na slici 1.6, `D2` za "binarneFeb", itd. do `D11` za "binarneNov". Naime, kako je "binarne" tablica s 11 stupaca, R prepoznaje da se svaki stupac odnosi na pojedini mjesec. Pritom je prvi stupac "Jan" – što na engleskom znači skraćenica za siječanj (*January*), itd. do posljednjeg jedanaestog stupca koji se odnosi na studeni (engl.

November, skraćena Nov). Vrijednosti se interpretiraju na sljedeći način. 101,68 je prosječna vrijednost IIP-a u promatranom razdoblju u indeksnim bodovima u prosincu; zbog učinka sezone je u siječnju vrijednost IIP-a u prosjeku za 12,11 indeksnih bodova manja u odnosu na prosinac; zbog učinka sezone je u veljači vrijednost IIP-a u prosjeku za 7,29 indeksnih bodova manja u odnosu na prosinac; ... do zbog učinka sezone je u studenom vrijednost IIP-a u prosjeku za 4 indeksna boda veća u odnosu na prosinac.

```
summary(lm(IIP~binarne))

##
## Call:
## lm(formula = IIP ~ binarne)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.6818  -4.7288   0.4565   6.1304  20.4783
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  101.6818    2.0036  50.748 < 2e-16 ***
## binarneJan   -12.1166    2.8026  -4.323 2.19e-05 ***
## binarneFeb    -7.2905    2.8026  -2.601 0.00982 **
## binarneMar     3.0138    2.8026   1.075 0.28321
## binarneApr    -0.9862    2.8026  -0.352 0.72522
## binarneMay     2.9269    2.8026   1.044 0.29730
## binarneJun    -0.1601    2.8026  -0.057 0.95450
## binarneJul     4.1877    2.8026   1.494 0.13633
## binarneAug    -7.1166    2.8026  -2.539 0.01169 *
## binarneSep     1.9091    2.8336   0.674 0.50108
## binarneOct     4.5909    2.8336   1.620 0.10641
## binarneNov     4.0000    2.8336   1.412 0.15925
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.398 on 260 degrees of freedom
## Multiple R-squared:  0.2433, Adjusted R-squared:  0.2113
## F-statistic: 7.599 on 11 and 260 DF, p-value: 2.405e-11
```

Slika 1.6. Procjena modela (1.10)

S druge strane, ako bismo htjeli razmotriti model (1.11), potrebno je definirati nove binarne varijable, u okviru paketa `tsutils`, pomoću naredbe

```
seasdummy(length(IIP),12,y=IIP,type = "bin",full = T);
```

gdje je prvi argument funkcije `seasdummy` duljina vremenskog niza, drugi argument se odnosi na frekvenciju podataka (12), a treći argument je naziv varijable koja je vremenski niz kako bi naredba prepoznala koji je početni datum. Slika 1.7. prikazuje naredbe za definiranje novih binarnih varijabli, te procjenu novog modela.

Sada možemo pisati sljedeći procijenjeni model:

$$\widehat{IIP}_t = 89.57 D1_t + 94.39 D2_t + 104.7 D3_t + \dots + 105.86 D11_t + 101.68 D12_t,$$

$$D1_t = \begin{cases} 1, & \text{za siječanj} \\ 0, & \text{inače} \end{cases}, \quad D2_t = \begin{cases} 1, & \text{za veljaču} \\ 0, & \text{inače} \end{cases}, \quad \dots, \quad D12_t = \begin{cases} 1, & \text{za prosinac} \\ 0, & \text{inače} \end{cases}.$$

Sada D1 označava binarnu varijablu za siječanj, koja je u ispisu slike 1.7. dana kao sve.binarne1, jer Rstudio prepoznaje tablicu sve.binarne kao tablicu vremenskih serija, u kojima se prvi stupac odnosi na siječanj, potom D2 označava binarnu varijablu za veljaču (sve.binarne2 u ispisu), itd. do posljednje, D12 koja označava prosinac (sve.binarne12). Interpretacija procijenjenih parametara je sljedeća: vrijednost 89,57 interpretira se kao prosječna razina IIP-a u indeksnim bodovima u siječnju u promatranom razdoblju; 94,39 je prosječna razina IIP-a u indeksnim bodovima u veljači u promatranom razdoblju, ... do 101,68, što je prosječna razina IIP-a u indeksnim bodovima u prosincu u promatranom razdoblju.

```
library(tsutils)
sve.binarne<-seasdummy(length(IIP),12,y=IIP,type="bin",full=T)
summary(lm(IIP~0+sve.binarne))

##
## Call:
## lm(formula = IIP ~ 0 + sve.binarne)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.6818  -4.7288   0.4565   6.1304  20.4783
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## sve.binarne1    89.565     1.960   45.71 <2e-16 ***
## sve.binarne2    94.391     1.960   48.17 <2e-16 ***
## sve.binarne3   104.696     1.960   53.43 <2e-16 ***
## sve.binarne4   100.696     1.960   51.39 <2e-16 ***
## sve.binarne5   104.609     1.960   53.38 <2e-16 ***
## sve.binarne6   101.522     1.960   51.81 <2e-16 ***
## sve.binarne7   105.870     1.960   54.03 <2e-16 ***
## sve.binarne8    94.565     1.960   48.26 <2e-16 ***
## sve.binarne9   103.591     2.004   51.70 <2e-16 ***
## sve.binarne10  106.273     2.004   53.04 <2e-16 ***
## sve.binarne11  105.682     2.004   52.74 <2e-16 ***
## sve.binarne12  101.682     2.004   50.75 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.398 on 260 degrees of freedom
## Multiple R-squared:  0.9918, Adjusted R-squared:  0.9914
## F-statistic: 2627 on 12 and 260 DF, p-value: < 2.2e-16
```

Slika 1.7. Definiranje novih binarnih varijabli i procjena modela (1.11)

Uočimo da su u slučaju ovog drugog modela sve binarne varijable značajne u modelu (stupci koji se odnose na t -test), dok u slučaju prethodnog modela (vidjeti sliku 1.6) se pojedinačni testovi značajnosti odnose na razliku u prosječnim razinama između pojedinog mjeseca i prosinca kao baznog. U tom slučaju je jedino postojala značajna razlika (pri razini $\alpha=5\%$) između prosječne razine IIP-a u prosincu i siječnju, prosincu i veljači te prosincu i kolovožu.

Primjer 1.2.

Grafički prikaz IIP-a iz prethodnog primjera već je predložen na slici 1.1. (a), gdje se uočava da zbog učinka krize iz 2008. godine dolazi do promjene u trendu ponašanja IIP-a, sve do oporavka koji se uočava od 2015. godine. U tu svrhu ćemo definirati uz sezonske binarne varijable još 2 dodatne. Prva binarna varijabla će poprimiti vrijednost 1 za vremensko razdoblje od siječnja 2008. do prosinca 2013. godine, a 0 inače, te će druga binarna varijabla poprimiti vrijednost 1 za razdoblje od siječnja 2014. do kolovoza 2020. godine, a 0 inače. Potrebno je definirati i

varijablu vrijeme, a kako imamo ukupno 272 podatka za varijablu IIP, linearni trend će poprimiti vrijednosti 1, 2, ... do 272. Potom ćemo procijeniti sljedeći regresijski model:

$$IIP_t = \beta_0 + \beta_1 D1_t + \beta_2 D2_t + \dots + \beta_{11} D11_t + \gamma_1 t + \gamma_2 t B1_t + \gamma_3 t B2_t + \varepsilon_t$$

$$t \in \{1, 2, \dots, 272\}, B1_t = \begin{cases} 1, & 1.2008. - 12.2013. \\ 0, & \text{inače} \end{cases}, B2_t = \begin{cases} 1, & 1.2014. - 8.2020. \\ 0, & \text{inače} \end{cases}$$

u kojemu se pretpostavlja da uz sezonske učinke koji su opisani binarnim varijablama, postoji promjena u ponašanju trenda IIP varijable, na način da postoji jedan linearni trend u prvom podrazdoblju, drugačiji linearni trend u drugom, te linearni trend oporavka u trećem podrazdoblju. Naredbe za definiranje varijable trend, kao i nove dvije binarne varijable predložene su na slici 5.8. Najprije je definirana varijabla trend pomoću naredbe `ts`, gdje je duljina vremenskog niza određena duljinom niza IIP (naredba `length(IIP)`), pri čemu trend poprima vrijednosti 1, 2, ... do 272. Binarna varijabla $B1_t$ je definirana na sljedeći način:

```
bin1<-c(rep(0,120),rep(1,72),rep(0,272-120-72))
```

gdje `c()` naredba znači spajanje podataka (`c` – za *combine*), gdje se spajaju tri serije, kod kojih naredba `rep()` znači da se ponavlja (`rep` – za *repeat*) broj koji je dan kao prvi argument, a toliko puta koliko iznosi drugi argument. Prvih 120 podataka (od početka uzorka do prosinca 2007. godine) vrijednosti jednake 0 (dio naredbe `rep(0,120)` što znači da se vrijednost nula ponovi 120 puta); potom je idućih 72 (siječanj 2008. – prosinac 2013. godine) vrijednosti jednako 1 (`rep(1,72)`) i preostalih $272 - 120 - 72$ (od siječnja 2014. godine do kraja uzorka) vrijednosti su jednake 0 ponovno. Binarna $B2_t$ varijabla je definirana na način da je prvih $120 + 72$ vrijednosti jednako 0, a posljednjih $272 - 120 - 72$ jednako 1. Sam procijenjeni model je sljedeći:

$$\widehat{IIP}_t = 81,64 - 10,37 D1_t - 5,73 D2_t + \dots + 4,18 D11_t + 0,27t - 0,13tB1_t - 0,18tB2_t$$

$$t \in \{1, 2, \dots, 272\}, B1_t = \begin{cases} 1, & 1.2008. - 12.2013. \\ 0, & \text{inače} \end{cases}, B2_t = \begin{cases} 1, & 1.2014. - 8.2020. \\ 0, & \text{inače} \end{cases}$$

```
trend<-ts(1:length(IIP),start=c(1998,1),frequency=12)
bin1<-c(rep(0,120),rep(1,72),rep(0,272-120-72))
bin2<-c(rep(0,120+72),rep(1,272-120-72))
summary(m1<-lm(IIP~binarne+trend+I(bin1*trend)+I(bin2*trend)))

##
## Call:
## lm(formula = IIP ~ binarne + trend + I(bin1 * trend) + I(bin2 *
##   trend))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.0738  -3.5930  -0.2028   2.8054  22.3634
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    81.64248    1.70639   47.845 < 2e-16 ***
## binarneJan     -10.37087    1.84773  -5.613 5.15e-08 ***
## binarneFeb      -5.72549    1.84721  -3.100 0.002154 **
## binarneMar       4.39814    1.84675   2.382 0.017968 *
## binarneApr       0.21743    1.84634   0.118 0.906349
## binarneMay       3.94976    1.84598   2.140 0.033325 *
```

```
## binarneJun      0.68209    1.84568    0.370 0.712015
## binarneJul      4.84920    1.84543    2.628 0.009114 **
## binarneAug     -6.63586    1.84523   -3.596 0.000387 ***
## binarneSep      2.46322    1.86533    1.321 0.187834
## binarneOct      4.96033    1.86519    2.659 0.008319 **
## binarneNov      4.18471    1.86511    2.244 0.025705 *
## trend           0.26826    0.01596   16.809 < 2e-16 ***
## I(bin1 * trend) -0.13086    0.01155  -11.326 < 2e-16 ***
## I(bin2 * trend) -0.17547    0.01251  -14.025 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.186 on 257 degrees of freedom
## Multiple R-squared:  0.6759, Adjusted R-squared:  0.6583
## F-statistic: 38.29 on 14 and 257 DF, p-value: < 2.2e-16
```

Slika 1.8. Definiranje novih binarnih varijabli i procjena modela

Sada bi interpretacija konstante u modelu bila sljedeća: prosječna razina IIP-a u prosincu 1997. godine iznosi 81,64 indeksnih bodova (jer se radi o slučaju kada su sve binarne varijable jednake 0, a to je slučaj mjeseca prosinca i kada je trend varijabla jednaka 0. S obzirom da je za siječanj 1998. godine trend varijabla jednaka 1, za prethodnu vrijednost, prosinac 1997. godine, trend varijabla iznosila bi 0). Procijenjeni parametri uz sezonske binarne varijable se interpretiraju kao razlika u prosječnoj vrijednosti IIP-a u pojedinom mjesecu u odnosu na prosinac, bez obzira na sam trend. Vrijednost 0,268 uz varijablu trend se interpretira da je u promatranom razdoblju do krize prosječno mjesečno povećanje IIP-a iznosilo 0,268 indeksnih bodova; za vrijeme krize je prosječno mjesečno povećanje IIP-a bilo za 0,13 indeksnih bodova manje u odnosu na razdoblje prije i nakon krize, dok je u posljednjem podrazdoblju nakon krize prosječno mjesečno povećanje IIP-a bilo za 0,18 indeksnih bodova manje u odnosu na prethodno razdoblje.

Napomena. Detalje o trendu i njegovom modeliranju u vremenskim nizovima vidjeti u odjeljku 1.10.

Primjer 1.3.

U prethodnom primjeru (1.2.) provedimo Waldov test značajnosti³ sezonskih učinaka u modelu, uz $\alpha=5\%$.

Testiraju se sljedeće hipoteze: $H_0 \dots \mathbf{R}\boldsymbol{\beta} = \mathbf{q}$, $\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 \end{bmatrix}$, $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \gamma_3 \end{bmatrix}$, $\mathbf{q} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$,

$H_0 \dots \mathbf{R}\boldsymbol{\beta} \neq \mathbf{q}$, gdje je matrica \mathbf{R} formata (11,15), jer se testira hipoteza $\beta_1 = \beta_2 = \dots = \beta_{11} = 0$, stoga se radi o 11 jednakosti, dok je 15 parametara u samome modelu.

```
library(car)
ogranicenje<-c("binarneJan=0", "binarneFeb=0", "binarneMar=0",
              "binarneApr=0", "binarneMay=0", "binarneJun=0",
              "binarneJul=0", "binarneAug=0", "binarneSep=0",
              "binarneOct=0", "binarneNov=0")
linearHypothesis(m1,ogranicenje,type="Chisq")
```

³ Vidjeti publikaciju Odabrane teme primijenjene ekonometrije: Linearni regresijski model.


```
## Linear hypothesis test
##
## Hypothesis:
## binarneJan = 0
## binarneFeb = 0
## binarneMar = 0
## binarneApr = 0
## binarneMay = 0
## binarneJun = 0
## binarneJul = 0
## binarneAug = 0
## binarneSep = 0
## binarneOct = 0
## binarneNov = 0
##
## Model 1: restricted model
## Model 2: IIP ~ binarne + trend + I(bin1 * trend) + I(bin2 * trend)
##
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1     268 16467.7
## 2     257  9833.8 11     6634 15.761 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Slika 1.9. Testiranje značajnosti sezonskih učinaka

Na slici 1.9. su predočene naredbe i ispis samog testa, gdje se uočava da je test veličina jednaka $W = 15,76$, s pripadajućom p -vrijednosti približno 0, stoga se uz razinu značajnosti od 5%, odbacuje hipoteza da niti jedan sezonalni učinak nije značajan u modelu.

Kada bi se pretpostavilo da će se buduće ponašanje vremenskog niza temeljiti na modelu (1.10), odnosno (1.11), tada se ti modeli mogu koristiti u svrhu prognoziranja budućih vrijednosti promatrane pojave.

Naime, procjena modela (1.11) nad uzorkom podataka $t \in \{1, 2, \dots, T\}$:

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 D1_t + \hat{\beta}_2 D2_t + \dots + \hat{\beta}_{11} D11_t, \quad (1.12)$$

koristi se za prognoziranje budućih vrijednosti zavisne varijable, jer su za vremensko razdoblje $T+1$ poznate vrijednosti svih binarnih varijabli. Naime, potrebno je razmotriti o kojem se mjesecu radi kada nastupi razdoblje $T+1$, te se u (1.9) uvrste vrijednosti 0 i 1, koje se potom uvrste u model (1.12) kako bi se definirale vrijednosti D_{if} , $i \in \{1, 2, \dots, 11\}$, te potom prognozirala vrijednost:

$$\hat{y}_f = \hat{\beta}_0 + \hat{\beta}_1 D1_f + \hat{\beta}_2 D2_f + \dots + \hat{\beta}_{11} D11_f. \quad (1.13)$$

Pogreška predviđanja i ostali pojmovi definirani u poglavlju 2.2.8. u prvome dijelu (Odabrane teme primijenjene ekonometrije: Linearni regresijski model) se slično i ovdje mogu definirati i interpretirati.

Primjer 1.4.

Temeljem procijenjenog modela u primjeru 1.1.:

$$\widehat{IIP}_t = 101,68 - 12,11 D1_t - 7,29 D2_t + \dots + 4 D11_t$$

$$D1_t = \begin{cases} 1, & \text{za siječanj} \\ 0, & \text{inače} \end{cases}, \quad D2_t = \begin{cases} 1, & \text{za veljaču} \\ 0, & \text{inače} \end{cases}, \dots, \quad D11_t = \begin{cases} 1, & \text{za studeni} \\ 0, & \text{inače} \end{cases},$$

prognozirajmo vrijednosti IIP-a za razdoblje rujan – prosinac 2020. godine.

```
model<-lm(IIP~binarne)
binarne<-seasonaldummy(ts(273:276,start=c(2020,9),frequency=12))
binarne<-ts(binarne,start=c(2020,9),frequency=12)
prognoza<-predict(model,newdata=binarne)
procjene<-(fitted(model))
zajedno<-c(procjene,prognoza)
zajedno<-ts(zajedno,start=c(1998,1),frequency=12)
cbind(IIP,zajedno)
```

Slika 1.10. Prognoziranje izvan uzorka temeljem binarnih varijabli

Slika 1.10. predočuje naredbe koje su potrebne za prognoziranje izvan uzorka (engl. *out of sample forecasting*), gdje je prvo spremljen model u kojemu se IIP regresira na skup binarnih varijabli. Potom su kao nove binarne spremljeni vremenski nizovi naredbom `seasonaldummy(...)`:

```
seasonaldummy(ts(273:276,start=c(2020,9),frequency=12)),
```

gdje su vrijednosti izvan uzorka $T+1$, $T+2$, $T+3$ i $T+4$ jednaki 273 do 276, pri čemu je prva početna vrijednost u rujnu 2020. godine. Pomoću naredbe `predict` (kao i u odjeljku 2.2.8 u prvome dijelu Odabrane teme primijenjene ekonometrije: Linearni regresijski model) se temeljem novih podataka sprema prognoza izvan uzorka. Pomoću naredbe `fitted(...)` su spremljene procijenjene vrijednosti unutar uzorka i one su spojene zajedno s vrijednostima izvan uzorka pomoću naredbe `c(...)` u vremenski niz „zajedno“. Na kraju je uspoređen niz IIP (originalni podaci) s nizom zajedno, te slika 1.11. predočava usporedbu posljednje 4 vrijednosti, koje naravno u trenutku modeliranja nisu poznate za IIP (stoga u ispisu piše NA – engl. *not available*), ali prognoze izvan uzorka iznose redom: 103,59, 106,27, 105,68 i 101,68.

```
## Jun 2020 102 101.52174
## Jul 2020 111 105.86957
## Aug 2020 91 94.56522
## Sep 2020 NA 103.59091
## Oct 2020 NA 106.27273
## Nov 2020 NA 105.68182
## Dec 2020 NA 101.68182
```

Slika 1.11. Usporedba vrijednosti IIP i niza „zajedno“

Ako razmatramo tromjesečne nizove, možemo razmatrati sljedeća dva osnovna modela:

$$y_t = \beta_0 + \beta_1 Q1_t + \beta_2 Q2_t + \beta_3 Q3_t + \varepsilon_t, \quad (1.14)$$

ili

$$y_t = \beta_1 Q1_t + \beta_2 Q2_t + \beta_3 Q3_t + \beta_4 Q4_t + \varepsilon_t, \quad (1.15)$$

gdje su

$$Q_{1t} = \begin{cases} 1, & \text{za 1. kvartal} \\ 0, & \text{inače} \end{cases}, Q_{2t} = \begin{cases} 1, & \text{za 2. kvartal} \\ 0, & \text{inače} \end{cases}, Q_{3t} = \begin{cases} 1, & \text{za 3. kvartal} \\ 0, & \text{inače} \end{cases}, Q_{4t} = \begin{cases} 1, & \text{za 4. kvartal} \\ 0, & \text{inače} \end{cases}.$$

Interpretacija je slična kao za modele (1.10) i (1.11), samo što se sada radi o slučaju tromjesečnih podataka pa treba paziti na interpretacije.

Ako razmatramo dnevne nizove, možemo razmatrati sljedeća dva osnovna modela:

$$y_t = \beta_0 + \beta_1 P_t + \beta_2 U_t + \beta_3 S_t + \beta_4 \check{C}_t + \varepsilon_t, \quad (1.16)$$

ili

$$y_t = \beta_1 P_t + \beta_2 U_t + \beta_3 S_t + \beta_4 \check{C}_t + \beta_5 P e_t + \varepsilon_t, \quad (1.17)$$

gdje su

$$P_t = \begin{cases} 1, & \text{ponedjeljak} \\ 0, & \text{inače} \end{cases}, U_t = \begin{cases} 1, & \text{utorak} \\ 0, & \text{inače} \end{cases}, S_t = \begin{cases} 1, & \text{srijeda} \\ 0, & \text{inače} \end{cases}, \check{C}_t = \begin{cases} 1, & \text{četvrtak} \\ 0, & \text{inače} \end{cases}, P e_t = \begin{cases} 1, & \text{petak} \\ 0, & \text{inače} \end{cases}.$$

U empirijskim primjenama se često razmatraju modeli (1.16) i (1.17) ili njihove modifikacije za analizu učinka dana u tjednu u prinosima dionica (engl. *day of the week effect*), vidjeti primjerice Šego i Škrinjarić (2012) i literaturi navedenoj u spomenutom radu.

Primjer 1.5.

Učitajmo datoteku „**crobox.txt**“ u RStudio, koja sadrži dnevne podatke o kretanju indeksa CROBEX u razdoblju od 2. siječnja 2019. do 31. srpnja 2020. godine. Definirajmo najprije prinose na vrijednost CROBEX-a pomoću formule (1.6), te pomoću binarnih varijabli za radne dane u tjednu definirane za model (1.17) procijenimo parametre modela i interpretirajmo. Kako postoji problem autokorelacije i heteroskedastičnosti reziduala, primijenit ćemo Newey-Westove korekcije standardnih pogrešaka procjenitelja (vidjeti odjeljak 4.3.3.2. u prvome dijelu Odabrane teme primijenjene ekonometrije: Linearni regresijski model), uz pretpostavku autokorelacije prvog reda.

Slika 1.12. predočava naredbe za izračun prinosa, na način da pomoću naredbe `diff(log(...))` računamo prinos koristeći formulu (1.6), dok naredbom `c(NA,...)` na prvo mjesto dodajemo nedostajuću vrijednost kako bi niz prinos imao jednaku duljinu kao i binarni nizovi za radne dane u tjednu (`pon`, `uto`, `sri`, `cet` i `pet`; s obzirom da u naredbi `lm(...)` svi nizovi trebaju biti jednake duljine). Ostale naredbe su opisane u prvome dijelu ovih publikacija za izvršenje potrebnih korekcija. Procijenjen je model:

$$\hat{r}_t = -0.25P_t + 0.20U_t - 0.12S_t - 0.09\check{C}_t + 0.14P e_t,$$

gdje su

$$P_t = \begin{cases} 1, & \text{ponedjeljak} \\ 0, & \text{inače} \end{cases}, U_t = \begin{cases} 1, & \text{utorak} \\ 0, & \text{inače} \end{cases}, S_t = \begin{cases} 1, & \text{srijeda} \\ 0, & \text{inače} \end{cases}, \check{C}_t = \begin{cases} 1, & \text{četvrtak} \\ 0, & \text{inače} \end{cases}, P e_t = \begin{cases} 1, & \text{petak} \\ 0, & \text{inače} \end{cases}.$$

Uočava se kako je značajna jedino binarna varijabla vezana uz utorak, pri čemu su natprosječno pozitivni prinosi upravo taj dan.

```

crobex<-read.table("crobex.txt",header=T,sep="\t")
prinos<-c(NA,diff(log(crobex$Crobex))*100)
cro<-lm(prinos~0+pon+uto+sri+cet+pet,data=crobex)
#Newey-West korekcija:
library(sandwich)
mat<-NeweyWest(cro,lag=1)
library(lmtest)
coefTest(cro,vcov=mat)

##
## t test of coefficients:
##
##      Estimate Std. Error t value Pr(>|t|)
## pon -0.254185  0.191298 -1.3287  0.18472
## uto  0.202248  0.079776  2.5352  0.01163 *
## sri -0.121787  0.086194 -1.4129  0.15848
## cet -0.090991  0.150527 -0.6045  0.54588
## pet  0.141479  0.093878  1.5071  0.13261
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

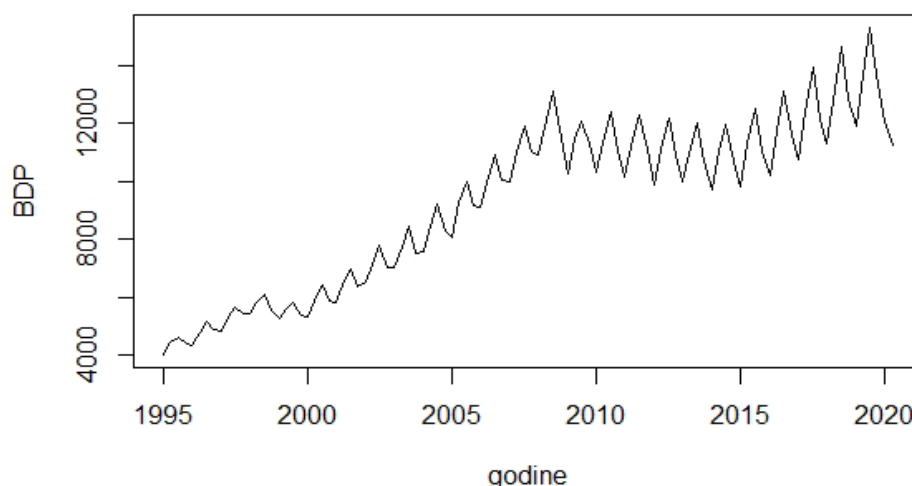
Slika 1.12. Procjena modela (1.17)

1.4. Primjer

Učitajmo datoteku „**bdp.txt**“ u RStudio. Datoteka sadrži tromjesečne podatke o kretanju BDP-a (milijuni eura) Hrvatske od prvog tromjesečja tromjesečja 1995. do drugog tromjesečja tromjesečja 2020. godine.

- a) Grafički predočimo kretanje varijable BDP kroz vrijeme te komentirajmo što se događa s pojavom kroz vrijeme.

Na slici 1.13. se uočava da postoji sezonalni učinak u kretanju BDP-a tijekom vremena, pri čemu se taj učinak promijenio nakon krize iz 2008. godine. Dodatno, uočava se jedan linearan trend ponašanja do spomenute krize, a drugačiji u vrijeme i nakon krize.



Slika 1.13. Vremenska dinamika BDP-a u razmatranom primjeru

- b) Procijenimo tri modela: u prvome neka vrijednost BDP-a kroz vrijeme ovisi o sezonskim binarnim varijablama i konstanti, u drugome neka učinak sezonskih binarnih varijabli ovisi o informaciji radi li se o razdoblju prije krize 2008. godine ili nakon, te u

- c) treći model uz binarne varijable za sezonski učinak uključimo i varijablu linearni trend, pri čemu će informacija o tome radi li se o razdoblju prije krize 2008. godine ili nakon utjecati i na linearni trend i na binarne varijable. Spojimo rezultate zajedno u jednu tablicu pomoću naredbe `stargazer(...)` i usporedimo kakvoću modela odgovarajućim mjerama prikladnosti.

S obzirom na definirane modele u tekstu zadatka, potrebno je procijeniti sljedeća tri modela:

$$m1: BDP_t = \beta_0 + \beta_1 Q1_t + \beta_2 Q2_t + \beta_3 Q3_t + \varepsilon_t,$$

$$m2: BDP_t = \beta_0 + \beta_1 Q1_t + \beta_2 Q2_t + \beta_3 Q3_t + \beta_4 Q1_t D_t + \beta_5 Q2_t D_t + \beta_6 Q3_t D_t + \varepsilon_t,$$

$$m3: BDP_t = \beta_0 + \beta_1 Q1_t + \beta_2 Q2_t + \beta_3 Q3_t + \beta_4 Q1_t D_t + \beta_5 Q2_t D_t + \beta_6 Q3_t D_t + \beta_7 t + \beta_8 t D_t + \varepsilon_t,$$

$$Q1_t = \begin{cases} 1, & \text{za 1. kvartal} \\ 0, & \text{inače} \end{cases}, \quad Q2_t = \begin{cases} 1, & \text{za 2. kvartal} \\ 0, & \text{inače} \end{cases}, \quad Q3_t = \begin{cases} 1, & \text{za 3. kvartal} \\ 0, & \text{inače} \end{cases}, \quad D_t = \begin{cases} 1, & \text{do 4Q 2007} \\ 0, & \text{inače} \end{cases},$$

$t \in \{1, 2, \dots, 102\}$. Naredbe za definiranje svih varijabli te procjene sva tri modela prikazane su na slici 1.14. S naredbom `seasonaldummy()` se definiraju sezonske binarne varijable, a za drugu binarnu varijablu koja se odnosi na razdoblje prije krize 2008. godine se definira niz jedinica kojih ima ukupno 52 (od početka uzorka pa do zadnjeg tromjesečja tromjesečja 2007. godine), dok se drugi dio niza definira kao niz nula od 53. do zadnje opservacija, 102. jer ukupno imamo 102 opservacije za BDP (u ispisu je to varijabla `bin2`, dobivena naredbom `c(...)` u kojoj se spajaju niz jedinica (u ispisu je to `rep(1,52)`); te niz nula, kojih ima ukupno $102-52$). Ispis sva tri modela prikazan na slici 1.15. Uočava se kako je korigirani koeficijent determinacije najveći u slučaju modela `m3`, stoga će se ostatak dijagnostike u preostalim postupcima odnositi na model `m3`.

```
library(forecast)
bin<-seasonaldummy(bdp)
bin2<-c(rep(1,52),rep(0,102-52))
trend<-ts(1:102,start=c(1995,1),frequency = 4)
m1<-lm(bdp~bin)
m2<-lm(bdp~bin+I(bin*bin2))
m3<-lm(bdp~bin+trend+I(bin*bin2)+I(trend*bin2))
library(stargazer)
stargazer(list(m1,m2,m3),type="text")
```

Slika 1.14. Naredbe potrebne za procjenu tri odabrana modela

```
##
## =====
##                               Dependent variable:
## -----
##                               bdp
##                               (1)      (2)      (3)
## -----
## binQ1                        -701.942    1,375.289**    -349.083
##                               (807.241)    (683.201)    (347.149)
##
## binQ2                         254.446    2,604.681***    794.594**
##                               (807.241)    (683.201)    (348.891)
##
## binQ3                         999.380    3,783.387***    2,059.016***
##                               (815.117)    (701.680)    (354.479)
```

```

##
## trend                                85.716***
##                                     (4.648)
##
## I(bin * bin2)Q1                      -4,154.462***
##                                     (783.686)
##                                     -415.856
##                                     (471.687)
##
## I(bin * bin2)Q2                      -4,700.469***
##                                     (783.686)
##                                     -990.609**
##                                     (474.755)
##
## I(bin * bin2)Q3                      -5,353.860***
##                                     (799.846)
##                                     -1,844.177***
##                                     (478.246)
##
## I(trend * bin2)                      28.745***
##                                     (7.013)
##
## Constant                             9,176.896***
##                                     (576.375)
##                                     9,176.896***
##                                     (399.603)
##                                     4,301.121***
##                                     (308.517)
## -----
## Observations                         102
## R2                                    0.044
## Adjusted R2                          0.015
## Residual Std. Error 2,881.873 (df = 98) 1,998.014 (df = 95) 895.741 (df = 93)
## F Statistic                          1.514 (df = 3; 98) 19.722*** (df = 6; 95) 121.051*** (df = 8; 93)
## =====
## Note:                                  *p<0.1; **p<0.05; ***p<0.01

```

Slika 1.15. Ispis sva tri modela zajedno

d) Interpretirajmo sve parametre u sva tri modela.

$$\widehat{BDP}_t = 9176,90 - 701,94 Q1_t + 254,45 Q2_t + 999,38 Q3_t :$$

- $\hat{\beta}_0 = 9176,90$ - u promatranom razdoblju, prosječna vrijednost BDP-a u Hrvatskoj u 4. tromjesečju iznosi 9176,90 milijuna eura.
- $\hat{\beta}_1 = -701,94$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 1. tromjesečju u prosjeku za 701,94 milijuna eura manji u odnosu na 4. tromjesečje.
- $\hat{\beta}_2 = 254,45$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 2. tromjesečju u prosjeku za 254,45 milijuna eura veći u odnosu na 4. tromjesečje.
- $\hat{\beta}_3 = 999,38$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 3. tromjesečju u prosjeku za 999,38 milijuna eura veći u odnosu na 4. tromjesečje.

$$\widehat{BDP}_t = 9176,90 + 1375,29 Q1_t + 2604,68 Q2_t + 3783,39 Q3_t - 4154,46 Q1_t D_t - 4700,47 Q2_t D_t - 5353,86 Q3_t D_t :$$

Ako se razmotri model prije krize (kada je binarna varijabla $D_t = 1$):

$$\widehat{BDP}_t = 9176,90 + 1375,29 Q1_t + 2604,68 Q2_t + 3783,39 Q3_t - 4154,46 Q1_t D_t - 4700,47 Q1_t D_t - 5353,86 Q3_t D_t = 9176,90 - 2779,17 Q1_t - 2095,79 Q2_t - 1570,47 Q3_t$$

s onim u vrijeme i nakon krize ($D_t = 0$):

$$\widehat{BDP}_t = 9176,90 + 1375,29 Q1_t + 2604,68 Q2_t + 3783,39 Q3_t,$$

uočava se da dolazi do promjena u vrijednostima procijenjenih parametara uz binarne varijable.

Stoga u razdoblju prije krize interpretiramo parametre (kod svakog parametra treba naglasiti da se radi o razdoblju prije krize):

- $\hat{\beta}_0 = 9176,90$ - u promatranom razdoblju, prosječna vrijednost BDP-a u Hrvatskoj u 4. tromjesečju iznosi 9176,90 milijuna eura.
- $\hat{\beta}_1 = -2779,17$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 1. tromjesečju u prosjeku za 2779,17 milijuna eura manji u odnosu na 4. tromjesečje.
- $\hat{\beta}_2 = -2095,79$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 2. tromjesečju u prosjeku za 2095,79 milijuna eura manji u odnosu na 4. tromjesečje.
- $\hat{\beta}_3 = -1570,47$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 3. tromjesečju u prosjeku za 1570,47 milijuna eura manji u odnosu na 4. tromjesečje.

S obzirom da ova specifikacija modela ne mora biti točna za opisivanje kretanja hrvatskog BDP-a, rezultati za razdoblje prije krize mogu biti neintuitivni, s obzirom da je u trećem tromjesečju BDP veći u odnosu na vrijednost u četvrtom tromjesečju (zbog učinka turističke sezone).

Za razdoblje u vrijeme i nakon krize, interpretacije su sljedeće (kod svakog parametra treba naglasiti da se radi o razdoblju prije krize):

- $\hat{\beta}_0 = 9176,90$ - u promatranom razdoblju, prosječna vrijednost BDP-a u Hrvatskoj u 4. tromjesečju iznosi 9176,90 milijuna eura.
- $\hat{\beta}_1 = 1375,29$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 1. tromjesečju u prosjeku za 1375,29 milijuna eura veći u odnosu na 4. tromjesečje.
- $\hat{\beta}_2 = 2604,68$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 2. tromjesečju u prosjeku za 2604,68 milijuna eura veći u odnosu na 4. tromjesečje.
- $\hat{\beta}_3 = 3783,39$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 3. tromjesečju u prosjeku za 3783,39 milijuna eura veći u odnosu na 4. tromjesečje.

Stoga vrijednosti $-4154,46$, $-4700,47$ i $-5353,86$ u početnom modelu predstavljaju prosječne razlike u učincima sezone za 1., 2. i 3. tromjesečje u odnosu na 4. tromjesečje, s obzirom na to radi li se o razdoblju prije krize ili u vrijeme i nakon krize.

$$\widehat{BDP}_t = 4301,12 - 349,08Q1_t + 794,59Q2_t + 2059,02Q3_t - 415,86Q1_tD_t - 990,61 Q2_tD_t - 1844,18Q3_tD_t + 85,72t + 28,75tD_t$$

Model u razdoblju prije krize je ($D_t = 1$):

$$\widehat{BDP}_t = 4301,12 - 349,08Q1_t + 794,59Q2_t + 2059,02Q3_t + 114,47t$$

Stoga u razdoblju prije krize interpretiramo parametre (kod svakog parametra treba naglasiti da se radi o razdoblju prije krize):

- $\hat{\beta}_0 = 4301,12$ - prosječna vrijednost BDP-a u Hrvatskoj u 4. tromjesečju 1994. godine iznosi 4301,12 milijuna eura (kada su sve binarne varijable jednake 0, te je $t = 0$).
- $\hat{\beta}_1 = -764,94$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 1. tromjesečju u prosjeku za 764,94 milijuna eura manji u odnosu na 4. tromjesečje.
- $\hat{\beta}_2 = -196,02$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 2. tromjesečju u prosjeku za 196,02 milijuna eura manji u odnosu na 4. tromjesečje.
- $\hat{\beta}_3 = 214,84$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 3. tromjesečju u prosjeku za 214,84 milijuna eura veći u odnosu na 4. tromjesečje.
- $\hat{\beta}_4 = 114,47$ - u promatranom razdoblju, prosječno tromjesečno povećanje BDP-a u Hrvatskoj iznosi 114,47 milijuna eura.

Model u razdoblju u vrijeme i nakon krize je ($D_t = 0$):

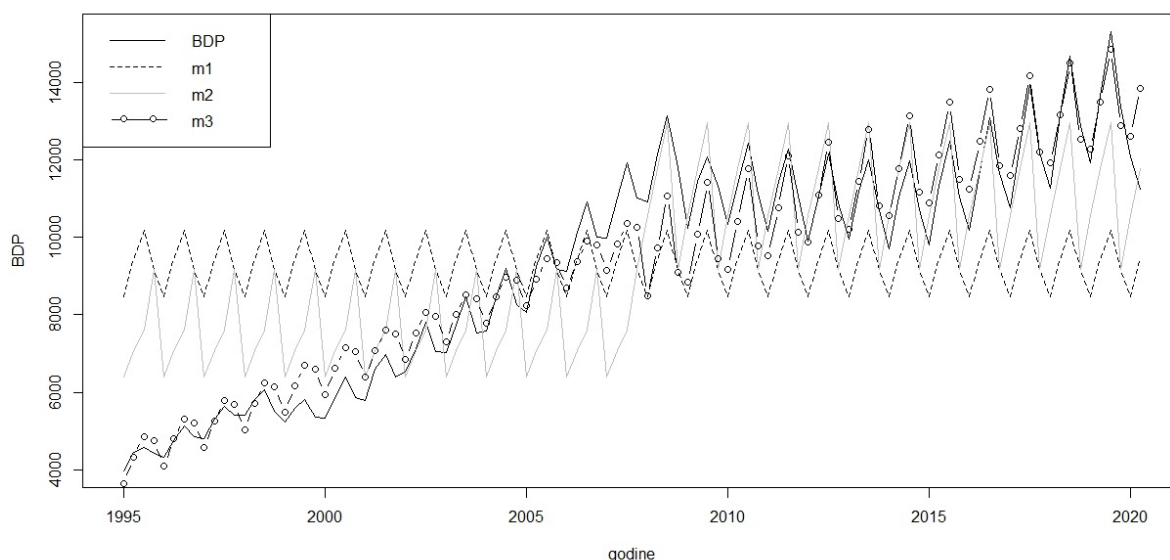
$$\widehat{BDP}_t = 4301,12 - 349,08Q1_t + 794,59Q2_t + 2059,02Q3_t + 85,72t$$

Stoga u razdoblju prije krize interpretiramo parametre (kod svakog parametra treba naglasiti da se radi o razdoblju prije krize):

- $\hat{\beta}_0 = 4301,12$ - prosječna vrijednost BDP-a u Hrvatskoj u 4. tromjesečju 1994. godine iznosi 4301,12 milijuna eura (kada su sve binarne varijable jednake 0, te je $t = 0$). Ovo je poseban slučaj kada se razmatra konstanta također za razdoblje prije prvog tromjesečja tromjesečja 1995. godine, s obzirom da se radi o prosječnoj razini pojave kada je $t = 0$.
- $\hat{\beta}_1 = -349,08$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 1. tromjesečju u prosjeku za 349,08 milijuna eura manji u odnosu na 4. tromjesečje.
- $\hat{\beta}_2 = 794,59$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 2. tromjesečju u prosjeku za 794,59 milijuna eura veći u odnosu na 4. tromjesečje.
- $\hat{\beta}_3 = 2059,02$ - u promatranom razdoblju je zbog učinka sezone vrijednost BDP-a u Hrvatskoj u 3. tromjesečju u prosjeku za 2059,02 milijuna eura veći u odnosu na 4. tromjesečje.
- $\hat{\beta}_4 = 85,72$ - u promatranom razdoblju, prosječno tromjesečno povećanje BDP-a u Hrvatskoj iznosi 85,72 milijuna eura.

Stoga vrijednosti $-415,86, -990,61$ i $-1844,18$ u početnom modelu predstavljaju razlike u učincima sezone za 1., 2. i 3. tromjesečje u odnosu na 4. tromjesečje, s obzirom na to radi li se o razdoblju prije krize ili u vrijeme i nakon krize. Također, $28,75$ je razlika u prosječnom mjesečnom povećanju BDP-a u ovisnosti razmatramo li razdoblje prije krize ili u vrijeme i nakon krize.

- e) Usporedimo grafički procijenjene vrijednosti u sva tri modela sa stvarnim vrijednostima BDP-a na jednome grafikonu te komentirajmo.



Slika 1.16. Grafička usporedba procijenjenih vrijednosti iz sva tri modela sa stvarnim vrijednostima BDP-a

Slika 1.16 predočava usporedbu spomenutih vrijednosti. Crna krivulja predstavlja stvarnu vrijednost BDP-a, dok crna iscrtna predstavlja procijenjene vrijednosti iz modela m1, siva puna krivulja predstavlja procijenjene vrijednosti temeljem modela m2, dok crna s točkama za m3. Ono što se uočava da model m3 najbolje opisuje kretanje stvarnih vrijednosti BDP-a.

- f) U najboljem modelu iz prethodnih postupaka testirajmo značajnost sezonskih binarnih varijabli nakon krize pomoću Waldova testa, uz $\alpha = 5\%$.

Kako je najbolji model m3, za njega provodimo Waldov test kako slijedi. Testiraju se sljedeće hipoteze:

$$H_0 \dots \mathbf{R} \cdot \boldsymbol{\beta} = \mathbf{q}, \quad \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_8 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad H_0 \dots \mathbf{R} \cdot \boldsymbol{\beta} \neq \mathbf{q},$$

gdje je matrica \mathbf{R} formata $(3,8)$, jer se testira hipoteza $\beta_1 = \beta_2 = \beta_3 = 0$, stoga se radi o 3 jednakosti, dok je 8 parametara u samome modelu. Rezultati su predočeni na slici 1.17. Kako test veličina iznosi $W = 17,71$, s pripadajućom p -vrijednosti gotovo 0, uz razinu značajnosti od 5%, odbacuje se hipoteza da niti jedan sezonski učinak u razdoblju u vrijeme i nakon krize nije značajan.

```

ogr<-c("binQ1=0","binQ2=0","binQ3=0")
linearHypothesis(m3,ogr,type="Chisq")

## Linear hypothesis test
##
## Hypothesis:
## binQ1 = 0
## binQ2 = 0
## binQ3 = 0
##
## Model 1: restricted model
## Model 2: bdp ~ bin + trend + I(bin * bin2) + I(trend * bin2)
##
##   Res.Df      RSS Df Sum of Sq    F    Pr(>F)
## 1      96 117242531
## 2      93  74618789   3  42623742 17.708 3.57e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Slika 1.17. Waldov test

- g) Temeljem najboljeg modela iz postupka b) prognozirajmo vrijednost BDP-a za iduća dva tromjesečja.

Temeljem naredbi i ispisa na slici 1.18., uočavamo da se prognozira vrijednost BDP-a u trećem tromjesečju 2020. godine u iznosu 16305,51 milijun eura, dok u četvrtom tromjesečju 16205,13 milijuna eura.

```

bin<-seasonaldummy(ts(103:104,start=c(2020,3),frequency=4))
bin2<-c(1,1)
trend<-ts(103:104,start=c(2020,3),frequency = 4)
novo<-cbind(bin,bin2,trend)
predict(m3,newdata=novo)

##           1           2
## 16305.51 16205.13

```

Slika 1.18. Prognoziranje BDP-a

- h) Testirajmo postoji li problem autokorelacije reziduala do zaključno 2. reda u modelu iz prethodnih postupaka, pomoću Ljung-Boxova testa, $\alpha = 5\%$.

Testiraju se sljedeće hipoteze: $H_0: \rho_1 = \rho_2 = 0$, $H_1: \exists \rho_j \neq 0, j \in \{1, 2\}$, dok test veličina iznosi $Q = 107,8$, a pripadajuća p -vrijednost je gotovo 0, stoga se razinu značajnosti od 5% odbacuje hipoteza da ne postoji autokorelacija reziduala do zaključno 2. reda.

```

Box.test(resid(m3),lag=2,type="Ljung-Box")

##
## Box-Ljung test
##
## data:  resid(m3)
## X-squared = 107.8, df = 2, p-value < 2.2e-16

```

Slika 1.19. Test autokorelacije reziduala u modelu m3

- i) Testirajmo postoji li problem heteroskedastičnosti reziduala pomoću Breusch-Paganova testa, $\alpha = 5\%$.

```
bptest<-lm(residuals(m3)^2~bin+trend+I(bin*bin2)+I(trend*bin2))
nobs(bptest)*(summary(bptest))$r.squared

## [1] 9.123594

1-pchisq(nobs(bptest)*(summary(bptest))$r.squared,8)

## [1] 0.3319773
```

Slika 1.20. Test heteroskedastičnosti reziduala u modelu m3

Testiraju se sljedeće hipoteze: $H_0: \alpha_1 = \dots = \alpha_8 = 0$, $H_1: \exists \alpha_j \neq 0, j \in \{1, \dots, 8\}$, dok test veličina iznosi $LM = 9,12$, a pripadajuća p -vrijednost⁴ jednaka je 0,33, stoga se razinu značajnosti od 5% ne odbacuje hipoteza da ne postoji heteroskedastičnost reziduala.

- j) S obzirom na rezultate u postupcima e) i f), izvršimo odgovarajuću korekciju standardnih pogrešaka i potom usporedimo ishode t -testova za slučaj bez korekcija i s uključenim korekcijama.

Kako se odbacila hipoteza o nepostojanju autokorelacije reziduala, dok se ne odbacuje hipoteza o heteroskedastičnosti reziduala u modelu, dovoljno je primijeniti Whiteovu korekciju standardnih pogrešaka procjenitelja (vidjeti odjeljak 4.3.3.1. u prvome dijelu Odabrane teme primijenjene ekonometrije: Linearni regresijski model), što je prikazano na slici 1.21.

```
#White korekcija:
library(car)
mat<-hccm(m3, type="hc0")
#t-test:
library(lmtest)
```

Slika 1.21. Whiteova korekcija standardnih pogrešaka procjenitelja u modelu

Sada se rezultati t -testova bez korekcija i uz Whiteovu korekciju uspoređuju na slici 1.22. Ono što se uočava da jedino dolazi do promjene u ishodu t -testa za binarnu varijablu $Q2_t$, koja je bez korekcija (gornji panel) značajna na razini značajnosti od 5% (jer je pripadajuća p -vrijednost 0,025), dok uz Whiteovu korekciju postaje neznačajna pri istoj razini značajnosti (pripadajuća p -vrijednost iznosi 0,06).

⁴ Naredba `pchisq(x,ss)` računa površinu ispod hi-kvadrat krivulje od vrijednosti 0 do vrijednosti x , za ss broj stupnjeva slobode. Kako se provodi test koji se odnosi na desni rep te distribucije, ta površina se računa kao ukupna površina koja iznosi 1, umanjena za vrijednost površine od 0 do x . Zato u ispisu izračuna p -vrijednosti za ovaj test pišemo `1-pchisq(x,ss)`. Vrijednost x u slučaju ovog testa iznosi N (broj opservacija) pomnožen s koeficijentom determinacije pomoćne regresijske jednadžbe, dok je 8 stupnjeva slobode. Detalje o testu vidjeti u prvoj publikaciji Odabrane teme primijenjene ekonometrije: Linearni regresijski model.

```

coefstest(m3)
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4301.1210   308.5172 13.9413 < 2.2e-16 ***
## binQ1         -349.0826   347.1488 -1.0056 0.3172314
## binQ2          794.5935   348.8911  2.2775 0.0250489 *
## binQ3         2059.0161   354.4793  5.8086 8.733e-08 ***
## trend          85.7162     4.6482 18.4405 < 2.2e-16 ***
## I(bin * bin2)Q1 -415.8556  471.6874 -0.8816 0.3802479
## I(bin * bin2)Q2 -990.6087  474.7555 -2.0866 0.0396641 *
## I(bin * bin2)Q3 -1844.1775  478.2465 -3.8561 0.0002121 ***
## I(trend * bin2)  28.7454     7.0128  4.0990 8.868e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

coefstest(m3,vcov=mat)
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4301.1210   313.8567 13.7041 < 2.2e-16 ***
## binQ1         -349.0826   407.7553 -0.8561  0.39414
## binQ2          794.5935   417.5905  1.9028  0.06016 .
## binQ3         2059.0161   371.2916  5.5455 2.733e-07 ***
## trend          85.7162     4.4314 19.3429 < 2.2e-16 ***
## I(bin * bin2)Q1 -415.8556  473.4488 -0.8784  0.38202
## I(bin * bin2)Q2 -990.6087  472.0818 -2.0984  0.03858 *
## I(bin * bin2)Q3 -1844.1775  451.0374 -4.0887 9.207e-05 ***
## I(trend * bin2)  28.7454     5.8540  4.9104 3.873e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Slika 1.22. Usporedba t -testova za model bez korekcija (gornji panel) i model s korekcijama (donji panel)

1.5. Pitanja za ponavljanje

1. Što je to vremenski niz? Od kojih komponenti se sastoji? Ukratko opišite svaku.
2. Koja je razlika između aditivnog i multiplikativnog modela za vremenski niz?
3. Kada koristimo binarne varijable za vremenske nizove?
4. Zapišite regresijski model u kojemu zavisna varijabla ovisi o dvije nezavisne varijable, pri čemu se radi o vremenskim nizovima, a učinak prve nezavisne varijable na zavisnu se mijenja nakon vremenskog trenutka $t = 20$. Ukupno je prikupljeno 100 opservacija za svaku varijablu.
5. Zapišite regresijski model u kojemu zavisna varijabla ovisi o dvije nezavisne varijable, pri čemu se radi o vremenskim nizovima, a učinak obje nezavisne varijable na zavisnu se mijenja nakon vremenskog trenutka $t = 20$. Ukupno je prikupljeno 100 opservacija za svaku varijablu.
6. Zapišite regresijski model u kojemu zavisna varijabla ovisi o dvije nezavisne varijable, pri čemu se radi o vremenskim nizovima, a učinak prve nezavisne varijable na zavisnu se mijenja nakon vremenskog trenutka $t = 20$, dok se učinak druge nezavisne varijable na zavisnu mijenja nakon vremenskog trenutka $t = 40$. Ukupno je prikupljeno 100 opservacija za svaku varijablu.
7. Zapišite regresijski model u kojemu se vremenski niz y_t opisuje s konstantom i sezonskim binarnim varijablama, a radi se o mjesečnim podacima.
8. Interpretirajte konstantu i koeficijent uz binarnu varijablu za veljaču u modelu u prethodnom zadatku.
9. Zapišite regresijski model u kojemu se vremenski niz y_t opisuje bez konstante i samo pomoću sezonskih binarnih varijabli, a radi se o mjesečnim podacima.
10. Interpretirajte koeficijent uz binarnu varijablu za veljaču i srpanj u modelu iz prethodnog zadatka.
11. Koja je razlika u interpretacijama procijenjenih parametara modela u zadatku 7 i zadatku 9?
12. Učitajte datoteku „bdp_agregati.txt“ u RStudio. Datoteka sadrži tromjesečne podatke o kretanju uvoza, izvoza, osobne potrošnje i državne potrošnje (indeksi, 2010=100, engl. *chain linked volumes*) od prvog tromjesečja 1995. do drugog tromjesečja 2020. godine (ukupno 102 opservacije za svaku varijablu).
 - a) Definirajte vremenski niz „izvoz“ pomoću naredbe `ts(...)` te grafički predočite i komentirajte sam niz.
 - b) Generirajte sezonske binarne varijable u okviru paketa `forecast`, varijablu linearni trend koja poprima vrijednosti 1, 2, ... do 102, te binarnu varijablu koja poprima vrijednost 0 za prve 52 opservacije te vrijednost 1 za preostale do kraja uzorka. Posljednja binarna varijabla odnosi se na razdoblje prije krize (1. tromjesečje 2008. godine) i razdoblje u vrijeme i nakon krize.
 - c) Procijenite sljedeća 4 regresijska modela: u prvome neka kretanje izvoza ovisi samo o sezonskim binarnim varijablama; u drugome o linearnom trendu i binarnim varijablama; u trećemu modelu proširite drugi s članom interakcije binarne varijable koja se odnosi na krizu i trendom; i četvrti model neka je drugi proširen i za interakciju sezonskih binarnih varijabli s binarnom varijablom koja se odnosi na krizu. Spremite sva 4 modela i pomoću naredbe `stargazer(...)` spojite rezultate ispisa.
 - d) Interpretirajte sve parametre u sva 4 modela.
 - e) Usporedite adekvatnom mjerom iz ispisa u postupku c) koji model je najprikladniji za opisivanje izvoza u promatranom razdoblju.

- f) Grafički predočite procijenjene vrijednosti iz svakog od 4 modela i usporedite s vremenskom dinamikom originalnog niza.
- g) U modelu m1 testirajte značajnost sezonskih učinaka pomoću Waldova testa. Zapišite hipoteze testa, test veličinu i donesite zaključak temeljem $\alpha = 5\%$.
- h) U modelu m3 testirajte značajnost binarne varijable koja se odnosi na krizu pomoću Waldova testa. Zapišite hipoteze testa, test veličinu i donesite zaključak temeljem $\alpha = 5\%$.
- i) Prognozirajte vrijednost izvoza za idućih osam tromjesečja za model m4.
- j) Izračunajte stopu rasta izvoza koristeći formulu (1.6) i potom formulu (1.8), predočite ih grafički i usporedite.

13. Ponovite postupke a) do i) iz prethodnog zadatka za varijablu uvoz iz iste datoteke.

Rješenja

Zadatak 4)

$$y_t = a + bx_{1t} + cx_{2t} + dx_{1t}D_t + e_t, D_t = \begin{cases} 1 & \text{za } t = 1 \text{ do } 20 \\ 0 & \text{za } t = 21 \text{ do } 100 \end{cases}$$

Zadatak 5)

$$y_t = a + bx_{1t} + cx_{2t} + dx_{1t}D_t + ex_{2t}D_t + u_t, D_t = \begin{cases} 1 & \text{za } t = 1 \text{ do } 20 \\ 0 & \text{za } t = 21 \text{ do } 100 \end{cases}$$

Zadatak 6)

$$y_t = a + bx_{1t} + cx_{2t} + dx_{1t}D_{1t} + ex_{2t}D_{2t} + u_t,$$

$$D_{1t} = \begin{cases} 1 & \text{za } t = 1 \text{ do } 20 \\ 0 & \text{za } t = 21 \text{ do } 100 \end{cases}, D_{2t} = \begin{cases} 1 & \text{za } t = 1 \text{ do } 40 \\ 0 & \text{za } t = 41 \text{ do } 100 \end{cases}$$

Zadatak 7)

$$y_t = \beta_0 + \beta_1 D_{1t} + \beta_2 D_{2t} + \dots + \beta_{11} D_{11t} + \varepsilon_t,$$

$$D_{1t} = \begin{cases} 1, & \text{za siječanj} \\ 0, & \text{inače} \end{cases}, D_{2t} = \begin{cases} 1, & \text{za veljaču} \\ 0, & \text{inače} \end{cases}, \dots, D_{11t} = \begin{cases} 1, & \text{za studeni} \\ 0, & \text{inače} \end{cases},$$

Zadatak 8)

Prosječna razina pojave y u prosincu iznosi β_0 jedinica.

Zadatak 9)

$$y_t = \beta_1 D_{1t} + \beta_2 D_{2t} + \dots + \beta_{12} D_{12t},$$

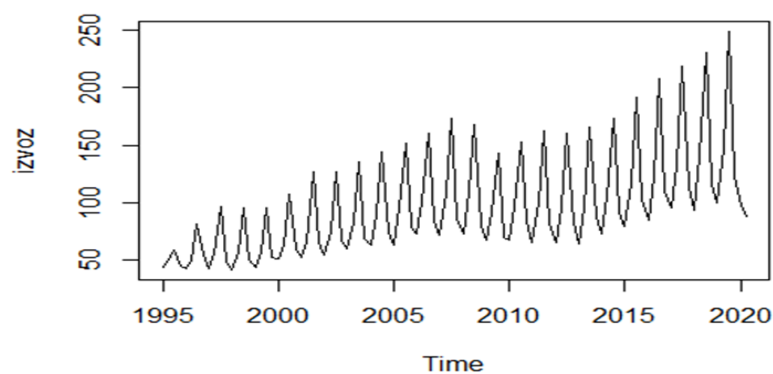
$$D_{1t} = \begin{cases} 1, & \text{za siječanj} \\ 0, & \text{inače} \end{cases}, D_{2t} = \begin{cases} 1, & \text{za veljaču} \\ 0, & \text{inače} \end{cases}, \dots, D_{12t} = \begin{cases} 1, & \text{za prosinac} \\ 0, & \text{inače} \end{cases}.$$

Zadatak 10)

Vrijednost pojave y u veljači u prosjeku iznosi β_2 jedinica, dok u srpnju iznosi u prosjeku β_7 jedinica.

Zadatak 11)

U zadatku 7 radi se o usporedbi koliko je u prosjeku vrijednost pojave veća ili manja u odnosu na prosječnu vrijednost u prosincu, dok se u zadatku 9 radi o prosječnoj vrijednosti pojave u određenom mjesecu.

Zadatak 12)

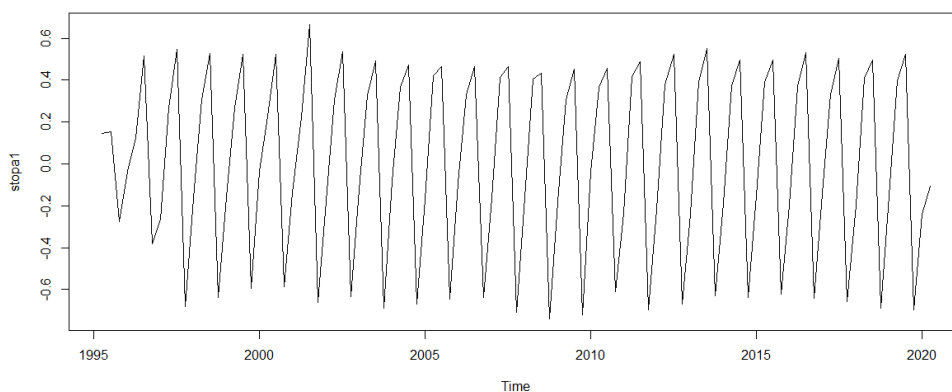
Dependent variable:				
	izvoz			
	(1)	(2)	(3)	(4)
## trend		0.878*** (0.052)	1.276*** (0.131)	0.953*** (0.069)
## binarneQ1	-12.813 (8.564)	-11.934*** (4.349)	-11.202*** (4.151)	0.474 (4.957)
## binarneQ2	12.756 (8.564)	12.756*** (4.349)	13.249*** (4.147)	18.606*** (4.932)
## binarneQ3	1.996*** (8.648)	72.874*** (4.392)	73.120*** (4.186)	64.183*** (4.907)
## I(trend * bin2)			-0.316*** (0.096)	
## I(binarne * bin2)Q1				-24.668*** (6.379)
## I(binarne * bin2)Q2				-11.699* (6.379)
## I(binarne * bin2)Q3				18.262*** (6.393)
## Constant	79.136*** (6.115)	33.455*** (4.125)	24.624*** (4.761)	29.576*** (4.490)
## Observations	102	102	102	102
## R2	0.537	0.882	0.894	0.914
## Adjusted R2	0.523	0.877	0.888	0.908

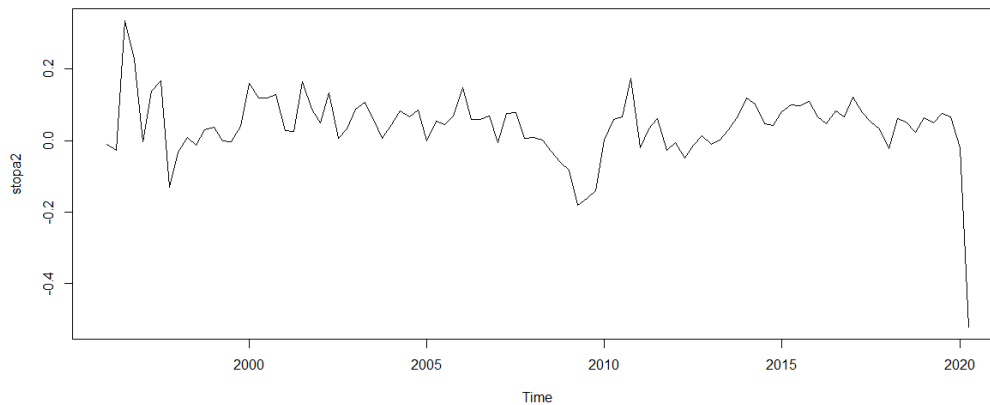
```
## Residual Std. Error 30.574 (df=98)      15.526 (df=97)      14.796 (df=96)      13.428 (df=94)
## F Statistic 37.934*** (df=3; 98) 181.070*** (df=4; 97) 161.664*** (df=5; 96) 143.429*** (df=7; 94)
## =====
## Note: *p<0.1; **p<0.05; ***p<0.01
```

```
## Linear hypothesis test
##
## Hypothesis:
## binarneQ1 = 0
## binarneQ2 = 0
## binarneQ3 = 0
##
## Model 1: restricted model
## Model 2: izvoz ~ binarne
##
## Res.Df  RSS Df Sum of Sq    F  Pr(>F)
## 1    101 197981
## 2     98  91605  3    106376 37.934 2.34e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Linear hypothesis test
##
## Hypothesis:
## I(trend * bin2) = 0
##
## Model 1: restricted model
## Model 2: izvoz ~ trend + binarne + I(trend * bin2)
##
## Res.Df  RSS Df Sum of Sq    F  Pr(>F)
## 1     97 23383
## 2     96 21017  1    2366.1 10.808 0.001413 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 2020 Q1 97.4 101.64331
## 2020 Q2 87.7 133.69715
## 2020 Q3 NA 210.18858
## 2020 Q4 NA 128.69650
## 2021 Q1 NA 105.45565
## 2021 Q2 NA 137.50950
## 2021 Q3 NA 214.00093
## 2021 Q4 NA 132.50884
## 2022 Q1 NA 109.26800
## 2022 Q2 NA 141.32184
```





1.6. Odabrane metode desezoniranja vremenskih nizova

Kada u ekonomskim analizama sezonalna komponenta nije od glavnog interesa kao što je bila u prethodnom odjeljku, već se želi sezonalnost ukloniti iz vremenskog niza, kako bi se tako „očišćen“ niz dalje koristio u regresijskoj ili nekoj drugoj analizi, postoje određene metode desezoniranja, koje su već ugrađene u većini matematičkih i statističkih programa, kojima se sezona isključuje iz originalnog vremenskog niza. U nastavu slijedi opis nekoliko odabranih metoda desezoniranja vremenskih nizova. Često se u literaturi na engleskom jeziku desezoniranje spominje kao dekomponiranje (engl. *decomposition*).

1.6.1. Metoda pomičnih prosjeka

Pomični prosjek (engl. MA, *moving average*) je aritmetička sredina M uzastopnih vrijednosti članova vremenskog niza. Niz pomičnih prosjeka je vremenski niz koji ima manju varijaciju u odnosu na izvorni vremenski niz, kako bi se na taj način vremenski niz izgadio. Jednostavni pomični prosjek računa se kao jednostavna aritmetička sredina M uzastopnih vrijednosti članova niza. Ako se računa temeljem neparnog broja članova ($M = 2m + 1$), računa se na sljedeći način:

$$\tilde{y}_t = \frac{\sum_{s=-m}^m y_{t+s}}{M}, t \in \{m+1, m+2, \dots, T-m\}, \quad (1.18)$$

dok se za paran broj članova ($M = 2m$), potrebno je provesti postupak centriranja, pa se pomični prosjek računa formulom:

$$\tilde{y}_t = \frac{0.5y_{t-m} + \sum_{s=-m+1}^{m-1} y_{t+s} + 0.5y_{t+m}}{M}, t \in \{m+1, m+2, \dots, T-m\}. \quad (1.19)$$

Primjer 1.6.

Učitana je datoteka „bdp.txt“ u RStudio (naslov 1.4.). Datoteka sadrži tromjesečne podatke o kretanju BDP-a (milijuni eura) Hrvatske od prvog tromjesečja 1995. do drugog tromjesečja 2020. godine. Pomoću metode pomičnih prosjeka desezonirajmo niz BDP, te grafički usporedimo originalni niz s desezoniranim.

Slika 1.23. Prikazuje naredbe potrebne za provođenje desezoniranja metodom pomičnih prosjeka u okviru paketa forecast. Naredba `ma(...)`:

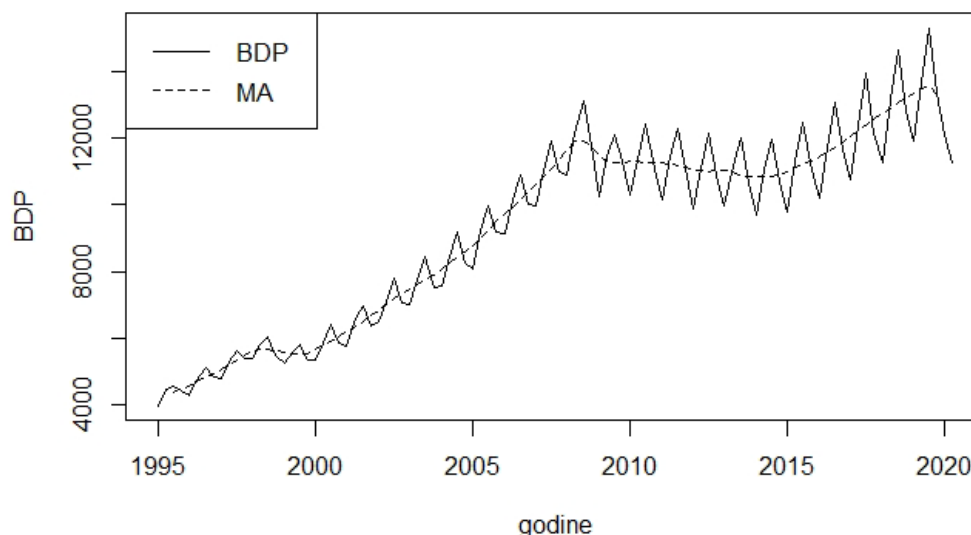
```
ma(bdp,4,centre=T)
```

se koristi za izračun pomičnih prosjeka, pri čemu kao prvi argument pišemo naziv niza koji se desezonira ovim putem, potom je drugi argument vrijednost M , i argument $centre=T$ se odnosi na postupak centriranja ako je M paran broj. U slučaju BDP-a smo odabrali $M = 4$, što znači da će se na temelju formule (1.19) računati pomični prosjek. Dodatno su prikazane naredbe za grafički prikaz vrijednosti BDP-a i desezoniranih vrijednosti na istome grafičkom prikazu. Na slici 1.24., gdje se uspoređuju originalni niz (BDP) i desezonirani (MA), uočava se kako je novi niz „izgladen“, tj. u njemu nema prisutnih sezonskih oscilacija.

```
bdp<-read.table("bdp.txt",header=T,sep="\t")
bdp<-ts(bdp$BDP,start=c(1995,1),frequency = 4)

library(forecast)
bdp.pp<-ma(bdp,4,centre = T)
plot(bdp)
lines(bdp.pp,col="black",lty="dashed")
legend("topleft", legend=c("BDP", "MA"),col=c("black", "black"), lty=c(1,2))
```

Slika 1.23. Provođenje desezoniranja pomičnim prosjecima



Slika 1.24. Usporedba originalnog niza (BDP) i desezoniranog (MA)

1.6.2. Klasično desezoniranje (dekomponiranje) vremenskog niza

U odjeljku 1.1. vremenski niz se pomoću aditivnog, odnosno multiplikativnog modela raščlanio na nekoliko komponenti: trend i cikličku, sezonsku i slučajnu komponentu. Ako grafički prikaz sezonskih amplituda upućuju na njihovu nepromjenjivost tijekom vremena, u tom slučaju se može koristiti aditivni pristup, dok se promjenjivost amplituda modelira multiplikativnim pristupom.

Klasično desezoniranje sastoji se od nekoliko koraka, ako se pretpostavi multiplikativni oblik modela:

1. Najprije se izračunaju pomični prosjeci (odjeljak 1.6.1.) \tilde{y}_t
2. Potom se izračunaju omjeri y_t / \tilde{y}_t , koje nazivamo prvim procjenama sezonskih faktora. Kako oni variraju, računa se njihova aritmetička sredina, na način da se faktori za isti mjesec/ tromjesečje svake godine koriste u izračunu u idućem koraku.
3. Izračun sezonskog faktora I_t , kao aritmetička sredina omjera y_t / \tilde{y}_t za isti mjesec/ tromjesečje iz svake godine.
4. Izračun desezoniranih vrijednosti y_t / I_t .
5. Izračun rezidualnih faktora, na način da se desezonirane vrijednosti podijele s pomičnim prosjecima kao procjenama trenda, tj. izračun $y_t / (I_t \tilde{y}_t)$.

U slučaju aditivnog modela se umjesto omjera koriste razlike u prethodno navedenim postupcima.

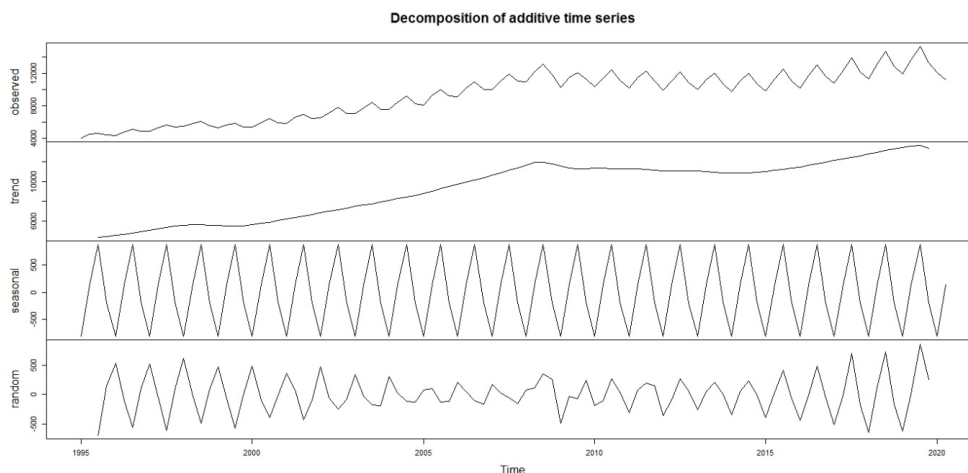
Primjer 1.7.

Učitana je datoteka „bdp.txt“ u RStudio (naslov 1.4.). Datoteka sadrži tromjesečne podatke o kretanju BDP-a (milijuni eura) Hrvatske od prvog tromjesečja 1995. do drugog tromjesečja 2020. godine. Pomoću aditivne, a potom multiplikativne metode klasičnog dekomponiranja vremenskog niza desezonirajmo niz BDP, te potom postupak predočimo grafički. Interpretirajmo sezonske faktore I_t .

Desezoniranje pomoću klasične metode dekomponiranja jednostavno je u RStudiu. Koristi se naredba `decompose(...)`, pri čemu se pod vrstu dekomponiranja piše `additive` ili `multiplicative` (vidjeti sliku 1.25). Grafički prikaz vremenskih nizova temeljem provedenih dekomponiranja prikazani su na slikama 1.26. (aditivni model) te 1.27. (multiplikativni model). Obje slike predočavaju u prvome dijelu originalni niz (observed), potom trend i cikličku komponentu (trend), sezonsku komponentu (seasonal) i na kraju slučajnu komponentu (random). Uočava se u slučajnoj komponenti kako i dalje postoje određene oscilacije koje upućuju na sezonalnosti koje nisu izbačene u obje metode. Razlog leži u činjenici što su obje metode relativno jednostavne, no postoje složenije, ali uspješnije metode desezoniranja (vidjeti EC, 2018 priručnik o desezoniranju).

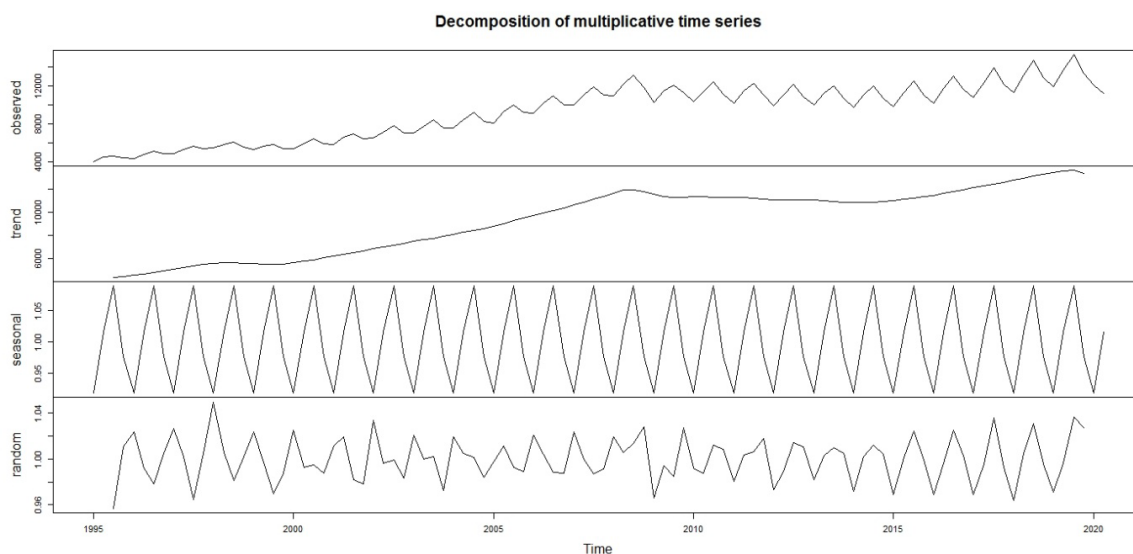
```
plot(decompose(bdp,type="additive"))
plot(decompose(bdp,type="multiplicative"))
```

Slika 1.25. Desezoniranje klasičnim pristupom dekomponiranja



Slika 1.26. Grafički prikaz desezoniranja klasičnim pristupom dekomponiranja, aditivna metoda

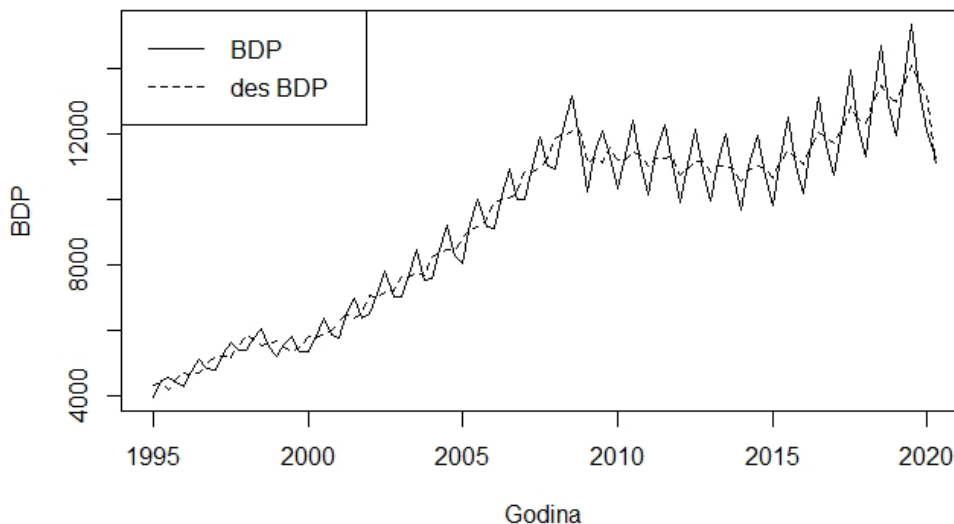
Naredbe potrebne za grafičko predočavanje originalnog niza zajedno s desezoniranim, u slučaju multiplikativne metode, predočene su na slici 1.28., dok slika 1.29. uspoređuje dva spomenuta niza. Uočava se kako je desezonirani niz više „izglađen“ u odnosu na originalni niz.



Slika 1.27. Grafički prikaz desezoniranja klasičnim pristupom dekomponiranja, multiplikativna metoda

```
dekompozicija<-decompose(bdp,type="multiplicative")
desezonirani.bdp<-bdp/dekompozicija$seasonal
plot(bdp,xlab="Godina",ylab="BDP")
lines(desezonirani.bdp,col="black",lty="dashed")
legend("topleft", legend=c("BDP", "des BDP"),col=c("black", "black"), lty=c(1,2))
```

Slika 1.28. Naredbe za usporedbu originalnog (BDP) i desezoniranog niza (des BDP)



Slika 1.29. Grafička usporedba originalnog (BDP) i desezoniranog niza (des BDP)

Konačno, za izračun sezonskih faktora i njihovu interpretaciju, koriste se naredbe predočene na slici 1.30. Naredba \$seasonal predočava faktore I_t iz postupka dekompozicije, pri čemu se uočava da su sljedeći faktori za 1., 2., 3. i 4. tromjesečje: 0,919, 1,015, 1,088 i 0,977. To znači da je vrijednost promatranog niza zbog utjecaja sezone na razini od 91,9%, 101,5%, 108,8% i

97,7% razine promatranog niza kada nema učinka sezone u BDP-u. Da bi se to lakše interpretiralo, računaju se stope na način da se od vrijednosti faktora oduzme 100% ili vrijednost 1 i pomnoži se sve sa 100% kako bi se interpretiralo u postocima. Stoga su sljedeće stope za svako tromjesečje redom: -8,07, 1,55, 8,81 i -2,28. To znači da je zbog utjecaja sezone u promatranom razdoblju BDP u prvom tromjesečju za 8,07% manji u odnosu na vrijednost kada nema utjecaja sezone; u drugom tromjesečju je zbog utjecaja sezone BDP za 1,55% veći u odnosu na vrijednost BDP-a bez učinka sezone (desezonirane vrijednosti BDP-a); u trećem tromjesečju je zbog učinka sezone BDP za 8,81% veći u odnosu na njegovu vrijednost kada nema sezone te je u četvrtom tromjesečju za 2,28% manja vrijednost BDP-a u odnosu na desezoniranu vrijednost.

```
dekompozicija$seasonal[1:4]
## [1] 0.9192557 1.0154841 1.0881047 0.9771555
stope<-(dekompozicija$seasonal-1)*100
stope[1:4]
## [1] -8.074431 1.548415 8.810466 -2.284449
```

Slika 1.30. Izračun sezonskih faktora i stopa promjena

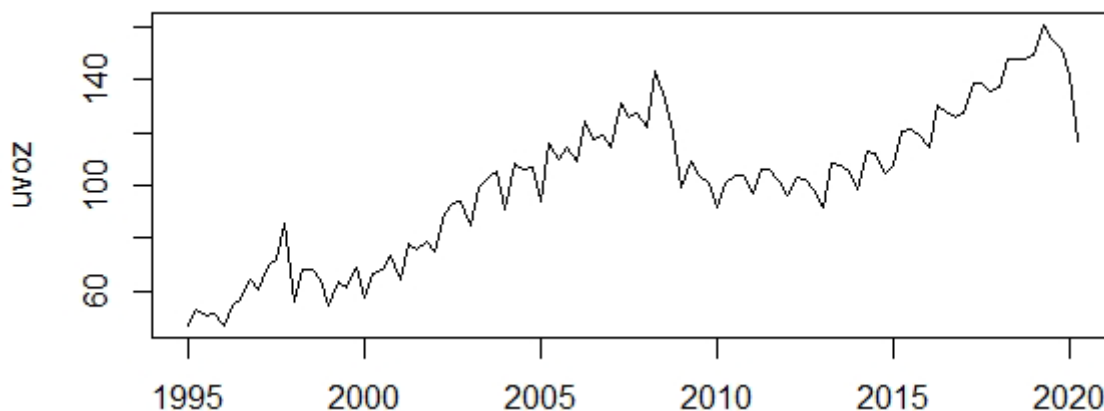
1.6.3. Primjer

Učitajmo datoteku „**bdp_agregati.txt**“ u RStudio. Datoteka sadrži tromjesečne podatke o kretanju uvoza (indeks, 2010=100, engl. *chain linked volumes*) od prvog tromjesečja 1995. do drugog tromjesečja 2020. godine (ukupno 102 opservacije). Definirajmo vremenski niz „uvoz“ pomoću naredbe `ts(...)` te grafički predočimo i komentirajmo sam niz. Pomoću metode pomičnih prosjeka desezonirajmo niz uvoz, te grafički usporedimo originalni niz s desezoniranim nizom. Pomoću multiplikativne metode klasičnog dekomponiranja vremenskog niza desezonirajmo niz uvoz, te potom postupak predočimo grafički. Interpretirajmo sezonske faktore za slučaj multiplikativne metode za prvo i drugo tromjesečje.

Najprije su na slici 1.31. predočene naredbe za učitavanje datoteke, definiranja vremenskog niza i grafičkog prikaza niza uvoz. Na slici 1.32. uočava se da postoji trend povećanja niza uvoz kroz vrijeme, ali je prisutan i učinak sezone, potrebno je niz desezonirati.

```
bdp<-read.table("bdp_agregati.txt",header=T,sep="\t")
uvoz<-ts(bdp$uvoz,start=c(1995,1),frequency = 4)
desezonirani.uvoz<-uvoz/decompose(uvoz,type="multiplicative")$seasonal
plot(uvoz)
```

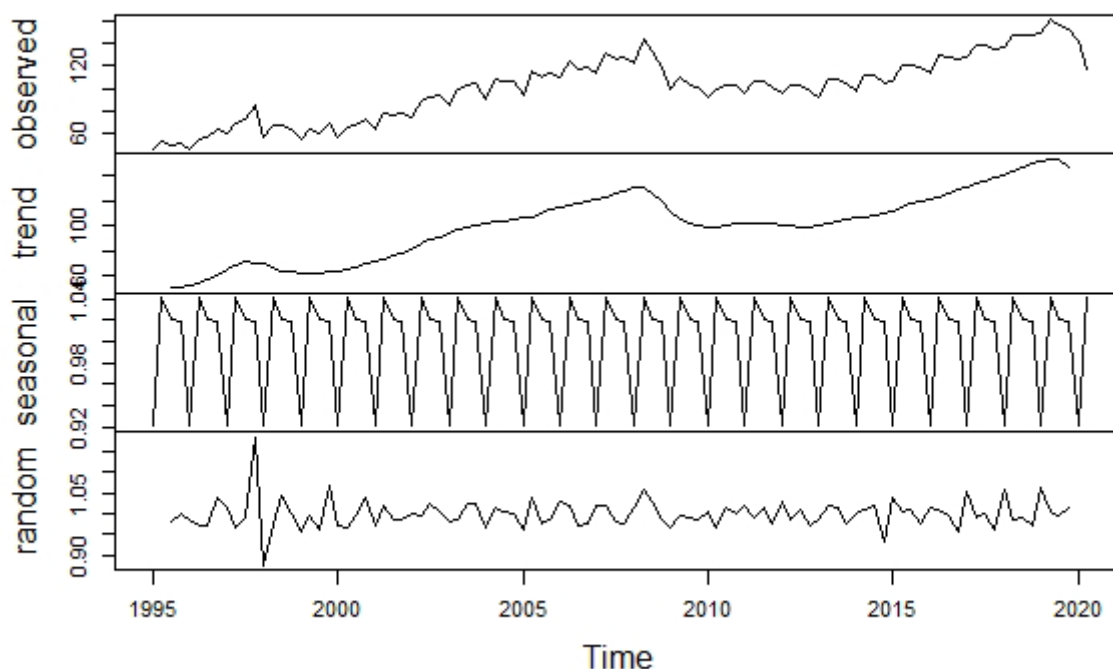
Slika 1.31. Učitavanje podataka, definiranje vremenskog niza i grafičko predočavanje



Slika 1.32. Grafički prikaz niza uvoz

Provedena je dekompozicija niza uvoz i slika 1.33. prikazuje grafički sve komponente samog niza dobivene naredbom `plot(decompose(uvoz,type="multiplicative"))`. Uočava se kako je sezonska komponenta isključena iz trend-cikličke komponente, te je niz izglađen. Konačno, kako bismo usporedili originalni niz s desezoniranim bolje, naredbe na slici 1.34. predočavaju što je potrebno upisati u RStudiju da bi se dobila slika 1.35., na kojoj je vidljivo da desezonirani niz nema sezonske oscilacije u odnosu na originalni niz.

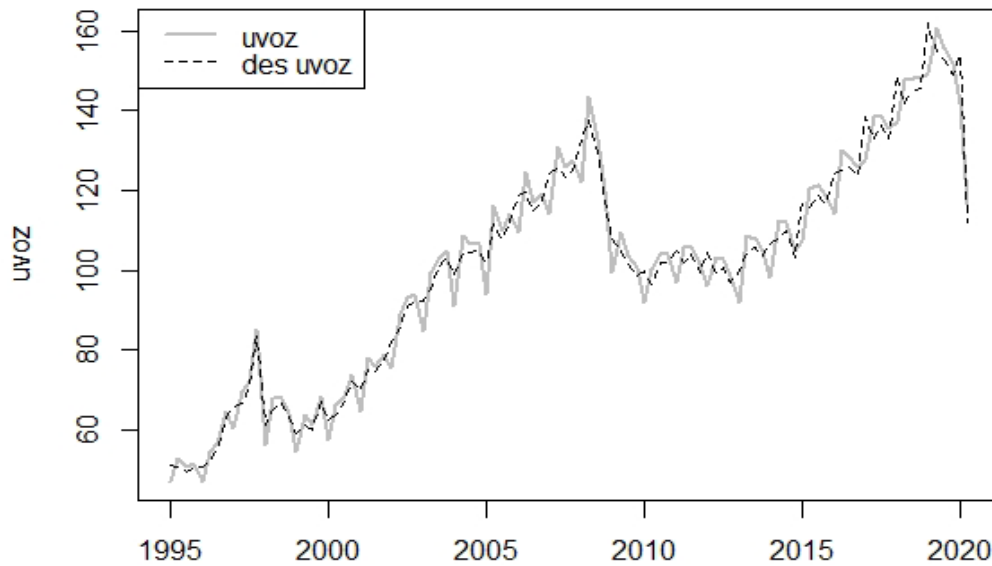
Decomposition of multiplicative time series



Slika 1.33. Grafički prikaz desezoniranja klasičnim pristupom dekomponiranja, multiplikativna metoda

```
plot(uvoz, xlab="Godina", ylab="uvoz", col="grey", lwd=2)
lines(desezonirani.uvoz, col="black", lty="dashed")
legend("topleft", legend=c("uvoz", "des uvoz"), col=c("grey", "black"), lty=c(2,2), lwd=c(2,1))
```

Slika 1.34. Naredbe za predočavanje slike 1.35.



Slika 1.35. Usporedba originalnog niza uvoz i desezoniranog niza (des uvoz)

Konačno, slika 1.36. predočava naredbe za izračun sezonskih faktora i stopa promjene koje su redom: $-7,96$, $4,11$, $1,99$ i $1,86$. Zbog učinka sezone je vrijednost uvoza u prvom tromjesečju za $7,96\%$ manja u odnosu kada nema učinka sezone, za $4,11\%$ veća u drugome tromjesečju u odnosu kad nema učinka sezone, itd.

```
dekompozicija$seasonal[1:4]
## [1] 0.9203829 1.0410812 1.0199325 1.0186033
stope<-(dekompozicija$seasonal-1)*100
stope[1:4]
## [1] -7.961708 4.108123 1.993255 1.860330
```

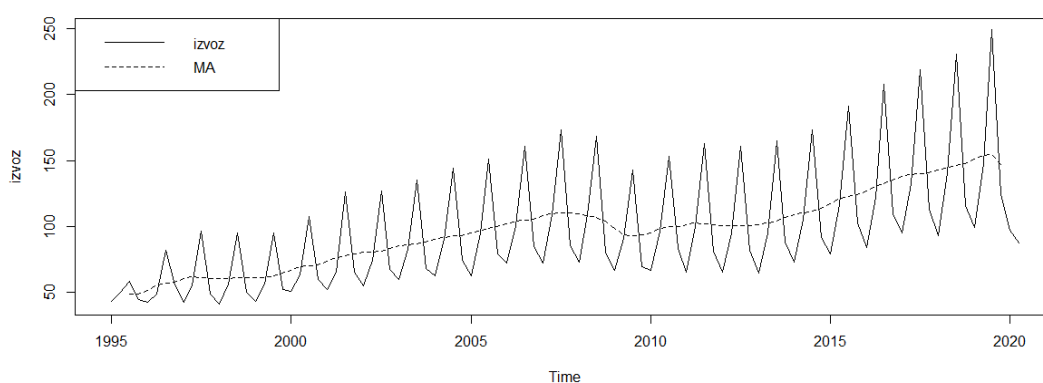
Slika 1.36. Izračun sezonskih faktora i stopa promjena

1.7. Pitanja za ponavljanje

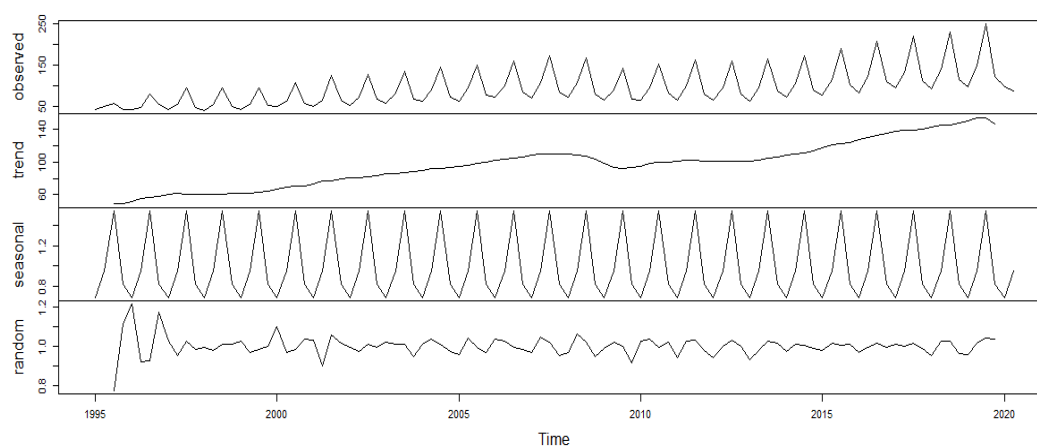
1. Ukratko opišite metodu pomičnih prosjeka za desezoniranje vremenskog niza.
2. Ukratko opišite korake metode klasičnog dekomponiranja (desezoniranja) vremenskog niza mutliplikativnom metodom.
3. Ukratko opišite korake metode klasičnog dekomponiranja (desezoniranja) vremenskog niza aditivnom metodom.
4. Učitajte datoteku „**bdp_agregati.txt**“ u RStudio. Datoteka sadrži tromjesečne podatke o kretanju izvoza (indeks, 2010=100, engl. *chain linked volumes*) od prvog tromjesečja 1995. do drugog tromjesečja 2020. godine (ukupno 102 opservacije). Definirajte vremenski niz „uvoz“ pomoću naredbe `ts(...)` te grafički predočite i komentirajte sam niz. Pomoću metode pomičnih prosjeka desezonirajte niz izvoz, te grafički usporedimo originalni niz s desezoniranim nizom. Pomoću multiplikativne metode klasičnog dekomponiranja vremenskog niza desezonirajte niz izvoz, te potom postupak predočite grafički. Interpretirajte sezonske faktore za slučaj multiplikativne metode za prvo i drugo tromjesečje.

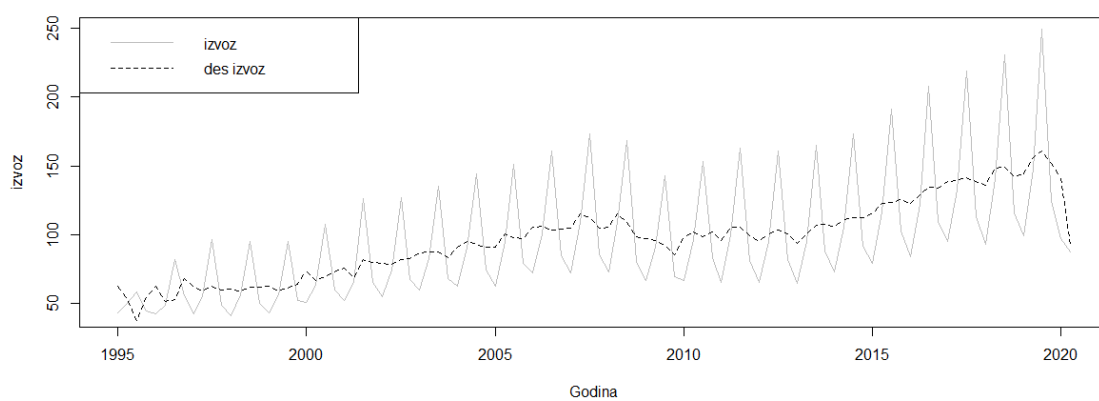
Rješenja

Zadatak 4)



Decomposition of multiplicative time series





```
#stope
dekompozicija$seasonal[1:4]

## [1] 0.6876825 0.9488591 1.5466701 0.8167883

stope<-(dekompozicija$seasonal-1)*100
stope[1:4]

## [1] -31.231747 -5.114095 54.667012 -18.321170
```

1.8. Model višestruke linearne regresije s vremenskim nizovima

Ako se razmatra regresijski model s jednom zavisnom varijablom te k nezavisnih, pri čemu se radi o vremenskim nizovima, model višestruke linearne regresije zapisuje se ovako:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t. \quad (1.20)$$

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$ su nepoznati parametri koje je potrebno procijeniti. Ako se model (1.20) želi procijeniti za sva opažanja u slučaju vremenskih nizova, $t \in \{1, 2, \dots, T\}$, tada se razmatra sustav od T jednadžbi:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_k x_{1k} + \varepsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_k x_{2k} + \varepsilon_2 \\ &\vdots \\ y_T &= \beta_0 + \beta_1 x_{T1} + \beta_2 x_{T2} + \dots + \beta_k x_{Tk} + \varepsilon_T \end{aligned} \quad (1.21)$$

Model je moguće zapisati u matricnoj formi:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1.22)$$

gdje je $\mathbf{y} \in \mathbb{R}^T$ vektor stupac čiji elementi su opažanja zavisne varijable, $\mathbf{X} \in \mathcal{M}_{T,k+1}$ je matrica čiji prvi stupac čine jedinice, a ostale stupce čine vrijednosti opažanja nezavisnih varijabli, $\boldsymbol{\beta} \in \mathbb{R}^{k+1}$ je vektor stupac nepoznatih parametara, dok je $\boldsymbol{\varepsilon} \in \mathbb{R}^T$ vektor stupac slučajne varijable.

Pretpostavke modela višestruke linearne regresije su obrađene u odjeljku 2.2.2. u prvom dijelu, te se jedino oznake koje su se do sada odnosile na presječne podatke mijenjaju u oznake za vremenske nizove:

1. Linearnost modela – pretpostavlja se linearna veza između zavisne i nezavisnih varijabli. Drugim riječima se može reći da je y linearna kombinacija k nezavisnih varijabli.
2. Egzogenost podataka u matrici \mathbf{X} : $E(\varepsilon_t | \mathbf{X}) = 0, \forall t$.
3. Greška relacije u prosjeku ne utječe na zavisnu varijablu: $E(\varepsilon_t) = 0, \forall t$, tj. $E(y_t | x_1, x_2, \dots, x_k) = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk}, \forall t$.
4. Varijanca greške relacije je konstantna (homoskedastična): $Var(\varepsilon_t) = Var(\varepsilon_t | \mathbf{X}) = \sigma^2, \forall t$
5. Nezavisnost slučajne varijable, tj. nekoreliranost: $E(\varepsilon_t, \varepsilon_s) = Cov(\varepsilon_t, \varepsilon_s) = Cov(\varepsilon_t, \varepsilon_s | x_i) = 0$ za $t \neq s$.
6. Slučajna varijabla normalno je distribuirana: $\varepsilon_t \sim N(0, \sigma^2), \forall t$.
7. Varijable x_j su međusobno nezavisne, što znači da vrijedi $r(\mathbf{X}'\mathbf{X})^{-1} = r(\mathbf{X}) = k + 1$, tj. postoji inverz $(\mathbf{X}'\mathbf{X})^{-1}$.

Dakle, ako se razmatra regresijska analiza u slučaju vremenskih nizova, potrebno je imati pristup kao što je obrađivano u prethodnim poglavljima vezanim uz procjenu parametara modela, interpretaciju dobivenih rezultata, kao i testiranje pretpostavki modela. Ono što je još potrebno analizirati za vremenske nizove jest (ne)stacionarnost vremenskih nizova, što se obrađuje u narednim odjeljcima. Dodatno, u odjeljku 1.10. obrađuju se dinamički modeli, stoga se model (1.20) naziva statički, s obzirom da se pretpostavlja da promjene nezavisnih varijabli

u vremenskom trenutku ili razdoblju t ima učinak na zavisnu varijablu u istom trenutku ili razdoblju t .

1.9. Stacionarni vremenski nizovi

Stohastički proces je familija vremenski indeksiranih slučajnih varijabli, u oznaci: $\{y_t, t \in \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}\}$. Sam proces $\{y_t, t \in \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}\}$ je populacija, a ako se razmatra vremenski niz $\{y_1, y_2, \dots, y_T\}$, radi se o realizaciji, odnosno uzorku iz populacije. Stohastički proces može biti stacionaran ili nestacionaran.

1.9.1. Definiranje stacionarnosti

Stohastički proces je stacionaran ako se njegova statistička svojstva ne mijenjaju tijekom vremena (Guidolin i Pedio, 2019). Formalno se stacionarnost definira pomoću striktno (engl. *strict stationarity*) i slabe (engl. *weak stationarity*) stacionarnosti.

1.9.1.1. Striktna stacionarnost

Proces je **striktno stacionaran** ako je zajednička distribucija F varijable za svaki niz vremenskih indeksa t_1, t_2, \dots, t_n jednaka zajedničkoj distribuciji varijable za niz svih vremenskih indeksa $t_{1+k}, t_{2+k}, \dots, t_{n+k}$, gdje je p pomak:

$$F_{y_{t_1}, y_{t_2}, \dots, y_{t_n}}(y_1, y_2, \dots, y_T) = F_{y_{t_1+p}, y_{t_2+p}, \dots, y_{t_n+p}}(y_1, y_2, \dots, y_T), \quad (1.23)$$

tj. drugim riječima: striktno stacionaran vremenski niz $\{y_t, t \in \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}\}$ ima sljedeća svojstva: slučajne varijable y_t su identično distribuirane, te dva slučajna vektora $[y_t, y_{t+1}]'$ i $[y_{t+p}, y_{t+1+p}]'$ imaju istu zajedničku distribuciju za bilo koji t i bilo koji p .

Striktna stacionarnost stoga zahtjeva da svi momenti distribucije budu nepromjenjivi s obzirom na vrijeme, što je veoma jaka pretpostavka za neki vremenski niz. U empirijskim primjenama se stoga češće primjenjuje pojam slabe stacionarnosti, što znači da relacija (1.23) vrijedi samo za $p \leq 1$.

1.9.1.2. Slaba stacionarnost

Proces je **slabo stacionaran** (ili **kovarijančno stacionaran**, engl. *covariance stationary*) ako prvi i drugi momenti ne ovise o vremenu, tj. za svaki $t \in \{1, 2, \dots\}$, vrijede sljedeći uvjeti:

$$E[y_t] = \mu < \infty, \quad (1.24)$$

$$E[(y_t - \mu)^2] = \sigma^2 < \infty, \quad (1.25)$$

$$E[(y_t - \mu)(y_{t-h} - \mu)] = \gamma_h < \infty, \quad (1.26)$$

$h \in \{\dots, -2, -1, 1, 2, \dots\}$. Izraz (1.24) čitamo da se radi o **konstantnoj i konačnoj očekivanoj vrijednosti** (ne ovisi o t stoga nema indeksa t uz μ), (1.25) znači da je **varijanca nepromjenjiva**, stoga ponovno nema indeksa t uz σ^2 , te (1.26) znači da je **kovarijanca između y_t i y_{t-h} jedino funkcija pomaka h** , a ne ovisi o t .

Kako je kovarijanca funkcija pomaka h u (1.26), i **korelacija ρ_h** između y_t i y_{t-h} je također funkcija samo pomaka h :

$$\rho_h = \frac{\gamma_h}{\sigma^2} < \infty, h \in \{\dots, -2, -1, 1, 2, \dots\}. \quad (1.27)$$

1.9.1.3. Autokorelacijska funkcija

Niz koeficijenata kovarijanca γ_h , $h \in \{\dots, -2, -1, 1, 2, \dots\}$, se definira kao autokovarijančna funkcija (engl. *autocovariance function*) za neki vremenski niz, no kako je lakše interpretirati koeficijente korelacije (jer su intervalu $[-1, 1]$), definira se **autokorelacijska funkcija** (engl. *autocorrelation function*, ACF) kao niz koeficijenata autokorelacije ρ_h , $h \in \{\dots, -2, -1, 1, 2, \dots\}$. Koeficijent autokorelacije reda h definira se formulom (1.27), koju možemo pisati i ovako:

$$\rho_h = \frac{\text{Cov}(y_t, y_{t-h})}{\text{Var}(y_t)} = \frac{\gamma_h}{\sigma^2}, h \in \{\dots, -2, -1, 1, 2, \dots\} \quad (1.28)$$

Stacionarne vremenske nizove karakteriziraju autokovarijančna i autokorelacijska funkcija koje brzo teže prema vrijednosti 0, pa se autokorelacijska funkcija nekog vremenskog niza koristi u određivanju (ne)stacionarnosti niza.

Za dani uzorak opservacija varijable y_t **procjenjuje se autokorelacijska funkcija** (engl. *sample autocorrelation function*, SACF), nazvana i empirijska autokorelacijska funkcija (engl. *empirical ACF*):

$$\hat{\rho}_h = \frac{\sum_{t=h+1}^T (y_t - \hat{\mu})(y_{t-h} - \hat{\mu})}{\sum_{t=1}^T (y_t - \hat{\mu})^2} = \frac{\hat{\gamma}_h}{\hat{\sigma}^2}, h \in \{\dots, -2, -1, 1, 2, \dots\}, \quad (1.29)$$

gdje $\hat{\mu}$ i $\hat{\sigma}^2$ predstavljaju prosječnu vrijednost i procjenu varijance od y_t . Procjena autokorelacijske funkcije nam daje informaciju o stupnju linearne povezanosti između članova procesa y_t za različite pomake h . Također se može reći da $\hat{\rho}_h$ daje informaciju o memoriji procesa (kolika je prediktivna snaga prethodnih ili lagiranih vrijednosti varijable y_t za predviđanje budućih vrijednosti).

Stacionarne procese karakteriziraju sljedeći koeficijenti autokorelacije: autokorelacija nultog

reda jednaka je 1, tj. $\hat{\rho}_0 = \frac{\sum_{t=1}^T (y_t - \hat{\mu})(y_t - \hat{\mu})}{\sum_{t=1}^T (y_t - \hat{\mu})^2} = \frac{\hat{\sigma}^2}{\hat{\sigma}^2} = 1$; koeficijent autokorelacije reda h je po

apsolutnoj vrijednosti manji ili jednak 1, tj. $|\hat{\rho}_h| \leq 1$, svojstvo simetričnosti: $\hat{\rho}_h = \hat{\rho}_{-h}$, te je

autokorelacijska funkcija pozitivno semidefinitna: $\sum_i \sum_j \alpha_i \alpha_j \rho_{i-j} \geq 0$, gdje su α_i i α_j neki

realni brojevi. Prva tri navedena svojstva su nužan uvjet da bi se radilo o autokorelacijskoj funkciji stohastičkog procesa, dok je svojstvo pozitivne semidefinitnosti dovoljan uvjet (vidjeti Wei, 2005).

Grafički prikaz empirijske autokorelacijske funkcije nazivamo korelogram.

Primjer 1.8.

Učitajmo datoteku „bdp_agregati.txt“ u RStudio. Datoteka sadrži tromjesečne podatke o kretanju uvoza, izvoza, osobne potrošnje i državne potrošnje (indeksi, 2010=100, engl. *chain linked volumes*) od prvog tromjesečja 1995. do drugog tromjesečja 2020. godine (ukupno 102 opservacije za svaku varijablu). Predočimo korelogoram empirijske autokorelacijske funkcije do pomaka 20 ($h = 20$) za varijablu državna potrošnja. Nakon toga predočimo korelogram empirijske autokorelacijske funkcije za stopu rasta državne potrošnje (izračunatu formulom (1.8)) i komentirajmo oba korelograma. Procijenimo koeficijente autokorelacije do pomaka 8 za oba niza i usporedimo ih.

Slika 1.37. predočava procjenu koeficijenata autokorelacijske funkcije u okviru paketa forecast, pomoću naredbe `Acf(...)`:

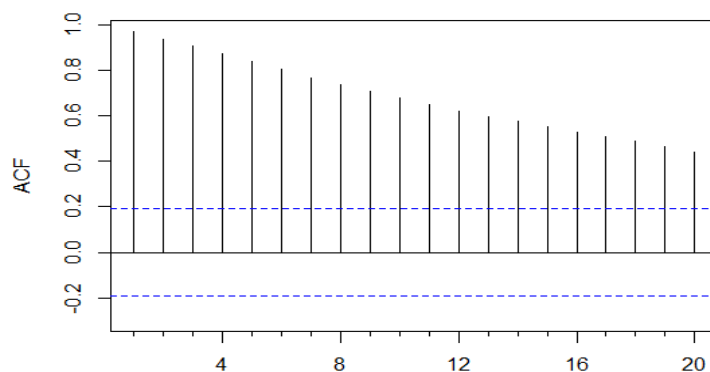
```
Acf(drzavna,lag.max=20)
```

gdje kao prvi argument upisujemo naziv niza, te kao drugi argument maksimalni pomak h do kojeg procjenjujemo koeficijente (`lag.max=20`), koji su prikazani na korelogramu na slici 1.38. Naredbom `Acf(...)` odmah se crta korelogram, stoga za same vrijednosti procijenjenih koeficijenata je potrebno tu naredbu spremiti (`acf.drz<-...`), te pomoću naredbe `acf.drz[,"acf"]` prikazati procijenjene vrijednosti na dnu ispisa slike 1.37. Napomenimo da je u okviru spomenute naredbe definirano da je prvi pomak u suštini pomak $h = 0$ u prethodnom tekstu. Zato je u ispisu prikazano devet pomaka (`[,9]`) jer je taj posljednji zapravo osmi.

```
drzavna<-ts(bdp$drzava,start=c(1995,1),frequency = 4)
library(forecast)
acf.drz<-Acf(drzavna,lag.max = 20)
acf.drz<-Acf(drzavna,lag.max = 8)
acf.koeficijenti<-matrix(acf.drz[,"acf"])
t(acf.koeficijenti)

##      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]      [,8]      [,9]
## [1,] 1 0.9683576 0.933438 0.9035819 0.8719975 0.8377952 0.8015676 0.766668 0.734566
```

Slika 1.37. Izračun koeficijenata autokorelacijske funkcije



Slika 1.38. Korelogram autokorelacijske funkcije za državnu potrošnju

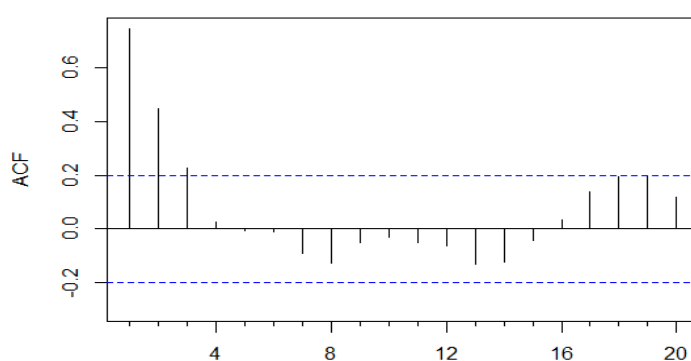
Uočava se da su vrijednosti procijenjenih koeficijenata autokorelacijske funkcije veliki (blizu jedinične vrijednosti) te se sporo smanjuju (vidljivo na slici 1.38). To je indikacija da se radi o nestacionarnom nizu, s obzirom da je spomenuto da stacionarne nizove karakterizira autokorelacijska funkcija koja brzo iščezava prema vrijednosti 0.

S druge strane, izračun stope rasta državne potrošnje i procjena koeficijenata autokorelacijske funkcije za stopu rasta predočen je na slici 1.39., uz korelogram predočen na slici 1.40. Ovdje se uočava da se vrijednosti procijenjenih koeficijenata autokorelacijske funkcije veoma brzo smanjuju prema 0. Ovo je karakteristika stacionarnih vremenskih nizova.

```
stopa<-diff(log(drzavna),4)
acf.drz<-Acf(drzavna,lag.max = 20)
acf.drz<-Acf(drzavna,lag.max = 8)
acf.koeficijenti<-matrix(acf.drz[["acf"]])
t(acf.koeficijenti)
```

##	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]
## [1,]	1	0.7458894	0.447207	0.2266629	0.02491716	-0.005796723	-0.009734007	-0.09135203	-0.1279069

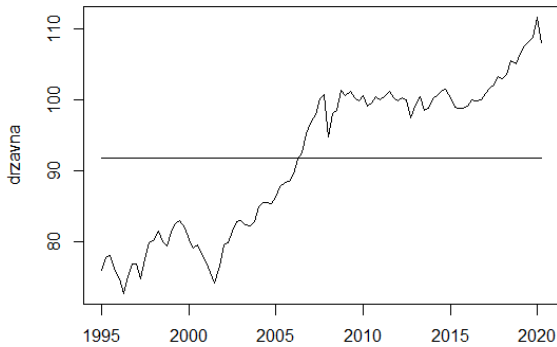
Slika 1.39. Izračun koeficijenata autokorelacijske funkcije



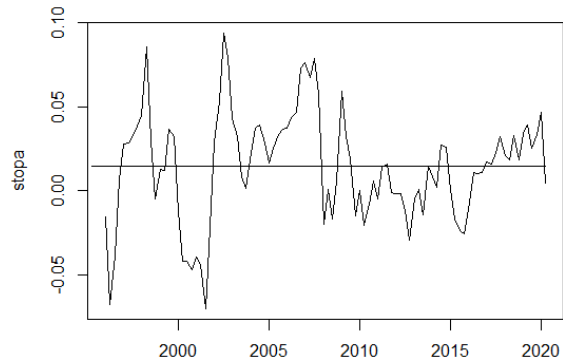
Slika 1.40. Korelogram autokorelacijske funkcije za stopu rasta

Ako bismo grafički predočili vremensku dinamiku državne potrošnje i usporedili s prosječnom vrijednosti u promatranom uzorku, kao i za stopu rasta (slika 1.41.), uočava se da je svojstvo (1.24) konstantnog očekivanja narušeno za državnu potrošnju, jer prosjek izračunat za cijeli uzorak nije reprezentativan (panel (a)), dok se većina vrijednosti stope rasta (panel (b)) mijenjaju oko prosjeka izračunatog u promatranom razdoblju.

(a) Usporedba državne potrošnje i prosjeka



(b) Usporedba stope rasta državne potrošnje i prosjeka



Slika 1.41. Usporedba vremenskih nizova i njihovih prosječnih vrijednosti

Ako je y_t niz identično i nezavisno distribuiran proces, s konačnom varijancom, za veliki uzorak je procijenjen niz koeficijenata autokorelacije normalno distribuiran, s očekivanjem 0 i varijancom $1/T$, T je veličina uzorka (Guidolin i Pedio, 2019). Upravo je ovo korisno za provođenje testiranja **hipoteze o značajnosti koeficijenata autokorelacijske funkcije**. 95% interval pouzdanosti (engl. *confidence interval*) se procjenjuje jednostavno, tako da se procijene donja i gornja granica:

$$\hat{\rho}_h - 1,96/\sqrt{T} \text{ i } \hat{\rho}_h + 1,96/\sqrt{T}. \quad (1.30)$$

Za testiranje neznačajnosti koeficijenta autokorelacije reda h (pretpostavka da je $\rho_h = 0$), razmatraju se donja i gornja granica $0 - 1,96 / \sqrt{T}$ i $0 + 1,96 / \sqrt{T}$. Upravo su donja i gornja granica $0 - 1,96 / \sqrt{T}$ i $0 + 1,96 / \sqrt{T}$ predložene na grafičkim prikazima korelograma na slikama 1.38. i 1.40., iscrtanim linijama. Samo testiranje značajnosti koeficijenata autokorelacije obrađeno je u odjeljku 1.9.1.5.

1.9.1.4. Parcijalna autokorelacijska funkcija

Osim autokorelacijske funkcije, možemo promatrati i **parcijalnu autokorelacijsku funkciju** (engl. *partial autocorrelation function*, PACF, *sample PACF*, SACF). Radi se o nizu koeficijenata autokorelacije između y_t i y_{t-h} koji se definiraju na način da se **ukloni linearna ovisnost** o varijablama $y_{t+1}, y_{t+2}, \dots, y_{t-h+1}$:

$$\phi_h = \text{Corr}(y_t, y_{t-h} \mid y_{t+1}, y_{t+2}, \dots, y_{t-h+1}), h \in \{\dots, -2, -1, 1, 2, \dots\}. \quad (1.31)$$

Procjena koeficijenata parcijalne autokorelacije može se dobiti metodom najmanjih kvadrata tako da se procijeni model (Guidolin i Pedio, 2019):

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_h y_{t-h} + \varepsilon_t, \quad (1.32)$$

pri čemu se procjena $\hat{\phi}_h$ koeficijenata parcijalne autokorelacije interpretira kao „dodana“ prediktivna moć pomaka h kada se kontrolira za preostale pomake $y_{t+1}, y_{t+2}, \dots, y_{t-h+1}$. Koeficijente dobivene procjenom modela (1.32) nazivamo empirijskim koeficijentima parcijalne autokorelacije (engl. *empirical PACF*, *sample PACF*, *SACF*). **Korelogram također može predočavati i empirijske koeficijente parcijalne autokorelacije.**

Primjer 1.9.

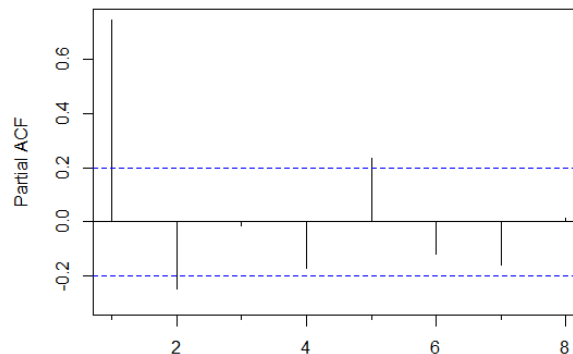
U prethodnom primjeru 1.8.) je učitana datoteka „bdp_agregati.txt“ u RStudio, pri čemu koristimo varijable državna potrošnja i stopa rasta iz tog primjera. Predočimo korelogram empirijske parcijalne autokorelacijske funkcije do pomaka 20 ($h = 20$) za varijablu državna potrošnja. Nakon toga predočimo korelogram empirijske parcijalne autokorelacijske funkcije za stopu rasta državne potrošnje (izračunatu formulom (1.8)) i komentirajmo oba korelograma. Procijenimo koeficijente parcijalne autokorelacije do pomaka 8 za oba niza i usporedimo ih.

Slike 1.42. i 1.44. predočavaju izračune koeficijenata parcijalne autokorelacijske funkcije za oba niza, dok slike 1.43. i 1.45. predočavaju odgovarajuće korelograme. Naredba `Pacf(...)` se koristi veoma slično kao naredba `Acf(...)` u prethodnom primjeru. Uočava se da se sada koeficijenti parcijalne autokorelacije veoma brzo smanjuju prema vrijednostima 0 za oba niza.

```
pacf.drz<-Pacf(drzavna,lag.max=20)
pacf.drz<-Pacf(stopa,lag.max=8)
pacf.koeficijenti<-matrix(pacf.drz[["acf"]])
t(pacf.koeficijenti)
```

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]
[1,]	0.7458894	-0.2460144	-0.01311639	-0.1695555	0.2353942	-0.1208143	-0.1609853	0.01225773

Slika 1.42. Izračun koeficijenata parcijalne autokorelacijske funkcije

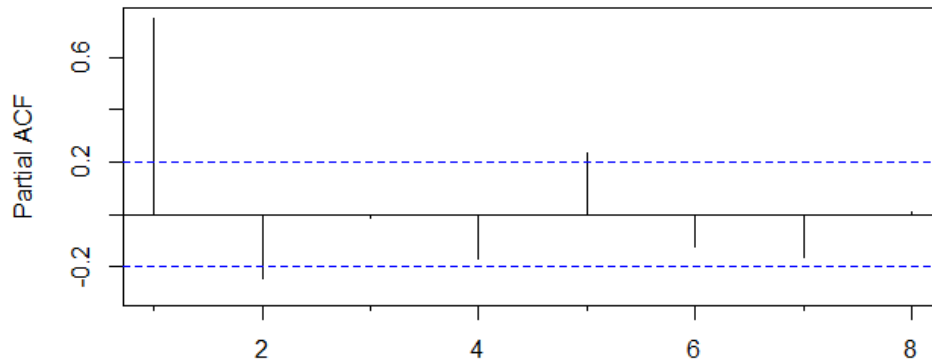


Slika 1.43. Korelogram parcijalne autokorelacijske funkcije za državnu potrošnju

```
acf.drz<-Pacf(stopa,lag.max=20)
acf.drz<-Pacf(stopa,lag.max=8)
acf.koeficijenti<-matrix(acf.drz[["acf"]])
t(acf.koeficijenti)
```

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]
[1,]	0.7458894	-0.2460144	-0.01311639	-0.1695555	0.2353942	-0.1208143	-0.1609853	0.01225773

Slika 1.44. Izračun koeficijenata parcijalne autokorelacijske funkcije



Slika 1.45. Korelogram parcijalne autokorelacijske funkcije za stopu rasta državne potrošnje

Slično kao i za testiranje značajnosti koeficijenata autokorelacije, može se provesti i testiranje značajnosti koeficijenata parcijalne autokorelacije, tako da se razmatra interval neznačajnosti $[0 - 1,96 / \sqrt{T}, 0 + 1,96 / \sqrt{T}]$. I ovaj interval je predočen na slikama 1.43. i 1.45. Samo testiranje značajnosti koeficijenata autokorelacije prikazano je u idućem odjeljku.

1.9.1.5. Testiranje značajnosti koeficijenata autokorelacije

Skupni test značajnosti koeficijenata autokorelacije do zaključno reda m može se provesti pomoću **Boxova i Pierceova (1970) testa**. Nulta hipoteza testa pretpostavlja da prvih m koeficijenata autokorelacije nije značajno, dok alternativna pretpostavlja suprotno:

$$\begin{aligned} H_0 : \rho_1 = \rho_2 = \dots = \rho_m = 0 \\ H_0 : \exists \rho_j \neq 0, j \in \{1, 2, \dots, m\} \end{aligned} \quad (1.33)$$

Test veličinu definirali su izrazom:

$$Q(m) = T \sum_{h=1}^m \hat{\rho}_h^2 \sim \chi^2(m), \quad (1.34)$$

gdje je T veličina uzorka, a test veličina $Q(m)$ asimptotski slijedi hi-kvadrat distribuciju s m stupnjeva slobode ako je nulta hipoteza točna. Ako je $\chi^2(m) > \chi_\alpha^2(m)$, nulta hipoteza se odbacuje, ili ako je p -vrijednost $< \alpha$, dok u slučaju $\chi^2(m) < \chi_\alpha^2(m)$ ne odbacujemo nultu hipotezu, odnosno ako je p -vrijednost $> \alpha$. **Nedostatak ovog testa** je što nije pogodan za male uzorke, jer prečesto vodi do krivog zaključka.

Stoga su **Ljung i Box (1978)** razvili varijantu testa koja je robusnija za male uzorke, čija je test veličina sljedeća:

$$Q^*(m) = T(T+2) \sum_{h=1}^m \frac{\hat{\rho}_h^2}{T-h} \sim \chi^2(m). \quad (1.35)$$

Primjer 1.10.

U primjeru 1.8. je učitana datoteka „bdp_agregati.txt“ u RStudio, pri čemu koristimo varijable državna potrošnja i stopa rasta iz tog primjera. Provedimo Box-Pierecov test o značajnosti koeficijenata autokorelacije do zaključno 3. reda ($m = 3$) za obje varijable, a potom i Ljung-Boxov test za isti red. Interpretirajmo rezultate za $\alpha = 5\%$.

Za varijablu državna potrošnja proveden je Box-Pierecov test kako slijedi:

$$H_0 : \rho_1 = \rho_2 = \rho_3 = 0$$

$$H_0 : \exists \rho_j \neq 0, j \in \{1, 2, 3\}, Q(3) = 267,8, p\text{-vrijednost} \approx 0 < \alpha,$$

te Ljung-Boxov test:

$$H_0 : \rho_1 = \rho_2 = \rho_3 = 0$$

$$H_0 : \exists \rho_j \neq 0, j \in \{1, 2, 3\}, Q^*(3) = 278,4, p\text{-vrijednost} \approx 0 < \alpha.$$

Stoga uz razinu značajnosti od 5% odbacujemo hipotezu da ne postoji autokorelacija među članovima procesa državna potrošnja do zaključno 3. reda u oba testa. Ispis oba testa predočen je na slici 1.46., gdje se koristi naredba `Box.test(...)`:

```
Box.test(drzavna,lag=3,type="Box-Pierce"),
```

u okviru koje se kao prvi argument navodi niz za koji se test provodi, potom pomak za koji se maksimalno testira nulta hipoteza testa ($\text{lag}=3$), te se kao tip testa navodi naziv samog testa (`type="Box-Pierce"` ili „Ljung-Box“).

```
Box.test(drzavna,lag=3,type="Box-Pierce")
##
## Box-Pierce test
##
## data:  drzavna
## X-squared = 267.8, df = 3, p-value < 2.2e-16

Box.test(drzavna,lag=3,type="Ljung-Box")
##
## Box-Ljung test
##
## data:  drzavna
## X-squared = 278.4, df = 3, p-value < 2.2e-16
```

Slika 1.46. Box-Pierceov i Ljung-Boxov test za državnu potrošnju

Za varijablu stopa rasta državne potrošnje proveden je Box-Pierecov test (slika 1.47.):

$$H_0 : \rho_1 = \rho_2 = \rho_3 = 0$$

$$H_0 : \exists \rho_j \neq 0, j \in \{1, 2, 3\}, Q(3) = 79,16, p\text{-vrijednost} \approx 0 < \alpha,$$

te Ljung-Boxov test:

$$H_0 : \rho_1 = \rho_2 = \rho_3 = 0$$

$$H_0 : \exists \rho_j \neq 0, j \in \{1, 2, 3\}, Q^*(3) = 81,93, p\text{-vrijednost} \approx 0 < \alpha .$$

Stoga uz razinu značajnosti od 5% odbacujemo hipotezu da ne postoji autokorelacija među članovima procesa stopa rasta državne potrošnje do zaključno 3. reda u oba testa.

```
Box.test(stopa, lag=3, type="Box-Pierce")
##
## Box-Pierce test
##
## data: stopa
## X-squared = 79.157, df = 3, p-value < 2.2e-16

Box.test(stopa, lag=3, type="Ljung-Box")
##
## Box-Ljung test
##
## data: stopa
## X-squared = 81.925, df = 3, p-value < 2.2e-16
```

Slika 1.47. Box-Pierceov i Ljung-Boxov test za stopu rasta državne potrošnje

Ovakvi ishodi testa su rezultat koeficijenata autokorelacije koji su značajno različiti od vrijednosti 0, što je vidljivo na korelogramima na slikama 1.43. i 1.45.

1.9.2. Čisti slučajni proces

Čisti slučajni proces (engl. *white noise process*, WN) ili **proces bijelog šuma** (inovacijski proces) je niz jednako distribuiranih, nekoreliranih slučajnih varijabli čije je očekivanje jednako 0 i varijanca je konstantna σ^2 :

$$\begin{aligned} & \{\varepsilon_t, t \in \{\dots, -2, -1, 0, 1, 2, \dots\}\}, \\ & E(\varepsilon_t) = 0, \\ & Var(\varepsilon_t) = \sigma^2, \\ & Cov(\varepsilon_t, \varepsilon_s) = 0, t \neq s. \end{aligned} \tag{1.36}$$

Ako se radi i o normalno distribuiranom procesu u (1.36), nazivamo ga **Gaussov čisti slučajni proces**. Dodatno, čisti slučajni proces je primjer striktno stacionarnog procesa. S obzirom na svojstvo kovarijance u (1.36), buduće članove čistog slučajnog procesa nije moguće predviđati temeljem prethodnih vrijednosti.

U regresijskoj analizi se upravo većina pretpostavki (vidjeti odjeljke 2.1.2. i 2.2.2. u prvom dijelu) odnosi na slučajnu varijablu, da se radi o Gaussovom čistom slučajnom procesu, stoga se testiranja pretpostavki u poglavlju 4 u prvome dijelu odnose na pretpostavke o definiranom čistom slučajnom procesu.

S obzirom na samu definiciju čistog slučajnog procesa, autokovarijance, autokorelacije i parcijalne autokorelacije reda h iznose:

$$\gamma_h = \begin{cases} \sigma^2, & h = 0 \\ 0, & h \neq 0 \end{cases}, \rho_h = \begin{cases} 1, & h = 0 \\ 0, & h \neq 0 \end{cases}, \phi_h = \begin{cases} 1, & h = 0 \\ 0, & h \neq 0 \end{cases}. \quad (1.37)$$

To znači da će korelogram za (parcijalnu) autokorelacijsku funkciju pokazivati vrijednost 1 za $h = 0$, a za ostale pomake su koeficijenti (parcijalne) autokorelacije jednaki 0. U empirijskim analizama se razmatra neznačajnost (parcijalnih) koeficijenata autokorelacije, s obzirom da te vrijednosti teorijski iznose 0. Za neznačajnost spomenutih koeficijenata se razmatra nalaze li se u spomenutom intervalu $[-1,96 / \sqrt{T}, 1,96 / \sqrt{T}]$.

Primjer 1.11.

Otvorimo praznu skriptu u RStudiju i generirajmo⁵ niz „epsilon“ sa 100 vrijednosti prema Gaussovom čistom slučajnom procesu i varijancom 1. Definirajmo da se radi o vremenskom nizu s mjesečnim frekvencijama, s početkom rujana 2011. godine. Grafički predočimo niz „epsilon“ i komentirajmo njegovu stacionarnost prema definiciji stacionarnog procesa. Potom predočimo korelograme autokorelacijske i parcijalne autokorelacijske funkcije do 10. reda i komentirajmo ih. Konačno, provedimo Box-Pierceov test o značajnosti koeficijenata autokorelacije do zaključno 10. reda ($m = 10$), a potom i Ljung-Boxov test za isti red. Interpretirajmo rezultate za $\alpha = 5\%$.

Slika 1.48. predočava naredbe za generiranje Gaussovog čistog slučajnog procesa, `rnorm(...)`, gdje se slovo *r* odnosi na *random* (slučajna varijabla) a `norm` na normalno distribuiranu varijablu. U okviru naredbe `rnorm(...)`:

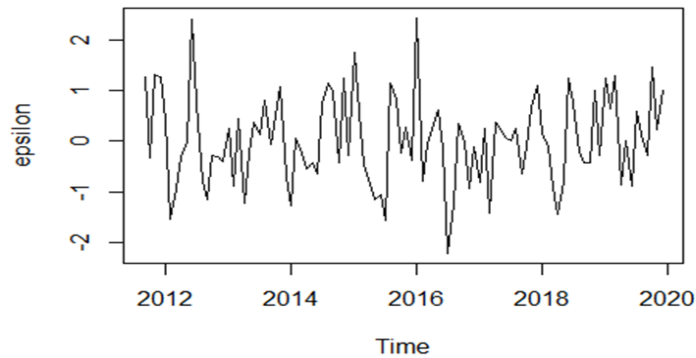
```
rnorm(100,0,1)
```

potrebno je kao prvi argument definirati koliko članova niza proces ima (100), potom očekivanu vrijednost (0, s obzirom na definiciju čistog slučajnog procesa) i konačno kao treći argument, varijancu procesa (1 s obzirom na uvjet zadatka). Grafički prikaz niza „epsilon“ predočen je na slici 1.49. S obzirom na definiciju stacionarnih procesa, uočava se da je očekivanje analiziranog niza konstantno (jednako 0), konstantna je varijanca (jer su odstupanja od prosječne vrijednosti podjednake u cijelom uzorku), dok se svojstvo (1.26) kovarijance razmatra u nastavku preko korelograma i testova.

```
epsilon<-rnorm(100,0,1)
epsilon<-ts(epsilon,start=c(2011,9),frequency=12)
plot(epsilon)
```

Slika 1.48. Generiranje i grafičko predočavanje Gaussova čistog slučajnog procesa

⁵ S obzirom na narav naredbe generiranja niza slučajne varijable, vrijednosti se u svakom ponavljanju generiranja niza slučajne varijable mijenjaju, no zaključci koji su prikazani u ovome primjeru vrijede bez obzira na te promjene.



Slika 1.49. Grafički prikaz „epsilon“

Nadalje, pomoću naredbi navedenih na slici 1.50. zajedno su predočene autokorelacijska i parcijalna autokorelacijska funkcija na korelogramima na slici 1.51. Uočava se da se svi koeficijenti na oba korelograma nalaze unutar granica neznačajnosti $[-1,96 / \sqrt{T}, 1,96 / \sqrt{T}]$. S obzirom na samu definiciju čistog slučajnog procesa, vrijednosti analiziranih koeficijenata su veoma blizu vrijednosti 0. To se dodatno potvrđuje na slici 1.52., gdje su provedeni Box-Pierceov i Ljung-Boxov test.

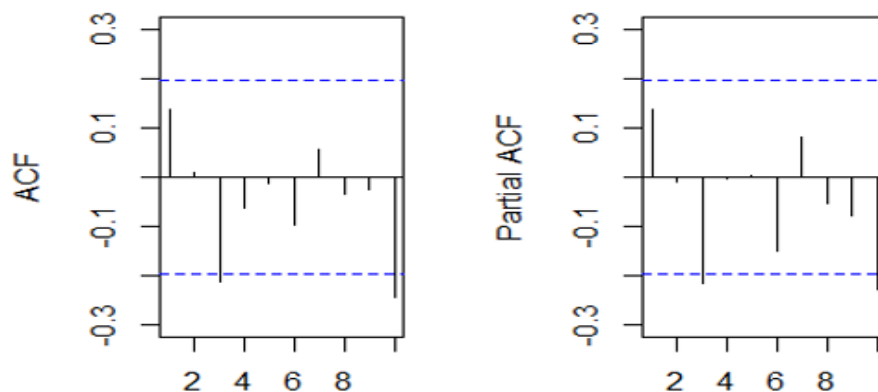
$$H_0 : \rho_1 = \rho_2 = \dots = \rho_{10} = 0$$

$$H_0 : \exists \rho_j \neq 0, j \in \{1, 2, \dots, 10\}, Q(10) = 14,31, p\text{-v} = 0,96 > \alpha \text{ i } Q^*(10) = 15,57, p\text{-v} = 0,11 > \alpha.$$

Stoga uz razinu značajnosti od 5%, ne odbacujemo nultu hipotezu da ne postoji autokorelacija među članovima procesa „epsilon“ do zaključno 10. reda.

```
library(forecast)
par(mfrow = c(1, 2))
Acf(epsilon, lag.max=10)
Pacf(epsilon, lag.max=10)
par(mfrow = c(1, 1))
```

Slika 1.50. Naredbe za prikaz korelograma



Slika 1.51. Korelogrami autokorelacijske i parcijalne autokorelacijske funkcije

```

Box.test(epsilon,lag = 10,type="Box-Pierce")

##
## Box-Pierce test
##
## data:  epsilon
## X-squared = 14.312, df = 10, p-value = 0.1592

Box.test(epsilon,lag = 10,type="Ljung-Box")

##
## Box-Ljung test
##
## data:  epsilon
## X-squared = 15.565, df = 10, p-value = 0.1128
    
```

Slika 1.52. Rezultati Box-Pierceova i Ljung-Boxova testa

1.9.3. Autoregresijski proces, AR

1.9.3.1. Definiranje autoregresijskog procesa

Autoregresijski proces (engl. *autoregressive process*, AR) je stohastički proces y_t koji se definira temeljem prethodnih p vrijednosti tog procesa $y_{t-1}, y_{t-2}, \dots, y_{t-p}$:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t, \quad (1.38)$$

gdje ε_t predstavlja slučajnu varijablu, $\phi_j, j \in \{1, \dots, p\}$ autoregresijske parametre, te ϕ_0 konstantu. Općenito se autoregresijski proces označava s $AR(p)$, kako bi se temeljem pomaka p znalo koliko je prethodnih pomaka (ili lagova, engl. *lag*) uključeno u model (1.38). Često se ekonomske varijable, tj. njihov razvoj tijekom vremena može opisati temeljem njihovih prethodnih vrijednosti. Dakle, pomoću $AR(p)$ modela definiranim u (1.38), koji se još naziva i stohastička diferencijska jednadžba. Kompaktnije se model (1.38) može zapisati i na sljedeći način:

$$y_t = \phi_0 + \sum_{j=1}^p \phi_j y_{t-j} + \varepsilon_t. \quad (1.39)$$

Češće se koristi sljedeći zapis $AR(p)$ procesa gdje se koristi operator pomaka (engl. *lag operator, backward shift*) L :

$$Ly_t = y_{t-1}, \quad (1.40)$$

koji pomiče vremenski indeks t za jednu vrijednost unazad, $t-1$. Dakle, to je linearni operator koji pomiče vrijednosti niza y_t za određeni pomak. Nadalje, vrijedi:

$$L^2 y_t = y_{t-2}, L^3 y_t = y_{t-3}, \dots, L^p y_t = y_{t-p}, \quad (1.41)$$

pa je općenito $L^j y_t = y_{t-j}$. Uvrštavanjem supstitucije (1.41) u (1.38) dobivamo sljedeće:

$$y_t = \phi_0 + \phi_1 L y_t + \phi_2 L^2 y_t + \dots + \phi_p L^p y_t + \varepsilon_t, \quad (1.42)$$

$$y_t - \phi_1 L y_t - \phi_2 L^2 y_t - \dots - \phi_p L^p y_t = \phi_0 + \varepsilon_t, \quad (1.43)$$

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) y_t = \phi_0 + \varepsilon_t, \quad (1.44)$$

$$\phi(L) y_t = \phi_0 + \varepsilon_t, \quad (1.45)$$

gdje je $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$ polinom reda p . Traženje nul-točaka polinoma $\phi(L)$ na način da se operator pomaka L zamijeni nepoznicama λ : $1 - \phi_1 \lambda - \phi_2 \lambda^2 - \dots - \phi_p \lambda^p = 0$, dobiva se **inverzna karakteristična jednadžba** procesa definiranog u (1.38). Rješenja karakteristične jednadžbe nazivaju se korijeni polinoma $\phi(L)$.

Polinom reda p ima ukupno p korijena (rješenja pripadne jednadžbe), koji određuju „ponašanje“ vremenskog niza. **Ako je apsolutna vrijednost svih korijena inverzne karakteristične jednadžbe veća od jedinice**, kažemo da je $AR(p)$ **stabilan**, a time i (slabo) **stacionaran**.

1.9.3.2. Svojstva autoregresijskih procesa

Razmotrimo nekoliko osnovnih autoregresijskih procesa s njihovim svojstvima.

Autoregresijski proces prvog reda, AR(1)

Autoregresijski proces prvog reda definiran je kao:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2), \quad (1.46)$$

gdje *iid* označava niz identično i nezavisno distribuiranih varijabli (engl. *iid – independent and identically distributed*). Očekivana vrijednost $AR(1)$ procesa iznosi:

$$E(y_t) = \phi_0 + \phi_1 E(y_{t-1}), \quad (1.47)$$

a zbog uvjeta stacionarnosti je: $E(y_t) = E(y_{t-1}) \equiv \mu$ pa vrijedi:

$$\mu = \phi_0 + \phi_1 \mu, \quad (1.48)$$

$$\mu - \phi_1 \mu = \phi_0 \Rightarrow (1 - \phi_1) \mu = \phi_0 \Rightarrow \mu = \frac{\phi_0}{1 - \phi_1}. \quad (1.49)$$

Kako bi očekivana vrijednost bila konačna, vrijedi: $1 - \phi_1 \neq 0$, tj. $\phi_1 \neq 1$. Međutim, dobiveni uvjet je **dovoljan**, ali **ne i nužan** da bi proces bio stacionaran. Naime, karakteristični polinom i traženje nul-točaka tog polinoma u slučaju $AR(1)$ procesa dobili bismo tako da najprije model (1.46) zapišemo pomoću operatora pomaka:

$$y_t = \phi_0 + \phi_1 Ly_t + \varepsilon_t \Rightarrow y_t - \phi_1 Ly_t = \phi_0 + \varepsilon_t \Rightarrow (1 - \phi_1 L)y_t = \phi_0 + \varepsilon_t, \quad (1.50)$$

te potom za polinom $(1 - \phi_1 L)$ riješimo sljedeću jednadžbu:

$$1 - \phi_1 \lambda = 0 \Rightarrow \lambda = \frac{1}{\phi_1}. \quad (1.51)$$

Da bi proces bio stacionaran, mora vrijediti $\left|\frac{1}{\phi_1}\right| > 1$, odnosno $|\phi_1| < 1$. Dakle, AR(1) proces je stacionaran ako vrijedi $|\phi_1| < 1$, tj. $-1 < \phi_1 < 1$.

Varijanca AR(1) procesa iznosi:

$$E[(y_t - \mu)^2] = \phi_1^2 E[(y_{t-1} - \mu)^2] + E(\varepsilon_t^2), \quad (1.52)$$

a zbog uvjeta stacionarnosti je $E[(y_t - \mu)^2] = E[(y_{t-1} - \mu)^2]$, pa slijedi:

$$Var(y_t) = \phi_1^2 Var(y_t) + \sigma^2, \quad (1.53)$$

$$Var(y_t) - \phi_1^2 Var(y_t) = \sigma^2 \Rightarrow Var(y_t) = \frac{\sigma^2}{1 - \phi_1^2}. \quad (1.54)$$

Kako varijanca procesa mora biti pozitivna, vrijedi: $1 - \phi_1^2 > 0$, tj. $\phi_1^2 < 1$. Ovo je ujedno i nužan i dovoljan uvjet za stacionarnost procesa jer uključuje već obrađen uvjet stacionarnosti. S obzirom na ovaj uvjet, i prethodni uvjet stacionarnosti, slijedi zaključak da je AR(1) proces (slabo) stacionaran ako vrijedi $-1 < \phi_1 < 1$.

Nadalje, razmatraju li se koeficijenti autokorelacijske funkcije za AR(1) proces za pomake h , možemo razmotriti AR(1) u (1.46), ali bez konstante ϕ_0 (s obzirom da konstanta ne utječe na kovarijancu): $y_t = \phi_1 y_{t-1} + \varepsilon_t$. Pomnožimo li lijevu i desnu stranu jednakosti s y_{t-h} :

$$y_t y_{t-h} = \phi_1 y_{t-1} y_{t-h} + \varepsilon_t y_{t-h}, \quad (1.55)$$

te se promotri očekivana vrijednost od (1.55) kako bi se odredila autokovarijanca između y_t i y_{t-h} :

$$\underbrace{E(y_t y_{t-h})}_{=\gamma_h} = \underbrace{E(\phi_1 y_{t-1} y_{t-h})}_{=\phi_1 \gamma_{h-1}} + \underbrace{E(\varepsilon_t y_{t-h})}_{=0}, \quad (1.56)$$

jer su nekorelirani

$$\gamma_h = \phi_1 \gamma_{h-1}. \quad (1.57)$$

Autokorelacijska funkcija tada je jednaka:

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \phi_1 \gamma_{h-1} = \phi_1 \underbrace{\phi_1 \gamma_{h-2}}_{=\gamma_{h-1}} = \phi_1 \phi_1 \underbrace{\phi_1 \gamma_{h-3}}_{=\gamma_{h-2}} = \dots = \phi_1^h \underbrace{\gamma_0}_{=1} = \phi_1^h. \quad (1.58)$$

Konačno, možemo pisati da je autokorelacijska funkcija jednaka:

$$\rho_h = \begin{cases} \phi_1, & h = 1 \\ \phi_1^h, & h > 1 \end{cases}, \quad (1.59)$$

pa je uz uvjet stacionarnosti $-1 < \phi_1 < 1$ posljedica sljedeća: **koeficijenti autokorelacijske funkcije se eksponencijalno ili sinusoidalno smanjuju.**

Koeficijenti parcijalne autokorelacijske funkcije (PACF) mogu se dobiti procjenom sljedećih modela redom:

$$\begin{aligned} y_t &= \phi_{01} + \phi_{11}y_{t-1} + \varepsilon_{1t} \\ y_t &= \phi_{02} + \phi_{12}y_{t-1} + \phi_{22}y_{t-2} + \varepsilon_{2t} \\ y_t &= \phi_{03} + \phi_{13}y_{t-1} + \phi_{23}y_{t-2} + \phi_{33}y_{t-3} + \varepsilon_{3t} \\ &\dots \end{aligned}, \quad (1.60)$$

gdje je ϕ_{0h} konstanta, a ϕ_{hh} koeficijent parcijalne autokorelacije reda h . S obzirom na definiciju autoregresijskog procesa, procijenjene vrijednosti $\hat{\phi}_{hh}$ konvergiraju prema vrijednosti 0 za pomake $h > p$. To znači da će u slučaju AR(1) procesa, vrijednost $\hat{\phi}_{11}$ iznositi upravo koliko iznosi koeficijent autokorelacije prvog reda, tj. $\hat{\phi}_{11} = \phi_1$, a **ostali $\hat{\phi}_{hh}$ kad je $h > 1$ (zbog $p = 1$) konvergiraju brzo nuli, tj. iščezavaju.**

Sumarno, ACF i PACF koeficijenti za AR(1) proces se ponašaju na sljedeći način:

$$\rho_h = \begin{cases} \phi_1, & h = 1 \\ \phi_1^h, & h > 1 \end{cases}, \quad \phi_{hh} = \begin{cases} \phi_1, & h = 1 \\ 0, & h > 1 \end{cases}, \quad (1.61)$$

pri čemu se ACF koeficijenti eksponencijalno smanjuju, a PACF iščezavaju nakon pomaka 1.

Primjer 1.12.

Otvorimo praznu skriptu u RStudiju i generirajmo niz „y“ sa 100 vrijednosti, pomoću formule: $y_t = 1 + 0,88y_{t-1} + \varepsilon_t$, gdje je slučajna varijabla iz prethodnog primjera (1.11) niz epsilon. Definirajmo da se radi o vremenskom nizu s mjesečnim frekvencijama, s početkom rujana 2011. godine. Grafički predočimo niz „y“ i komentirajmo njegovu stacionarnost prema definiciji stacionarnog procesa. Potom predočimo korelograme autokorelacijske i parcijalne autokorelacijske funkcije do 10. reda i komentirajmo ih. Provedimo Box-Pierceov test o značajnosti koeficijenata autokorelacije do zaključno 10. reda ($m = 10$), a potom i Ljung-Boxov test za isti red. Interpretirajmo rezultate za $\alpha = 5\%$. Konačno, komentirajmo stacionarnost procesa temeljem korijena karakteristične jednadžbe.

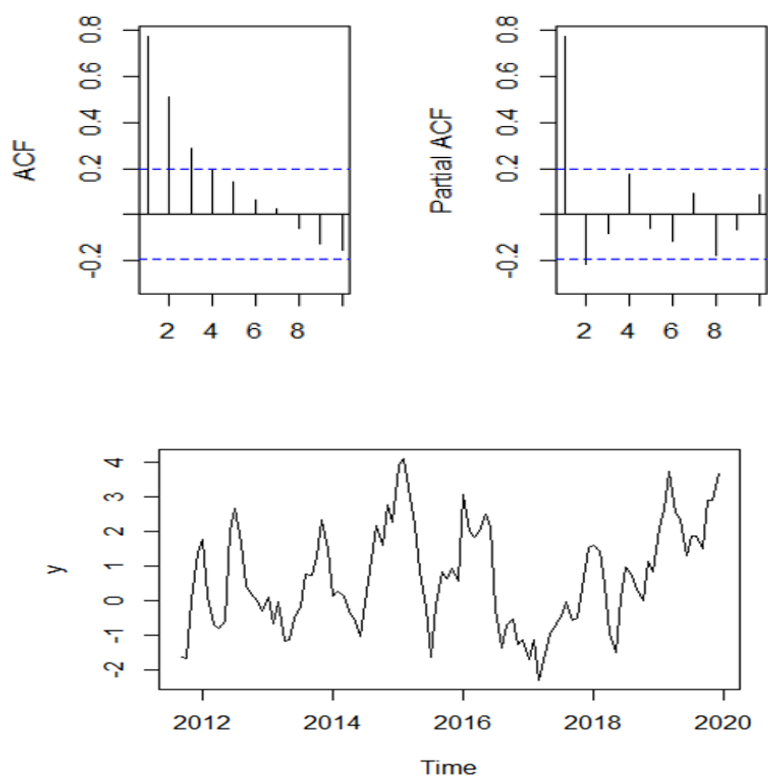
Da bismo generirali niz „y“ (vidjeti sliku 1.53.), potrebno je u okviru paketa stats koristiti naredbu kojom se simuliraju podaci, `arima.sim(...)`,

```
arima.sim(model=list(ar=0.88),n=100,innov=epsilon),
```

gdje za prvi argument pišemo tip procesa koji se želi simulirati (`model=list(...)`), gdje se navodi da je koeficijent uz pomak varijable y jednak 0,88 prema uvjetu zadatka, uz broj članova procesa jednak 100 ($n=100$) i da se za slučajnu varijablu koriste vrijednosti varijable „epsilon“ iz prethodnog primjera. Kako naredba `arima.sim(...)` generira niz čije je očekivanje jednako 0, u dodatnoj naredbi definiramo „y“ na način da prethodno definiranom nizu dodamo vrijednost 1 (naredba `y<-1+y`). Nakon što smo definirali da se radi o vremenskom nizu pomoću naredbe `ts(...)`, grafički je taj niz predočen na slici 1.54. Može se reći da zadovoljava svojstvo slabe stacionarnosti konstantnog očekivanja.

```
library(stats)
y<-arima.sim(model=list(ar = 0.88), n=100, innov = epsilon)
y<-1+y
y<-ts(y,start=c(2011,9),frequency=12)
plot(y)
```

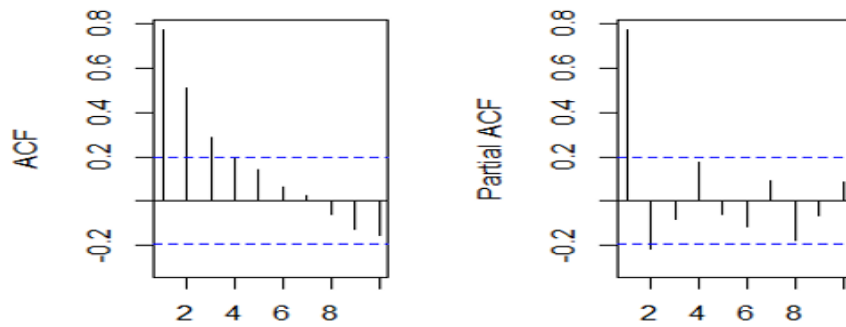
Slika 1.53. Naredbe za definiranje i grafičko predočavanje AR(1) procesa



Slika 1.54. AR(1) proces, $y_t = 1 + 0,88y_{t-1} + \varepsilon_t$

Na slici 1.55. predočeni su korelogrami autokorelacijske (ACF) i parcijalne autokorelacijske (PACF) funkcije. Uočava se da se vrijednosti ACF funkcije smanjuju veoma brzo (eksponencijalno), dok koeficijenti PACF funkcije iščezavaju nakon prvog pomaka.

Nadalje, i Box-Pierceov i Ljung-Boxov test za razinu značajnosti ukazuju na ne odbacivanje hipoteze o ne postojanju autokorelacije među članovima procesa „y“ do zaključno 10. reda (vidjeti sliku 1.56.).

Slika 1.55. ACF i PACF korelogrami za AR(1), $y_t = 1 + 0,88y_{t-1} + \varepsilon_t$

Kako bi se procijenio odgovarajući AR(1) model, koristi se naredba `auto.arima(...)`, vidjeti sliku 1.57. U okviru spomenute naredbe, navodi se naziv niza za koji se provodi Box-Jenkinsov postupak (vidjeti odjeljak 5.11.), koji odabire najbolji model koji će opisivati podatke temeljem određenih kriterija, pri čemu se navodi da se niz ne diferencira ($d=0$, vidjeti odjeljak 1.10.2). Dobiven je model: $y_t = 1,86 + 0,82y_{t-1} + \varepsilon_t$. S naredbom `autoplot(...)` možemo grafički predložiti inverze korijena karakterističnog polinoma (slika 1.58), ili pak ispisati same vrijednosti korijena karakterističnog polinoma.

```
Box.test(y,lag = 10,type="Box-Pierce")

##
## Box-Pierce test
##
## data: y
## X-squared = 291.76, df = 10, p-value < 2.2e-16

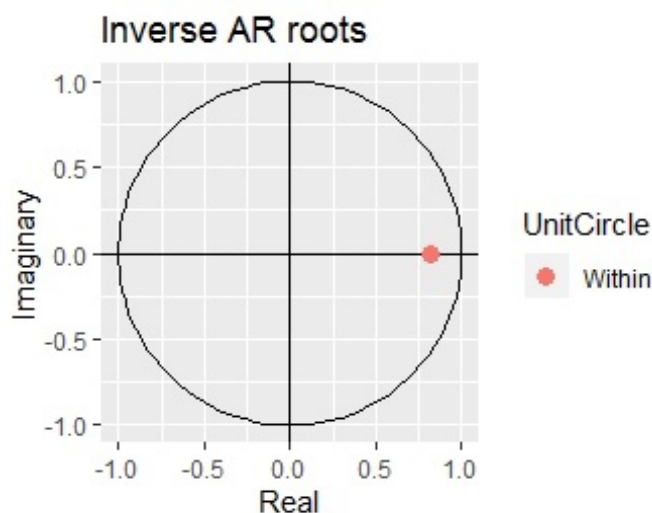
Box.test(y,lag = 10,type="Ljung-Box")

##
## Box-Ljung test
##
## data: y
## X-squared = 308.04, df = 10, p-value < 2.2e-16
```

Slika 1.56. Testiranje autokorelacije među članovima procesa „y“ do zaključno 10. reda

```
ar_1<-auto.arima(y,d=0)
summary(ar_1)
## Series: y
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##      ar1      mean
##    0.8907  2.0976
## s.e. 0.0435  0.7598
##
## sigma^2 = 0.811: log likelihood = -131.2
## AIC=268.4  AICc=268.65  BIC=276.22
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
## Training set 0.0320541 0.8915311 0.7027303 31.00786 99.40404 0.3196439 0.01296325
korijeni<-autoplot(ar_1)
plot(korijeni)
korijeni$data$roots
## [1] 1.122691+0i
```

Slika 1.57. Procjena AR(1) modela i izračun korijena inverzne karakteristične jednadžbe



Slika 1.58. Grafički prikaz inverza korijena karakteristične jednadžbe

Napomena: Kako se najprije računa korijen inverznog karakterističnog polinoma prema formuli (1.51), naredba `autoplot()` crta recipročnu (inverznu) vrijednost od tog korijena λ . Zato je na prikazu apsolutna vrijednost manja od 1. S druge strane, ako ispišemo vrijednost korijena λ , ona iznosi 1,123. Dakle, odabir naredbe grafičkog predočavanja će prikazati vrijednost ϕ_1 , dok sam ispis vrijednosti korijena daje ispis od $\lambda = 1 / \phi_1$. Kako korijeni karakterističnog polinoma mogu biti realni ili kompleksni brojevi, RStudio daje ispis istovremeno i realnog i imaginarnog dijela pojedinog broja. Tako na slici 1.57. vrijednost „1.122691+0i“ znači da se radi o realnom broju jer je imaginarni dio jednak 0·i.

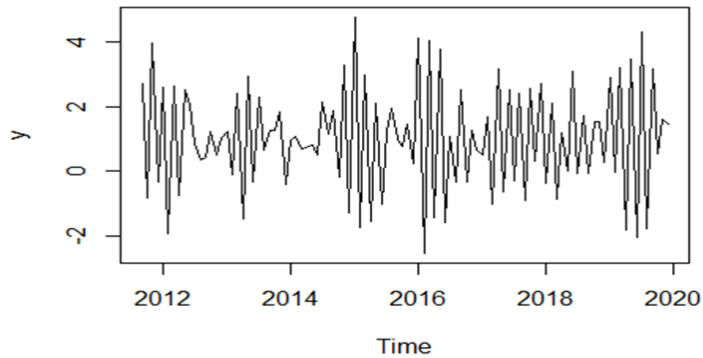
Napomena: Naredba `auto.arima(...)` prema određenim kriterijima (vidjeti odjeljak 2.6.2.) odabire najbolji model, stoga se kod generiranja nekih procesa može dogoditi da ova naredba procijeni drugačiji oblik modela u odnosu na onaj model kojim je proces generiran. Stoga se može koristiti naredba `arima(...)` u okviru koje korisnik sam definira pomake procesa koji se razmatraju.

Kako je korijen inverznog karakterističnog polinoma po apsolutnoj vrijednosti veći od 1 (tj. $1,123 > 1$) ili njegov inverz zadovoljava uvjet stacionarnosti: $\phi_1 = 1/\lambda = 1/1,123 = 0,89$, $-1 < \phi_1 < 1$, zaključuje se da je AR(1) niz „y“ stacionaran.

Primjer 1.13.

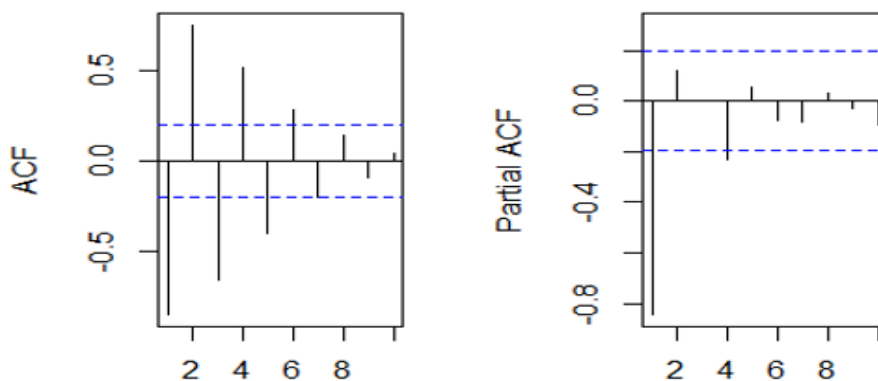
Otvorimo praznu skriptu u RStudio i generirajmo niz „y“ sa 100 vrijednosti, formulom: $y_t = 1 - 0,88y_{t-1} + \varepsilon_t$, gdje je slučajna varijabla iz primjera 1.11 niz epsilon. Definirajmo da se radi o vremenskom nizu s mjesečnim frekvencijama, s početkom rujana 2011. godine. Grafički predočimo niz „y“ i komentirajmo njegovu stacionarnost prema definiciji stacionarnog procesa. Potom predočimo korelograme autokorelacijske i parcijalne autokorelacijske funkcije do 10. reda i komentirajmo ih. Provedimo Box-Pierceov test o značajnosti koeficijenata autokorelacije do zaključno 10. reda ($m = 10$), a potom i Ljung-Boxov test za isti red. Interpretirajmo rezultate za $\alpha = 5\%$. Konačno, komentirajmo stacionarnost procesa temeljem korijena karakteristične jednadžbe.

Sada su naredbe iz prethodnog primjera identične naredbama potrebnim za ovaj primjer, uz izuzetak stavljanja predznaka minus kod definiranja niza „y“. Grafički prikaz samog niza u ovome slučaju dan je na slici 1.59., gdje se uočava drugačije ponašanje u odnosu na prethodni primjer. Naime, s obzirom na predznak minus, slijedi da će vrijednost niza u trenutku t smanjiti (povećati) ako se vrijednost niza u prethodnom trenutku povećao (smanjio).



Slika 1.59. AR(1) proces, $y_t = 1 - 0,88y_{t-1} + \varepsilon_t$

Sada se koeficijenti ACF funkcije smanjuju sinusoidalno (slika 1.60.), dok je i dalje koeficijent PACF funkcije pri prvom pomaku jedini značajan, dok ostali iščezavaju nakon tog prvog pomaka.



Slika 1.60. ACF i PACF korelogrami za AR(1), $y_t = 1 - 0,88y_{t-1} + \varepsilon_t$

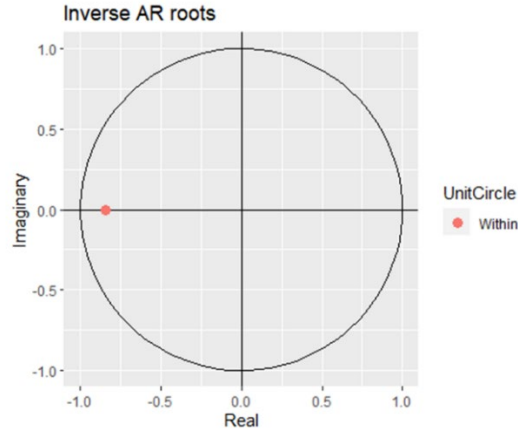
Ishodi Box-Pierceova i Ljung-Boxova testa su jednaki onima u prethodnom primjeru (slika 1.61.).

Konačno, kako je korijen inverznog karakterističnog polinoma po apsolutnoj vrijednosti veći od 1 ili njegov inverz zadovoljava uvjet stacionarnosti (slika 1.62.), zaključuje se da je AR(1) niz „y“ stacionaran

```
Box.test(y,lag = 10,type="Box-Pierce")
## Box-Pierce test
##
## data: y
## X-squared = 276.49, df = 10, p-value < 2.2e-16
Box.test(y,lag = 10,type="Ljung-Box")
```

```
## Box-Ljung test
##
## data: y
## X-squared = 293.24, df = 10, p-value < 2.2e-16
```

Slika 1.61. Testiranje autokorelacije među članovima procesa „y“ do zaključno 10. reda



Slika 1.62. Grafički prikaz inverza korijena karakteristične jednadžbe

Autoregresijski proces drugog reda, AR(2)

Autoregresijski proces drugog reda definiran je kao:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t, \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2). \quad (1.62)$$

Očekivana vrijednost AR(2) procesa iznosi:

$$E(y_t) = \phi_0 + \phi_1 E(y_{t-1}) + \phi_2 E(y_{t-2}), \quad (1.63)$$

a zbog uvjeta stacionarnosti je: $E(y_t) = E(y_{t-1}) = E(y_{t-2}) \equiv \mu$ pa vrijedi:

$$\mu = \phi_0 + \phi_1 \mu + \phi_2 \mu, \quad (1.64)$$

$$\mu - \phi_1 \mu - \phi_2 \mu = \phi_0 \Rightarrow (1 - \phi_1 - \phi_2) \mu = \phi_0 \Rightarrow \mu = \frac{\phi_0}{1 - \phi_1 - \phi_2}. \quad (1.65)$$

Stoga mora vrijediti $1 - \phi_1 - \phi_2 \neq 0$.

Nadalje, karakteristični polinom i traženje nul-točaka tog polinoma u slučaju AR(2) procesa je sljedeći:

$$y_t = \phi_0 + \phi_1 L y_t + \phi_2 L^2 y_t + \varepsilon_t \Rightarrow (1 - \phi_1 L - \phi_2 L^2) y_t = \phi_0 + \varepsilon_t, \quad (1.66)$$

pa je

$$1 - \phi_1 \lambda - \phi_2 \lambda^2 = 0 \Rightarrow \lambda_{1,2} = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}. \quad (1.67)$$

Da bi proces bio stacionaran, mora vrijediti $\left| \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2} \right| > 1$ i $\left| \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2} \right| > 1$.

Rješavanjem ovih nejednadžbi dobiva se rezultat (vidjeti izvod u Wei, 2005):

$$\begin{aligned} \phi_1 + \phi_2 &< 1 \\ \phi_2 - \phi_1 &< 1, \\ |\phi_2| &< 1 \end{aligned} \quad (1.68)$$

što je dovoljan uvjet za stacionarnost AR(2) modela.

Varijanca AR(2) modela bez konstante $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$ iznosi:

$$E[(y_t - \mu)^2] = \phi_1^2 E[(y_{t-1} - \mu)^2] + \phi_2^2 E[(y_{t-2} - \mu)^2] + E(\varepsilon_t^2), \quad (1.69)$$

a zbog uvjeta stacionarnosti je $E[(y_t - \mu)^2] = E[(y_{t-1} - \mu)^2] = E[(y_{t-2} - \mu)^2]$, pa slijedi:

$$Var(y_t) = \phi_1^2 Var(y_t) + \phi_2^2 Var(y_t) + \sigma^2, \quad (1.70)$$

$$Var(y_t) = \frac{\sigma^2}{1 - \phi_1^2 - \phi_2^2}. \quad (1.71)$$

Kako varijanca procesa mora biti pozitivna, vrijedi: $1 - \phi_1^2 - \phi_2^2 > 0$.

Nadalje, razmatraju li se koeficijenti autokorelacijske funkcije za AR(2) proces za pomake h , možemo razmotriti $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$. Pomnožimo li lijevu i desnu stranu jednakosti s y_{t-h} :

$$y_t y_{t-h} = \phi_1 y_{t-1} y_{t-h} + \phi_2 y_{t-2} y_{t-h} + \varepsilon_t y_{t-h}, \quad (1.72)$$

te se promotri očekivana vrijednost od (1.72) kako bi se odredila autokovarijanca između y_t i y_{t-h} , dobivamo kako slijedi:

$$\underbrace{E(y_t y_{t-h})}_{=\gamma_h} = \underbrace{E(\phi_1 y_{t-1} y_{t-h})}_{=\phi_1 \gamma_{h-1}} + \underbrace{E(\phi_2 y_{t-2} y_{t-h})}_{=\phi_2 \gamma_{h-2}} + \underbrace{E(\varepsilon_t y_{t-h})}_{=0 \text{ jer su nekorelirani}}, \quad (1.73)$$

$$\gamma_h = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2}. \quad (1.74)$$

Autokorelacijska funkcija tada se računa rekurzivno formulom:

$$\rho_h = \phi_1 \rho_{h-1} + \phi_2 \rho_{h-2}. \quad (1.75)$$

$$\text{Sada su } \rho_0 = 1, \rho_1 = \phi_1 \underbrace{\rho_0}_{=1} + \phi_2 \underbrace{\rho_{-1}}_{=0} \Rightarrow \rho_1 = \frac{\phi_1}{1 - \phi_2}, \rho_2 = \phi_1 \underbrace{\rho_1}_{=\frac{\phi_1}{1-\phi_2}} + \phi_2 \underbrace{\rho_0}_{=1} \Rightarrow \rho_2 = \frac{\phi_1^2 + \phi_2(1 - \phi_2)}{1 - \phi_2}.$$

Ostali ACF koeficijenti mogu se vidjeti u Wei (2005), i zbog uvjeta **stacionarnosti ACF koeficijenti za AR(2) se smanjuju eksponencijalno ili sinusoidalno prema vrijednosti 0.**

Koeficijenti parcijalne autokorelacijske funkcije (PACF) mogu se dobiti procjenom modela (1.60). Sada vrijednosti ϕ_{hh} konvergiraju prema vrijednosti 0 za pomake $h > p = 2$. To znači da će u slučaju AR(2) procesa vrijednosti ϕ_{hh} **kad je $h > 2$ (zbog $p = 2$) konvergirati brzo nuli, tj. iščezavaju.**

Sumarno, za AR(2) proces se ACF koeficijenti eksponencijalno ili sinusoidalno smanjuju, a PACF iščezavaju nakon pomaka 2.

Primjer 1.14.

Otvorimo praznu skriptu u RStudiju i generirajmo niz „y“ sa 100 vrijednosti, formulom: $y_t = 1 + 0,6y_{t-1} - 0,5y_{t-2} + \varepsilon_t$, gdje je slučajna varijabla iz primjera 1.11 niz epsilon. Definirajmo da se radi o vremenskom nizu s mjesečnim frekvencijama, s početkom rujana 2011. godine. Grafički predočimo niz „y“ i komentirajmo njegovu stacionarnost prema definiciji stacionarnog procesa. Potom predočimo korelograme autokorelacijske i parcijalne autokorelacijske funkcije do 10. reda i komentirajmo ih. Provedimo Box-Pierceov test o značajnosti koeficijenata autokorelacije do zaključno 10. reda ($m = 10$), a potom i Ljung-Boxov test za isti red. Interpretirajmo rezultate za $\alpha = 5\%$. Konačno, komentirajmo stacionarnost procesa temeljem korijena karakteristične jednadžbe.

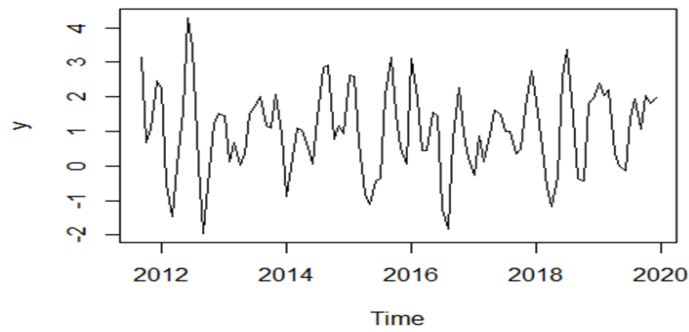
Naredbe za rješavanje ovog primjera su gotovo identične prethodnim primjerima, uz jedini izuzetak definiranja niza „y“, koji se sada definira naredbama:

```
y<-arima.sim(model=list(ar=c(0.6,-0.5)), n=100, innov = epsilon); y<-y+1,
```

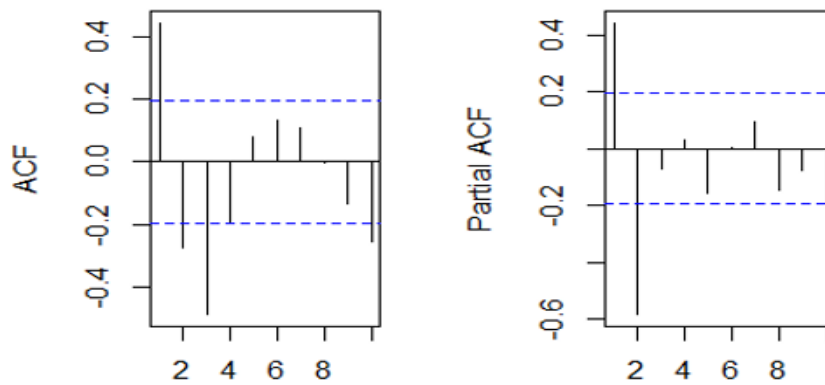
gdje se u listi koeficijenata za AR(2) model navode dva koeficijenta, a potom se definirani niz y redefinira kao on uvećan za vrijednost 1. Slika 1.63. predočava vremensku dinamiku „y“, gdje se uočava da zadovoljava svojstvo slabe stacionarnosti da je konstantno očekivanje niza.

Ono što se uočava na korelogramima za ACF i PACF funkcije jest da ACF vrijednosti se smanjuju veoma brzo, dok koeficijenti PACF funkcije iščezavaju nakon pomaka 2 (slika 1.64.).

Konačno, slike 1.65. i 1.66. predočavaju korijene inverznog karakterističnog polinoma i grafički prikaz za lakše određivanje stacionarnosti samog niza, te se zaključuje kako je niz stacionaran jer su korijeni inverznog karakterističnog polinoma po apsolutnoj vrijednosti veći od 1 (ili njihov inverz je manji od jedan po apsolutnoj vrijednosti, slika 1.66.).



Slika 1.63. Grafički prikaz niza „y“, $y_t = 1 + 0,6 y_{t-1} - 0,5 y_{t-2} + \varepsilon_t$



Slika 1.64. ACF i PACF korelogrami za AR(2), $y_t = 1 + 0,6 y_{t-1} - 0,5 y_{t-2} + \varepsilon_t$

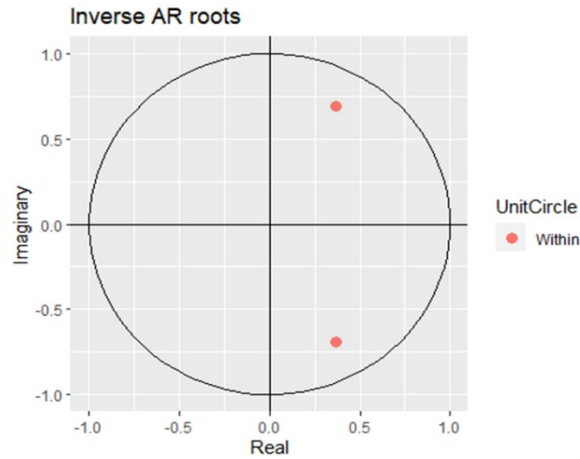
```
ar_2<-auto.arima(y,d=0)
summary(ar_2)

korijeni<-autoplot(ar_2)
plot(korijeni)

korijeni$data$roots

## [1] 0.5915354+0.987176i 0.5915354-0.987176i 0.9455458+1.646573i
## [4] 0.9455458-1.646573i
```

Slika 1.65. Korijeni inverznog karakterističnog polinoma



Slika 1.66. Grafički prikaz inverza korijena karakteristične jednadžbe

Autoregresijski proces reda p, AR(p)

Autoregresijski proces reda p definiran je kao:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t, \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2), \quad (1.76)$$

Očekivana vrijednost AR(p) procesa iznosi:

$$E(y_t) = \phi_0 + \phi_1 E(y_{t-1}) + \phi_2 E(y_{t-2}) + \dots + \phi_p E(y_{t-p}), \quad (1.77)$$

$$\mu = \phi_0 + \phi_1 \mu + \phi_2 \mu + \dots + \phi_p \mu, \quad (1.78)$$

$$\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p}. \quad (1.79)$$

Zbog uvjeta stacionarnosti mora vrijediti $1 - \phi_1 - \phi_2 - \dots - \phi_p \neq 0$. Nadalje, karakteristični polinom i traženje nul-točaka tog polinoma u slučaju AR(p) procesa je sljedeći:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) y_t = \phi_0 + \varepsilon_t, \quad (1.80)$$

gdje nul-točke polinoma $1 - \phi_1 \lambda - \phi_2 \lambda^2 - \dots - \phi_p \lambda^p = 0$ moraju po apsolutnoj vrijednosti biti veće od 1.

Varijanca AR(p) modela bez konstante $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$ iznosi:

$$Var(y_t) = \phi_1^2 Var(y_t) + \phi_2^2 Var(y_t) + \dots + \phi_p^2 Var(y_t) + \sigma^2, \quad (1.81)$$

$$\text{Var}(y_t) = \frac{\sigma^2}{1 - \phi_1^2 - \phi_2^2 - \dots - \phi_p^2}. \quad (1.82)$$

Kako varijanca procesa mora biti pozitivna, vrijedi: $1 - \phi_1^2 - \phi_2^2 - \dots - \phi_p^2 > 0$.

Autokorelacijska funkcija $\text{AR}(p)$ procesa se računa tako da se proces $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$ pomnoži s y_{t-h} , te se promatra očekivana vrijednost novodefiniranog procesa, pa su koeficijenti autokovarijance sljedeći:

$$\gamma_h = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \dots + \phi_p \gamma_{h-p}, \quad (1.83)$$

a koeficijenti autokorelacije:

$$\rho_h = \phi_1 \rho_{h-1} + \phi_2 \rho_{h-2} + \dots + \phi_p \rho_{h-p}, \quad (1.84)$$

čije se vrijednosti smanjuju eksponencijalno ili sinusoidalno.

Koeficijenti parcijalne autokorelacijske funkcije (PACF) mogu se dobiti procjenom modela (1.60). Sada vrijednosti ϕ_{hh} konvergiraju prema vrijednosti 0 za pomake $h > p$. To znači da će u slučaju $\text{AR}(p)$ procesa vrijednosti ϕ_{hh} kad je $h > p$ konvergirati brzo nuli, tj. iščezavaju.

Sumarno, za $\text{AR}(p)$ proces se ACF koeficijenti eksponencijalno ili sinusoidalno smanjuju, a PACF iščezavaju nakon pomaka p .

1.9.4. Proces pomičnih prosjeka, MA

1.9.4.1. Definiranje procesa pomičnih prosjeka

Proces pomičnih prosjeka (engl. *moving average*, MA) je proces koji je definiran temeljem slučajne varijable iz razdoblja t i prethodnih razdoblja $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$:

$$y_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \quad (1.85)$$

gdje ε_t predstavlja slučajnu varijablu, $\theta_j, j \in \{1, \dots, q\}$ parametre modela, te θ_0 konstantu. Općenito se proces pomičnih prosjeka označava s $\text{MA}(q)$, kako bi se temeljem pomaka q znalo koliko je prethodnih pomaka (ili lagova, engl. *lag*) uključeno u sam model. Kompaktnije se model (1.85) može zapisati i na sljedeći način:

$$y_t = \theta_0 + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t. \quad (1.86)$$

Uvedemo li operator pomaka L , $L^j \varepsilon_t = \varepsilon_{t-j}$, model možemo pisati i ovako:

$$y_t = \theta_0 + (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t, \quad (1.87)$$

gdje je $\theta(L) = (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)$ polinom operatora pomaka L reda q ,

$$y_t = \theta_0 + \theta(L) \varepsilon_t. \quad (1.88)$$

Proces pomičnih prosjeka je **uvijek (slabo) stacionaran** jer je linearna kombinacija stacionarne slučajne varijable i njenih prethodnih realizacija. S druge strane, da bi $MA(q)$ proces bio **invertibilan**, mora se moći prikazati kao $AR(\infty)$ proces (vidjeti 1.113). U tom slučaju vrijedi:

$$\frac{y_t}{\theta(L)} = \frac{\theta_0}{\theta(L)} + \varepsilon_t, \quad (1.89)$$

te svi korijeni polinoma $\theta(L)$ moraju biti izvan jediničnog kruga da bi $MA(q)$ proces bio invertibilan.

1.9.4.2. Svojstva procesa pomičnih prosjeka

Model pomičnih prosjeka prvog reda, MA(1)

Proces pomičnih prosjeka prvog reda, $MA(1)$, zapisuje se na način:

$$\begin{aligned} y_t &= \theta_0 + \theta_1 \varepsilon_{t-1} + \varepsilon_t = \theta_0 + \theta_1 L \varepsilon_t + \varepsilon_t \\ &= \theta_0 + (\theta_1 L + 1) \varepsilon_t, \end{aligned} \quad (1.90)$$

$$\frac{y_t}{(\theta_1 L + 1)} = \frac{\theta_0}{(\theta_1 L + 1)} + \varepsilon_t$$

čija je očekivana vrijednost: $E(y_t) = E \left[\theta_0 + \theta_1 \underbrace{\varepsilon_{t-1}}_{E(\varepsilon_{t-1})=0} + \underbrace{\varepsilon_t}_{E(\varepsilon_t)=0} \right] = \theta_0$, a varijanca je

$$Var(y_t) = Var[\theta_0 + \theta_1 \varepsilon_{t-1} + \varepsilon_t] = \theta_1^2 \sigma^2 + \sigma^2.$$

Proces je po definiciji stacionaran, a da bi bio invertibilan, nakon što se izračuna nul-točka karakteristične jednadžbe: $\theta_1 \lambda + 1 = 0 \Rightarrow \lambda = -\frac{1}{\theta_1}$, njezina apsolutna vrijednost mora biti veća

od jedinične: $\left| -\frac{1}{\theta_1} \right| > 1$, stoga slijedi da $|\theta_1| < 1$ ili $-1 < \theta_1 < 1$.

Autokorelacijska funkcija $MA(1)$ procesa dobiva se tako da se sam proces bez konstante (ponovimo, konstanta ne utječe na kovarijance) $y_t = \theta_1 \varepsilon_{t-1} + \varepsilon_t$ pomnoži s $y_{t-h} = \theta_1 \varepsilon_{t-h-1} + \varepsilon_{t-h}$ i promotri se očekivana vrijednost novodefiniranog procesa:

$$y_t y_{t-h} = \theta_1 \varepsilon_{t-1} (\theta_1 \varepsilon_{t-h-1} + \varepsilon_{t-h}) + \varepsilon_t (\theta_1 \varepsilon_{t-h-1} + \varepsilon_{t-h}), \quad (1.91)$$

$$\begin{aligned}
 E(y_t y_{t-h}) &= E(\theta_1 \varepsilon_{t-1} (\theta_1 \varepsilon_{t-h-1} + \varepsilon_{t-h})) + E(\varepsilon_t (\theta_1 \varepsilon_{t-h-1} + \varepsilon_{t-h})) \\
 &= E(\theta_1^2 \varepsilon_{t-1} \varepsilon_{t-h-1} + \theta_1 \varepsilon_{t-1} \varepsilon_{t-h}) + E(\theta_1 \varepsilon_t \varepsilon_{t-h-1} + \varepsilon_t \varepsilon_{t-h}) \quad , (1.92) \\
 &= \theta_1^2 E(\varepsilon_{t-1} \varepsilon_{t-h-1}) + \theta_1 E(\varepsilon_{t-1} \varepsilon_{t-h}) + \theta_1 E(\varepsilon_t \varepsilon_{t-h-1}) + E(\varepsilon_t \varepsilon_{t-h})
 \end{aligned}$$

i za $h = 0$ je

$$\begin{aligned}
 E(y_t y_{t-1}) &= \text{Var}(y_t) = \gamma_0 \\
 &= \theta_1^2 \underbrace{E(\varepsilon_{t-1} \varepsilon_{t-1})}_{=\sigma^2} + \theta_1 \underbrace{E(\varepsilon_{t-1} \varepsilon_t)}_{=0} + \theta_1 \underbrace{E(\varepsilon_t \varepsilon_{t-1})}_{=0} + \underbrace{E(\varepsilon_t \varepsilon_t)}_{=\sigma^2} \quad , (1.93) \\
 &= \theta_1^2 \sigma^2 + \sigma^2
 \end{aligned}$$

(varijanca koju smo već izračunali)

za $h = 1$ je

$$\begin{aligned}
 E(y_t y_{t-2}) &= \gamma_1 \\
 &= \theta_1^2 \underbrace{E(\varepsilon_{t-1} \varepsilon_{t-2})}_{=0} + \theta_1 \underbrace{E(\varepsilon_{t-1} \varepsilon_{t-1})}_{=\sigma^2} + \theta_1 \underbrace{E(\varepsilon_t \varepsilon_{t-2})}_{=0} + \underbrace{E(\varepsilon_t \varepsilon_{t-1})}_{=0} \quad , (1.94) \\
 &= \theta_1 \sigma^2
 \end{aligned}$$

a za $h > 1$ je

$$\begin{aligned}
 E(y_t y_{t-k}) &= \gamma_k \\
 &= \theta_1^2 \underbrace{E(\varepsilon_{t-1} \varepsilon_{t-k-1})}_{=0} + \theta_1 \underbrace{E(\varepsilon_{t-1} \varepsilon_{t-k})}_{=0} + \theta_1 \underbrace{E(\varepsilon_t \varepsilon_{t-k-1})}_{=0} + \underbrace{E(\varepsilon_t \varepsilon_{t-k})}_{=0} \quad . (1.95) \\
 &= 0
 \end{aligned}$$

To znači da je funkcija autokovarijance za MA(1) proces sljedeća:

$$\gamma_h = \begin{cases} \theta_1^2 \sigma^2 + \sigma^2, & h = 0 \\ \theta_1 \sigma^2 & , h = 1 \\ 0 & , h > 1 \end{cases} \quad , (1.96)$$

pa je funkcija autokorelacije ACF $\rho_h = \frac{\gamma_h}{\gamma_0}$:

$$\rho_h = \begin{cases} 1 & , h = 0 \\ \frac{\theta_1 \sigma^2}{\theta_1^2 \sigma^2 + \sigma^2} = \frac{\theta_1 \sigma^2}{(\theta_1^2 + 1) \sigma^2} = \frac{\theta_1}{\theta_1^2 + 1} & , h = 1 \\ 0 & , h > 1 \end{cases} \quad . (1.97)$$

Dakle, autokorelacijska funkcija (ACF) za proces MA(1) iščezava nakon prvog pomaka.
Prisjetimo se da je i PACF funkcija za proces AR(1) iščezavala nakon prvog pomaka.

Parcijalna autokorelacijska funkcija (PACF) definirana je na sljedeći način (izvod vidjeti u Wei, 2005): $\phi_{hh} = \frac{\theta_1^h (1 - \theta_1^2)}{1 - \theta_1^{2(h+1)}}$, $h \geq 1$, pa se njeni koeficijenti **manjuju eksponencijalno ili sinusoidalno prema vrijednosti 0**. Prisjetimo se da se i ACF funkcija za AR(1) proces smanjivala eksponencijalno ili sinusoidalno prema vrijednosti 0.

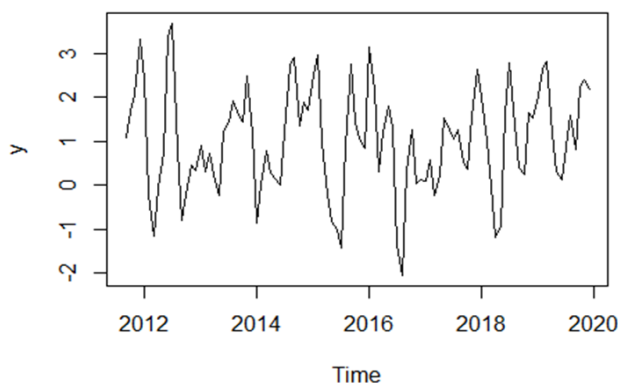
Primjer 1.15.

Otvorimo praznu skriptu u RStudiju i generirajmo niz „y“ sa 100 vrijednosti, pomoću formule: $y_t = 1 + 0,8\varepsilon_{t-1} + \varepsilon_t$, gdje je slučajna varijabla iz primjera 1.11 niz epsilon. Definirajmo da se radi o vremenskom nizu s mjesečnim frekvencijama, s početkom rujana 2011. godine. Grafički predočimo niz „y“ i komentirajmo njegovu stacionarnost prema definiciji stacionarnog procesa. Potom predočimo korelograme autokorelacijske i parcijalne autokorelacijske funkcije do 10. reda i komentirajmo ih. Konačno, komentirajmo invertibilnost procesa temeljem korijena karakteristične jednadžbe.

```
y<- arima.sim(model = list(ma =0.8), n = 100,innov=epsilon)
y<-1+y
y<-ts(y, start=c(2011,9), frequency=12)
plot(y)
```

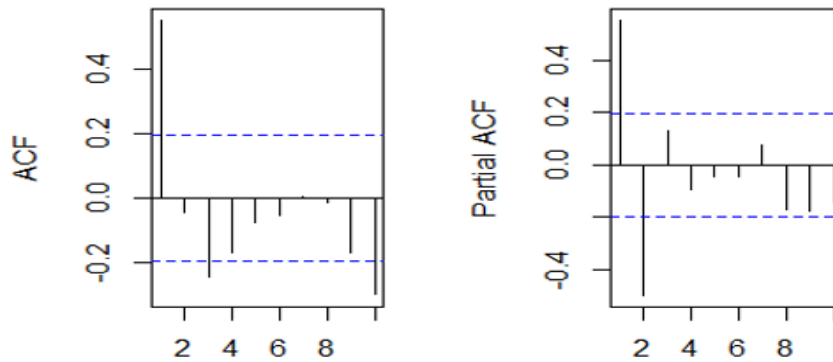
Slika 1.67. Generiranja niza „y“, $y_t = 1 + 0,8\varepsilon_{t-1} + \varepsilon_t$

Slika 1.67. predočava generiranje MA(1) procesa, gdje u listu parametara sada pišemo ma=0.8, što se odnosi na parametar uz pomak slučajne varijable, dok slika 1.68. predočava kako izgleda niz „y“, gdje se uočava da je očekivanje ovog procesa konstantno.



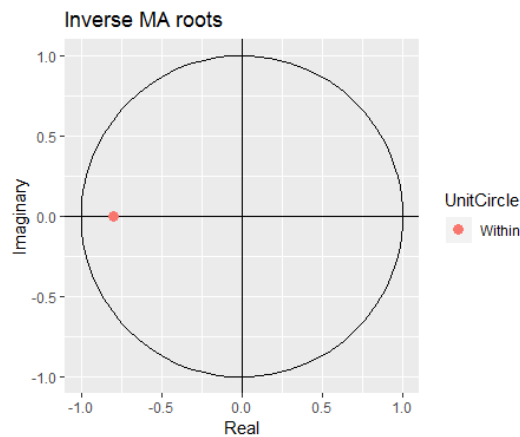
Slika 1.68. Niz „y“, $y_t = 1 + 0,8\varepsilon_{t-1} + \varepsilon_t$

Korelogrami ACF i PACF funkcije predočeni su na slici 1.69. Sada uočavamo svojstvo MA(1) procesa da ACF funkcija iščezava nakon prvog pomaka, dok PACF funkcija smanjuje svoje vrijednosti sinusoidalno prema vrijednosti 0. Slika 1.70. predočava procjenu MA(1) modela, uz ispis korijena karakteristične jednadžbe. Kako je on po apsolutnoj vrijednosti veći od 1, proces je invertibilan (ili je njegov inverz po apsolutnoj vrijednosti manji od 1, vidljivo na slici 1.71.).

Slika 1.69. Korelogrami ACF i PACF funkcija za MA(1), $y_t = 1 + 0,8\varepsilon_{t-1} + \varepsilon_t$

```
ma_1<-auto.arima(y,d=0)
summary(ma_1)
korijeni<-autoplot(ma_1)
plot(korijeni)
korijeni$data$roots
## [1] -1.334015+0i
```

Slika 1.70. Procjena i ispis korijena inverznog karakterističnog polinoma



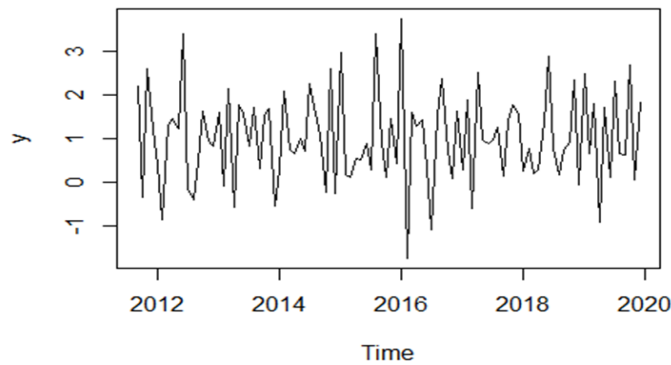
Slika 1.71. Grafički prikaz inverza korijena karakteristične jednadžbe

Primjer 1.16.

Otvorimo praznu skriptu u RStuidiu i generirajmo niz „y“ sa 100 vrijednosti, formulom: $y_t = 1 - 0,8\varepsilon_{t-1} + \varepsilon_t$, gdje je slučajna varijabla iz primjera 1.11 niz epsilon. Definirajmo da se radi o vremenskom nizu s mjesečnim frekvencijama, s početkom rujana 2011. godine. Grafički predočimo niz „y“ i komentirajmo njegovu stacionarnost prema definiciji stacionarnog procesa. Potom predočimo korelograme autokorelacijske i parcijalne autokorelacijske funkcije do 10. reda i komentirajmo ih. Konačno, komentirajmo invertibilnost procesa temeljem korijena karakteristične jednadžbe.

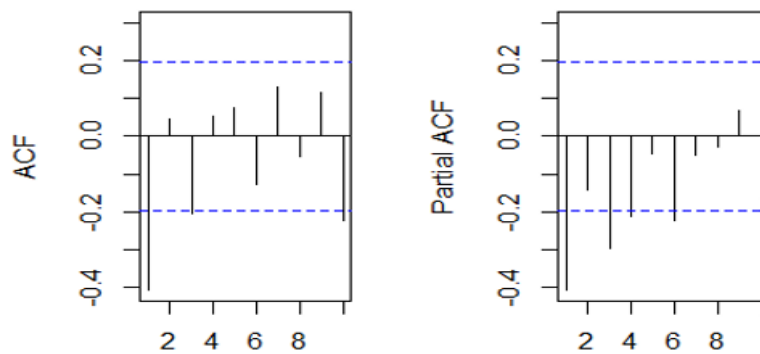
Naredbe su identične kao u prethodnom primjeru, uz napomenu predznaka minus uz 0.8. Slika 1.72. prikazuje kako izgleda MA(1) proces u ovome primjeru.

Sada se koeficijenti PACF funkcije smanjuju eksponencijalno, dok ponovno koeficijenti ACF funkcije iščezavaju nakon prvog pomaka (slika 1.73.).



Slika 1.72. Niz „y“, $y_t = 1 - 0,8\varepsilon_{t-1} + \varepsilon_t$

Konačno, korijen inverznog karakterističnog polinoma iznosi 1,42, što zadovoljava svojstvo invertibilnosti procesa (ili ekvivalentno, inverzna vrijednost od 1,2 je manje od 1).



Slika 1.73. Korelogrami ACF i PACF funkcija za MA(1), $y_t = 1 - 0,8\varepsilon_{t-1} + \varepsilon_t$

Model pomičnih prosjeka drugog reda, MA(2)

Model pomičnih prosjeka MA(2) definiran je kao:

$$\begin{aligned}
 y_t &= \theta_0 + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \varepsilon_t = \theta_0 + \theta_1L\varepsilon_t + \theta_1L^2\varepsilon_t + \varepsilon_t \\
 &= \theta_0 + (\theta_1L^2 + \theta_1L + 1)\varepsilon_t, \quad (1.98) \\
 \frac{y_t}{\theta_1L^2 + \theta_1L + 1} &= \frac{\theta_0}{\theta_1L^2 + \theta_1L + 1} + \varepsilon_t
 \end{aligned}$$

čija je očekivana vrijednost: $E(y_t) = E[\theta_0 + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \varepsilon_t] = \theta_0$, a varijanca je $Var(y_t) = Var[\theta_0 + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \varepsilon_t] = \theta_1^2\sigma^2 + \theta_2^2\sigma^2 + \sigma^2$.

Proces je po definiciji stacionaran, a da bi bio invertibilan, nakon što se izračunaju nul-točke karakteristične jednadžbe: $\theta_1\lambda^2 + \theta_1\lambda + 1 = 0$,

$$\lambda_{1,2} = \frac{\theta_1 \pm \sqrt{\theta_1^2 - 4\theta_2}}{2\theta_1}. \quad (1.99)$$

Da bi proces bio invertibilan, mora vrijediti $\frac{\theta_1 + \sqrt{\theta_1^2 - 4\theta_2}}{2\theta_1} > 1$ i $\frac{\theta_1 - \sqrt{\theta_1^2 - 4\theta_2}}{2\theta_1} > 1$.

Rješavanjem ovih nejednadžbi, dobiva se rezultat:

$$\begin{aligned} \theta_1 + \theta_2 &< 1 \\ \theta_2 - \theta_1 &< 1, \\ |\theta_2| &< 1 \end{aligned} \quad (1.100)$$

što je dovoljan uvjet za invertibilnost MA(2) modela.

Autokorelacijsku funkciju za MA(2) proces bismo slično definirali kao za slučaj MA(1), pri čemu je ona za slučaj MA(2):

$$\rho_h = \begin{cases} \frac{\theta_1(1+\theta_2)}{1+\theta_1^2+\theta_2^2} & , h=1 \\ \frac{\theta_2}{1+\theta_1^2+\theta_2^2} & , h=2 \\ 0 & , h>2 \end{cases} \quad (1.101)$$

Dakle, koeficijenti ACF funkcije za MA(2) proces iščezavaju nakon drugog pomaka. Prisjetimo se da su i koeficijenti PACF funkcije za AR(2) iščezavali nakon drugog pomaka.

Parcijalna autokorelacijska funkcija (PACF) MA(2) procesa (izvod vidjeti u Wei, 2005) sastoji se od **članova čija se vrijednost eksponencijalno ili sinusoidalno smanjuje**. Prisjetimo se da i ACF funkcija za AR(2) proces se sastoji od članova koji se eksponencijalno ili sinusoidalno smanjuju.

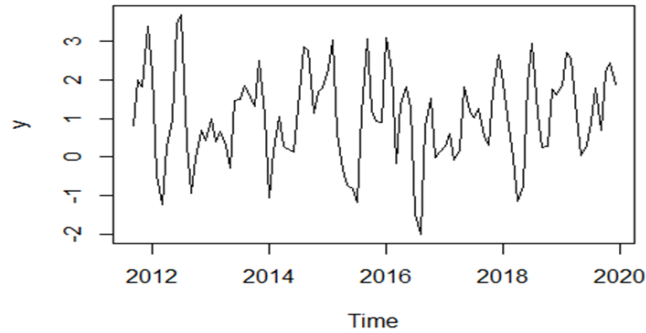
Primjer 1.17.

Otvorimo praznu skriptu u RStudiu i generirajmo niz „y“ sa 100 vrijednosti, formulom: $y_t = 1 + 0,8\varepsilon_{t-1} - 0,2\varepsilon_{t-2} + \varepsilon_t$, gdje je slučajna varijabla iz primjera 1.11 niz epsilon. Definirajmo da se radi o vremenskom nizu s mjesečnim frekvencijama, s početkom rujana 2011. godine. Grafički predočimo niz „y“ i komentirajmo njegovu stacionarnost prema definiciji stacionarnog procesa. Potom predočimo korelograme autokorelacijske i parcijalne autokorelacijske funkcije do 10. reda i komentirajmo ih. Konačno, komentirajmo invertibilnost procesa temeljem korijena karakteristične jednadžbe.

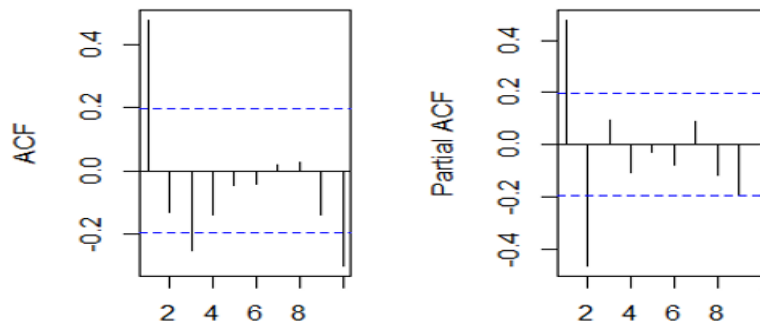
Na slici 1.74. predočeno je definiranje MA(2) procesa, dok grafički prikaz 1.75. upućuje da se radi o stacionarnom procesu. Slika 1.76. predočava korelograme, koji ukazuju da koeficijenti ACF funkcije iščezavaju nakon drugog pomaka, dok se koeficijenti PACF funkcije sinusoidalno smanjuju prema vrijednosti 0. Korijeni inverzne karakteristične jednadžbe iznose $-1,04$ i $-15,33$, te kako su po apsolutnoj vrijednosti veći od 1, radi se o invertibilnom procesu (slika 1.77., ili se razmatraju njihove inverzne vrijednosti na slici 1.78.).

```
y<- arima.sim(model = list(ma =c( 0.8,-0.2)), n = 100,innov=epsilon)
y<-1+y
y<-ts(y,start=c(2011,9),frequency=12)
plot(y)
```

Slika 1.74. Definiranje MA(2) procesa, $y_t = 1 + 0,8\varepsilon_{t-1} - 0,2\varepsilon_{t-2} + \varepsilon_t$



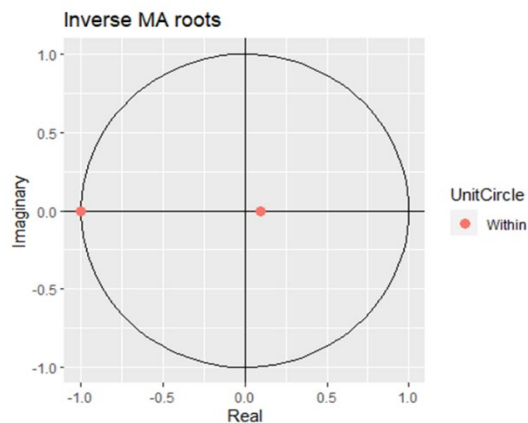
Slika 1.75. Niz „y“, $y_t = 1 + 0,8\varepsilon_{t-1} - 0,2\varepsilon_{t-2} + \varepsilon_t$



Slika 1.76. Korelogrami ACF i PACF funkcija za MA(2), $y_t = 1 + 0,8\varepsilon_{t-1} - 0,2\varepsilon_{t-2} + \varepsilon_t$

```
korijeni$data$roots
## [1] -1.03729-0i -15.33440+0i
```

Slika 1.77. Procjena i ispis korijena karakterističnog polinoma



Slika 1.78. Grafički prikaz inverza korijena karakteristične jednadžbe

Model pomičnih prosjeka reda q , MA(q)

Model pomičnih prosjeka q definiran je kao:

$$y_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2). \quad (1.102)$$

Očekivana vrijednost MA(q) procesa iznosi:

$$E(y_t) = \theta_0. \quad (1.103)$$

Nadalje, karakteristični polinom i traženje nul-točaka tog polinoma u slučaju AR(p) procesa je sljedeći:

$$\frac{y_t}{(1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)} = \frac{\theta_0}{(1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)} + \varepsilon_t, \quad (1.104)$$

gdje nul-točke polinoma $1 - \theta_1 \lambda - \theta_2 \lambda^2 - \dots - \theta_q \lambda^q = 0$ moraju po apsolutnoj vrijednosti biti veće od 1.

Varijanca MA(q) modela bez konstante $y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$ iznosi:

$$Var(y_t) = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma^2. \quad (1.105)$$

Koeficijenti autokorelacijske funkcije (ACF) (vidjeti Wei, 2005) iščezavaju nakon pomaka q , dok se koeficijenti parcijalne autokorelacijske funkcije (PACF) smanjuju eksponencijalno ili sinusoidalno prema vrijednosti 0. Prisjetimo se da za AR(p) proces se i ACF koeficijenti eksponencijalno ili sinusoidalno smanjuju, a PACF iščezavaju nakon pomaka p .

1.9.5. Woldov teorem dekompozicije

Jedan od važnijih nalaza u analizi vremenskih nizova je Woldov (1960) teorem dekompozicije (engl. *Wold's decomposition theorem*) koji glasi: svaki slabo stacionarni (kovarijančno stacionarni) stohastički proces $(y_t - \mu)$ se može prikazati kao beskonačna linearna kombinacija niza čistih slučajnih procesa:

$$y_t - \mu = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}. \quad (1.106)$$

Teorem se može iskazati i na način da se svaki slabo stacionarni stohastički proces y_t može prikazati kao zbroj dva neovisna procesa, od kojih je prvi deterministički (μ), a drugi

stohastički, $\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$.

Drugačiji način iskazivanja ovog teorema je sljedeći: Pretpostavimo da je y_t kovarijančno stacionarni proces, za koji vrijedi $E(y_t) = 0$ i $E(y_t, y_{t-h}) = \gamma_h, \forall h$. Tada je $y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} + \eta_t$,

gdje je $\psi_0 = 1, \sum_{i=0}^{\infty} \psi_i^2 < \infty, E(\varepsilon_t^2) = \sigma^2, E(\varepsilon_t, \varepsilon_s) = 0, t \neq s, E(\varepsilon_t) = 0, E(\varepsilon_t, \eta_t) = 0, \forall t, s$ i

$P(\eta_{t+s} | y_{t-1}, y_{t-2}, \dots) = \eta_{t+s}, s \geq 0$. η_t je deterministička komponenta.

Naime, ako bi se AR(1) model raspisao na sljedeći način:

$$y_t = \underbrace{\phi_0}_{=(1-\phi_1)\mu} + \phi_1 y_{t-1} + \varepsilon_t = (1-\phi_1)\mu + \phi_1 y_{t-1} + \varepsilon_t, \quad (1.107)$$

$$y_t - \mu = \phi_1 (y_{t-1} - \mu) + \varepsilon_t, \quad (1.108)$$

$$y_t - \mu = \phi_1 (\underbrace{y_{t-1} - \mu}_{=\phi_1(y_{t-2} - \mu) + \varepsilon_{t-1}}) + \varepsilon_t = \phi_1 (\phi_1 (y_{t-2} - \mu) + \varepsilon_{t-1}) + \varepsilon_t, \quad (1.109)$$

$$y_t - \mu = \phi_1 (\phi_1 (\underbrace{y_{t-2} - \mu}_{=\phi_1(y_{t-3} - \mu) + \varepsilon_{t-2}}) + \varepsilon_{t-1}) + \varepsilon_t = \phi_1 (\phi_1 (\phi_1 (y_{t-3} - \mu) + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t, \quad (1.110)$$

$$\begin{aligned} y_t - \mu &= \phi_1 (\phi_1 (\phi_1 (y_{t-3} - \mu) + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t \\ &= \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \phi_1^3 \varepsilon_{t-3} + \dots, \quad (1.111) \\ &= \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i} \end{aligned}$$

uočava se da je

$$y_t - \mu = \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i} = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}. \quad (1.112)$$

Dakle, neki autoregresijski proces y_t reda p može se prikazati na način kao što je predočeno u (1.85), odnosno može se prikazati kao beskonačni proces pomičnih prosjeka, tj. MA(∞) (vidjeti odjeljak 1.9.4.). Ako je stacionaran, $\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$ mora konvergirati ($\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} < \infty$).

1.9.6. Invertibilnost stohastičkog procesa

Važno svojstvo stohastičkih procesa je i svojstvo invertibilnosti. Ako je stohastički proces invertibilan, tada se može prikazati kao linearna kombinacija svojih prethodnih vrijednosti:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \varepsilon_t, \quad (1.113)$$

$$y_t = \sum_{i=0}^{\infty} \phi_i L^i y_{t-i} + \varepsilon_t, \quad (1.114)$$

tj. radi se o $AR(\infty)$ procesu. Ovo svojstvo je važno za procese pomičnih prosjeka MA (idući odjeljak), jer ako je moguće proces pomičnih prosjeka $MA(q)$ prikazati kao $AR(\infty)$ proces, tada MA proces nazivamo invertibilnim. S druge strane, uočava se da je **po definiciji $AR(p)$ proces uvijek invertibilan.**

1.9.7. Funkcija impulsnog odziva ARMA procesa

Za bilo koji $AR(p)$, $MA(q)$ ili $ARMA(p,q)$ proces (vidjeti 1.9.6.) može se postaviti pitanje: Koji je učinak jediničnog šoka u slučajnoj varijabli na stohastički proces y_t ? Dakle, ako se stohastički proces ponaša prema nekome ARMA modelu, moguće je analizirati reakciju tog procesa u slučaju šokova koji se mogu dogoditi. Primjerice, pretpostavimo da se kamatna stopa ponaša prema nekome od ARMA procesa i u ekonomiji se dogodi određen šok na tržištu na kojem se formira kamatna stopa, pitamo se kako će tijekom vremena to utjecati na kretanje kamatne stope, hoće li učinak biti pozitivan ili negativan, koliko jako će reagirati kamatna stopa i hoće li i kada taj učinak iščeznuti. **Grafički prikaz reakcije stohastičkog procesa na šokove u slučajnoj varijabli ε_t naziva se funkcija impulsnog odziva** (engl. *impulse response function*, IRF), jer se predočavaju koeficijenti učinka šoka u slučajnoj varijabli na stohastički proces y_t kroz vrijeme. Obično se pretpostavlja da je šok jednak jednoj standardnoj devijaciji slučajne varijable.

U tu svrhu se definirao Woldov teorem dekompozicije (odjeljak 1.9.5.), jer je potrebno proces y_t prikazati u obliku MA reprezentacije kako bi se razmatrao učinak ε_t i njegovih pomaka na sam proces y_t . Pretpostavi se da se u razdoblju t dogodi jedinični šok u slučajnoj varijabli, dakle, jednokratni šok, te se grafički predoči reakcija procesa y_t na taj šok. Posebno se funkcije impulsnog odziva koriste za analizu ARDL modela (engl. *autoregressive distributed lags*), vidjeti odjeljak 1.11, te VAR modela (engl. *vector autoregression*), vidjeti poglavlje 2.

Može se razmatrati sam učinak šoka kroz vrijeme, ali i njegova akumulirana vrijednost, tj. zbroj učinaka kroz svako razdoblje $t, t+1, \dots$.

Ako se razmatra $MA(1)$ proces za t : $y_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \varepsilon_t$, jedinični šok u ε_t ima trenutni učinak 1 na y_t , jer je $\frac{\partial y_t}{\partial \varepsilon_t} = 1$, za $t+1$: $y_{t+1} = \theta_0 + \theta_1 \varepsilon_t + \varepsilon_{t+1}$, jedinični šok u ε_t ima učinak θ_1 , za $t+2$: $y_{t+2} = \theta_0 + \theta_1 \varepsilon_{t+1} + \varepsilon_{t+2}$ učinak nestaje. Stoga bi funkcija impulsnog odziva za $MA(1)$ proces imala vrijednosti 1, θ_1 i dalje nula za preostala razdoblja.

Slično se može razmatrati $MA(2)$ proces: $y_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$, za t jedinični šok u ε_t ima trenutni učinak 1, za $t+1$ ($y_{t+1} = \theta_0 + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} + \varepsilon_{t+1}$) jedinični šok u ε_t ima učinak θ_1 , za $t+2$ ($y_{t+2} = \theta_0 + \theta_1 \varepsilon_{t+1} + \theta_2 \varepsilon_t + \varepsilon_{t+2}$) jedinični šok u ε_t ima učinak θ_2 , dok za $t+3$ učinak nestaje. Stoga bi funkcija impulsnog odziva za $MA(2)$ proces imala vrijednosti 1, θ_1 , θ_2 i dalje nula za preostala razdoblja.

Stoga se može poopćiti da za slučaj $MA(q)$ procesa funkcija impulsnog odziva ima koeficijente 1, $\theta_1, \theta_2, \dots, \theta_q, 0, 0, \dots$

AR(1) proces $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$ može se prikazati kao MA proces (vidjeti (1.86) do (1.91)):

$$y_t - \mu = \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i} = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}, \quad \text{pa ga možemo raspisati na sljedeći način:}$$

$y_t = \mu + \phi_1^1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots + \varepsilon_t$. Jedinični šok za $t = 1$ u ε_t ima učinak 1, za $t+1$ je učinak ϕ_1^1 , za $t+2$ učinak iznosi ϕ_1^2 , za $t+h$ učinak iznosi ϕ_1^h pa se uočava da za sve veću vrijednost h učinak se eksponencijalno ili sinusoidalno smanjuje. Slično se može izvesti za preostale AR(p) procese.

Primjer 1.18.

Predočimo grafički funkciju impulsnog odziva za procese:

- a) $y_t = 20 + 0,75y_{t-1} + \varepsilon_t$; b) $y_t = 20 - 0,75y_{t-1} + \varepsilon_t$; c) $y_t = 20 + 0,75\varepsilon_{t-1} + \varepsilon_t$;
 d) $y_t = 1 - 0,75\varepsilon_{t-1} + \varepsilon_t$; e) $y_t = 20 + 0,75y_{t-1} - 0,2y_{t-2} + \varepsilon_t$; f) $y_t = 1 + 0,75\varepsilon_{t-1} - 0,2\varepsilon_{t-2} + \varepsilon_t$

te komentirajmo rezultat, pri čemu ćemo razmoriti idućih 10. razdoblja.

Potrebno je uključiti paket FitARMA, kako bi se naredbom ImpulseCoefficientsARMA (...):

```
ImpulseCoefficientsARMA(phi=.75, theta=0, lag.max=10)
```

definirali koeficijenti AR ili MA procesa, pri čemu se phi= arugment odnosi na koeficijente za AR dio procesa, dok theta= na MA dio procesa; dok se lag.lag=10 odnosi na pomaka treba razmatrati IRF funkciju (vidjeti sliku 1.79). U slučaju procesa u a), phi=0.75 i theta=0; te je potrebno za ostale procese mijenjati vrijednosti AR ili pak MA dijela naredbe. Slika 1.79. predočava naredbu za proces pod a), dok slika 1.80. predočava sve IRF funkcije za procese koji se razmatraju u ovome primjeru.

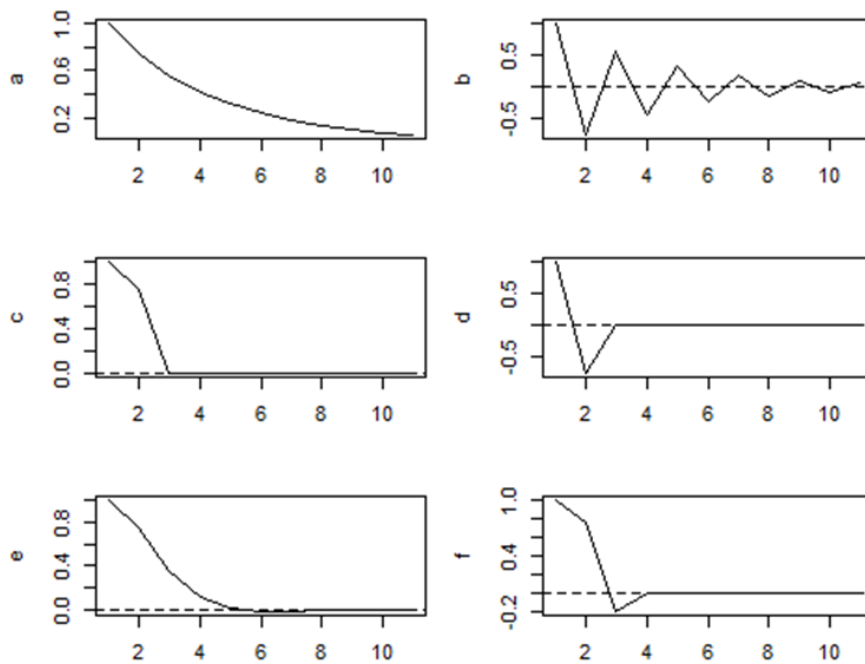
Napomena: naredba ImpulseCoefficientsARMA(...) interpretira koeficijente uz pomake slučajne varijable sa suprotnim predznakom (npr. pretpostavlja oblik MA(1) procesa:

$$y_t = \theta_0 - \theta_1 \varepsilon_{t-1} + \varepsilon_t), \text{ stoga treba pisati negativne predznake zadanih MA procesa u primjeru.}$$

Uočava se sljedeće: IRF funkcija eksponencijalno smanjuje svoju vrijednost za proces definiran u a), pri čemu šok u slučajnoj varijabli ima pozitivan učinak na y_t , koji je već oko 10. razdoblja gotovo neznačajan. Za slučaj b) je inicijalno učinak šoka u slučajnoj varijabli na y_t pozitivan, no IRF funkcija ima sinusoidalno ponašanje, što znači da se učinak kroz vrijeme mijenja; za slučaj c) je pozitivan učinak u prva dva razdoblja i nakon toga potpuno nestaje; za d); je učinak najprije pozitivan, potom negativan u idućem razdoblju i nakon toga potpuno nestaje; za e) je učinak pozitivan za prva tri razdoblja i nakon toga potpuno nestaje, dok za slučaj f) je učinak prisutan u prva tri razdoblja, IRF ima sinusoidalno ponašanje i nakon toga nestaju učinci.

```
library("FitARMA")
z<-ImpulseCoefficientsARMA(phi=.75, theta=0, lag.max=10)
plot(z,type="l",ylab="IRF",xlab="pomak")
```

Slika 1.79. Naredbe za IRF funkciju procesa $y_t = 20 + 0,75y_{t-1} + \varepsilon_t$

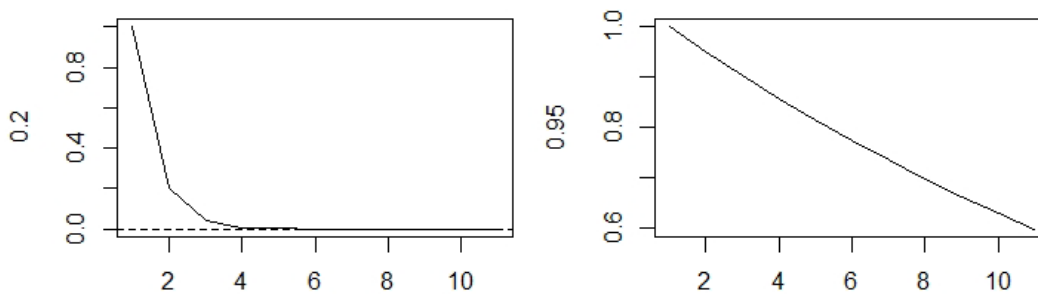


Slika 1.80. IRF funkcije za procese u primjeru

Primjer 1.19.

Usporedimo IRF funkcije za dva procesa: $y_t = 0,2y_{t-1} + \varepsilon_t$ i $y_t = 0,95y_{t-1} + \varepsilon_t$ te komentirajmo razlike za idućih 10. razdoblja.

Slika 1.81. predočava dvije IRF funkcije, na kojima se uočava da što je veći koeficijent uz pomak varijable y_t , da će šok imati veći učinak koji će dulje trajati kroz vrijeme. Stoga se može reći da AR(1) proces s većim koeficijentom ϕ ima veću „memoriju“ jer će se šokovi u slučajnoj varijabli dulje zadržati u samome procesu.



Slika 1.81. IRF funkcije za dva AR(1) procesa

1.9.8. Autoregresijski proces pomičnih prosjeka, ARMA

Kada bi se u empirijskim analizama razmatrali pojedinačno AR(p) i MA(q) modeli, da bi se neki proces dobro⁶ opisao, potrebno, bilo bi potrebno uključiti veliki pomak p ili q . Stoga se može razmatrati kombinacija i AR i MA procesa istovremeno, kako bi se smanjio broj

⁶ Dobro se odnosi na dobru dijagnostiku modela, vidjeti odjeljak 1.11.

parametara koje je potrebno procijeniti. Radi se o autoregresijskom procesu pomičnih prosjeka, engl. *autoregressive moving average*, ARMA(p, q). Upravo se u analizi financijskih varijabli često koriste modeli kojima se opisuje ARMA(p, q) proces.

ARMA(p, q) proces definiran je kao:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2), \quad (1.115)$$

$$y_t = \phi_0 + \sum_{j=1}^p \phi_j y_{t-j} + \sum_{j=1}^q \theta_j \varepsilon_{t-j}, \quad \theta_0 = 1 \quad (1.116)$$

$$\left(1 - \sum_{j=1}^p \phi_j L^j\right) y_t = \phi_0 + \left(1 + \sum_{j=1}^q \theta_j L^j\right) \varepsilon_t, \quad (1.117)$$

$$\phi(L) y_t = \phi_0 + \theta(L) \varepsilon_t. \quad (1.118)$$

ARMA(p, q) proces je invertibilan ako se može prikazati kao AR(∞) proces:

$\frac{\phi(L)}{\theta(L)} y_t = \frac{\phi_0}{\theta(L)} + \varepsilon_t$, te je i stacionaran ako se može prikazati kao MA(∞) proces:

$y_t = \frac{\phi_0}{\phi(L)} + \frac{\theta(L)}{\phi(L)} \varepsilon_t$. Da bi bio invertibilan, svi korijeni inverzne karakteristične jednadžbe

polinoma $\phi(L)$ moraju po apsolutnoj vrijednosti biti veći od 1, dok za stacionarnost svi korijeni inverzne karakteristične jednadžbe polinoma $\theta(L)$ moraju po apsolutnoj vrijednosti biti veći od 1.

Očekivanje ARMA(p, q) modela je $E(y_t) = \phi_0 + \phi_1 E(y_{t-1}) + \dots + \phi_p E(y_{t-p}) + 0$, pa je

$$\mu = \frac{\phi_0}{1 - \phi_1 - \dots - \phi_p}.$$

Autokorelacijska funkcija ARMA (p, q) procesa (vidjeti Wei, 2005) je definirana na način da se njeni koeficijenti smanjuju eksponencijalno ili sinusoidalno nakon pomaka q , dok je parcijalna autokorelacijska funkcija definirana na način da se njeni koeficijenti smanjuju eksponencijalno ili sinusoidalno nakon pomaka p .

ARMA (1,1) model je sljedeći:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t, \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2), \quad (1.119)$$

čije je očekivanje $E(y_t) = \phi_0 + \phi_1 E(y_{t-1})$, pa je $\mu = \frac{\phi_0}{1 - \phi_1}$ (kao i za slučaj AR(1) modela).

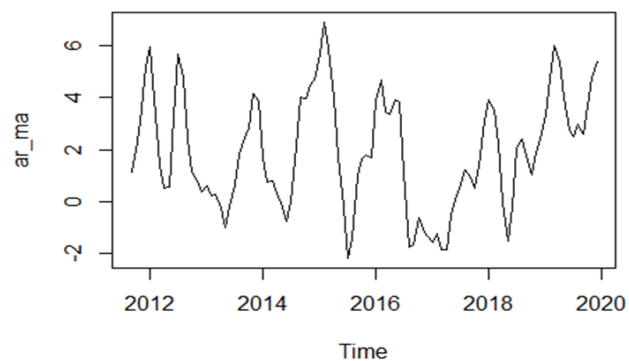
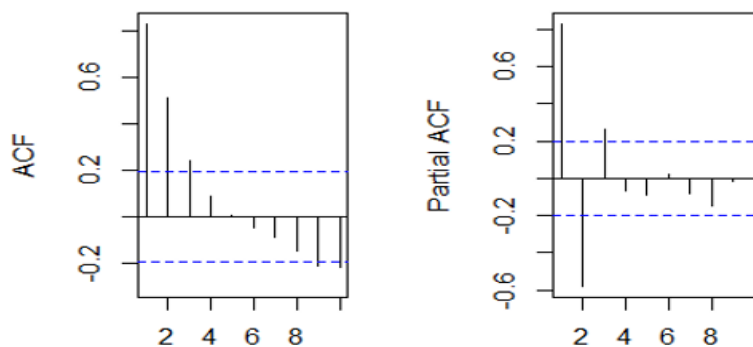
Autokorelacijska funkcija ARMA (1,1) procesa je definirana na način da se njeni koeficijenti smanjuju eksponencijalno ili sinusoidalno nakon pomaka 1, dok je parcijalna autokorelacijska funkcija definirana na način da se njeni koeficijenti smanjuju eksponencijalno ili sinusoidalno nakon pomaka 1.

Primjer 1.20.

Otvorimo praznu skriptu u RStudiju i generirajmo niz „y“ sa 100 vrijednosti, formulom: $y_t = 1 + 0,8y_{t-1} + 0,8\varepsilon_{t-1} + \varepsilon_t$, gdje je slučajna varijabla iz primjera 1.11 niz epsilon. Definirajmo da se radi o vremenskom nizu s mjesečnim frekvencijama, s početkom rujana 2011. godine. Grafički predočimo niz „y“ i komentirajmo njegovu stacionarnost prema definiciji stacionarnog procesa. Potom predočimo korelograme autokorelacijske i parcijalne autokorelacijske funkcije do 10. reda i komentirajmo ih. Komentirajmo invertibilnost procesa temeljem korijena karakteristične jednačbe. Konačno, predočimo IRF funkciju za razmatrani proces do pomaka 10.

Generiranje ARMA(1,1) procesa u ovome primjeru prikazano je na slici 1.82., s grafičkim prikazom samog procesa na slici 1.83., te se uočava na korelogramima na slici 1.84. da nakon pomaka 1 obje funkcije smanjuju svoje vrijednosti, ACF eksponencijalno, dok PACF sinusoidalno. Korijeni inverzne karakteristične jednačbe jednaki su 1,36 i -1,18 (slika 1.85). Prvi korijen odnosi se na AR dio procesa, što znači da je proces stacionaran jer je prvi korijen veći od jedinične vrijednosti, dok se drugi korijen odnosi na MA dio procesa, što znači da je proces invertibilan jer je i taj korijen veći od jedan gledajući apsolutno (ili slika 1.86. koja razmatra inverze tih korijena posebno za AR dio i posebno za MA dio).

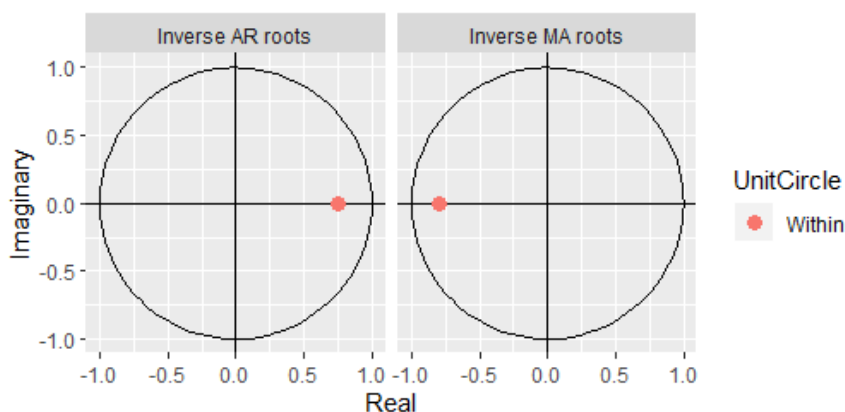
```
ar_ma<-arima.sim(model=list(ar=0.8,ma=0.8),n=100,innov=epsilon)
ar_ma<-2+ar_ma
ar_ma<-ts(ar_ma,start=c(2011,9),frequency=12)
plot(ar_ma)
```

Slika 1.82. Generiranje ARMA(1,1) procesa, $y_t = 1 + 0,8y_{t-1} + 0,8\varepsilon_{t-1} + \varepsilon_t$ Slika 1.83. ARMA(1,1) proces, $y_t = 1 + 0,8y_{t-1} + 0,8\varepsilon_{t-1} + \varepsilon_t$ Slika 1.84. Korelogrami ACF i PACF funkcija za ARMA(1,1) proces, $y_t = 1 + 0,8y_{t-1} + 0,8\varepsilon_{t-1} + \varepsilon_t$

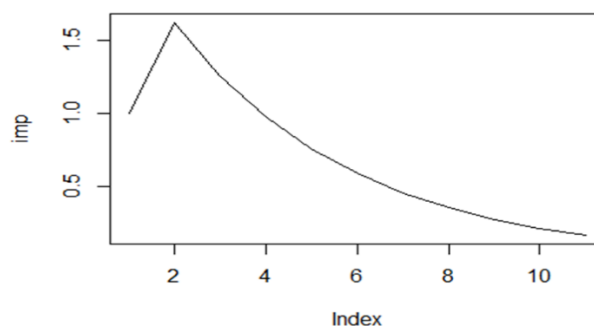
```
ar_ma_11<-arima(ar_ma,order=c(1,0,1))
korijeni<-autoplot(ar_ma_11)
plot(korijeni)
korijeni$data$roots
## [1] 1.355337+0i -1.173552+0i
```

Slika 1.85. Procjena i ispis korijena inverznog karakterističnog polinoma

IRF funkcija predočena na slici 1.87. pokazuje da je učinak šoka u slučajnoj varijabli pozitivan u slučaju razmatranog ARMA(1,1) procesa, najjači je učinak u drugom razdoblju i nakon toga se relativno brzo smanjuje.



Slika 1.86. Grafički prikaz inverza korijena karakteristične jednadžbe



Slika 1.87. IRF funkcija za ARMA(1,1) proces

Sumarno na kraju, tablica 1.1. sažima ponašanje koeficijenata ACF i PACF funkcija za odabrane procese.

Napomena: ponovimo, autoregresijski procesi su invertibilni po definiciji, dok su procesi pomičnih prosjeka po definiciji stacionarni. Za autoregresijske procese je potrebno ispitati jesu li stacionarni, dok je za procese pomičnih prosjeka potrebno ispitati jesu li invertibilni. Za mješovite ARMA procese je potrebno ispitati i invertibilnost i stacionarnost.

Tablica 1.1. Karakteristike ACF i PACF funkcija za odabrane procese

Proces	ACF koeficijenti	PACF koeficijenti
Čisti slučajni	0	0
$AR(p)$	Smanjuju se eksponencijalno ili sinusoidalno	Iščezavaju nakon pomaka p
$MA(q)$	Iščezavaju nakon pomaka q	Smanjuju se eksponencijalno ili sinusoidalno
$ARMA(p,q)$	Smanjuju se eksponencijalno ili sinusoidalno nakon pomaka q	Smanjuju se eksponencijalno ili sinusoidalno nakon pomaka p

Izvor: Guidolin i Pedio (2019)

1.9.9. Primjer

Učitajmo datoteku „**gen_nizovi.txt**“ u RStudio. Datoteka sadrži 200 opservacija za 7 različitih procesa (Y1 do Y7 nazivi). Definirajmo za svaki proces da se radi o vremenskom nizu, s početnom vrijednošću u rujnu 2003. godine, podaci su mjesečni te potom:

- Grafički predočimo svaki niz i komentirajmo njegovo ponašanje.
- Predočimo korelograme za ACF i PACF funkcije i komentirajmo o kakvom procesu se radi temeljem ponašanja koeficijenata obiju funkcija.
- Provedimo Box-Pierceov i Ljung-Boxov test autokorelacije do zaključno 10. reda, uz razinu značajnosti od 5%.
- Procijenimo odgovarajući $ARMA(p,q)$ model i komentirajmo stacionarnost i invertibilnost pojedinog procesa u ovisnosti o korijenima inverznog karakterističnog polinoma svakog procesa.
- Predočimo funkciju impulsnog odziva za svaki proces i komentirajmo što se događa s reakcijama svakog procesa na jedinični šok u slučajnoj varijabli za 10 pomaka.

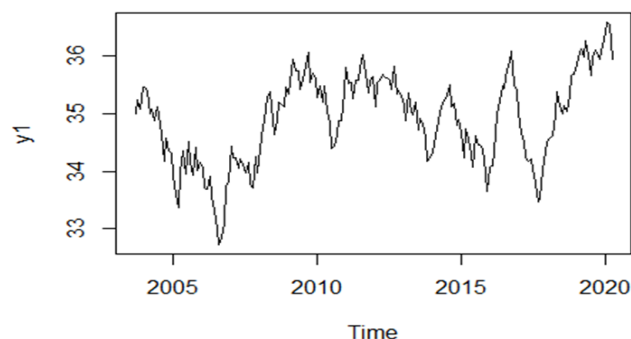
Slike 1.88. do 1.92. prikazuju rezultate postupaka a) do e) za niz „Y1“. Na slici 1.88. uočava se da određena (pozitivna) autokorelacija procesa postoji, s obzirom da povećanje vrijednosti slijede prethodna povećanja, a slično vrijedi i za smanjenja. Nadalje, slika 1.89. predočuje korelograme, na kojima se uočava da se koeficijenti ACF funkcije smanjuju eksponencijalno, a koeficijenti PACF funkcije iščezavaju nakon prvog pomaka, što bi moglo upućivati na $AR(1)$ proces. S obzirom na prvih 10 vrijednosti ACF funkcije, očekuje se koeficijent uz pomak varijable y_t da će biti relativno velik (bliži jedinici nego nuli).

Kako korelogrami upućuju na $AR(1)$ proces, očekuje se da će testiranje autokorelacije do zaključno 10. reda ukazivati na odbacivanje nulte hipoteze. Hipoteze testa su sljedeće:

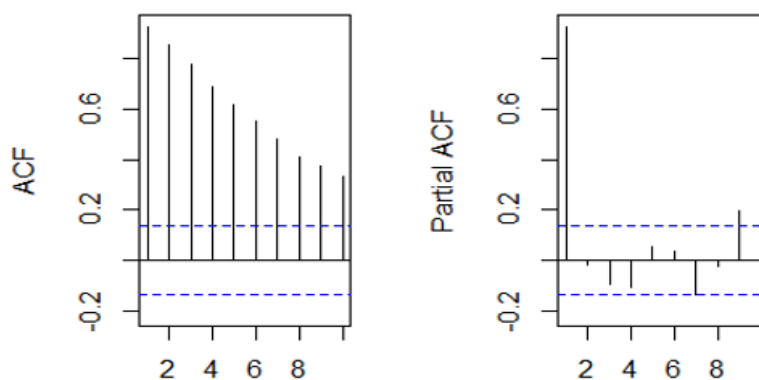
$$H_0 : \rho_1 = \dots = \rho_{10} = 0$$

$$H_1 : \exists \rho_j \neq 0, j \in \{1, \dots, 10\}$$

s test veličinama $Q(10) = 801,42$, p -vrijednost $\approx 0 < \alpha$ za Box-Pierceov test te $Q^*(10) = 825,14$, p -vrijednost $\approx 0 < \alpha$ za Ljung-Boxov test (slika 1.90.), što znači da se odbacuje hipoteza o nepostojanju autokorelacije među članovima procesa Y1 do zaključno 10. reda.



Slika 1.88. Grafički prikaz niza Y1



Slika 1.89. Korelogrami ACF i PACF funkcija za Y1

```
Box.test(y1,lag=10,type="Box-Pierce")
##
## Box-Pierce test
##
## data: y1
## X-squared = 801.42, df = 10, p-value < 2.2e-16

Box.test(y1,lag=10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: y1
## X-squared = 825.14, df = 10, p-value < 2.2e-16
```

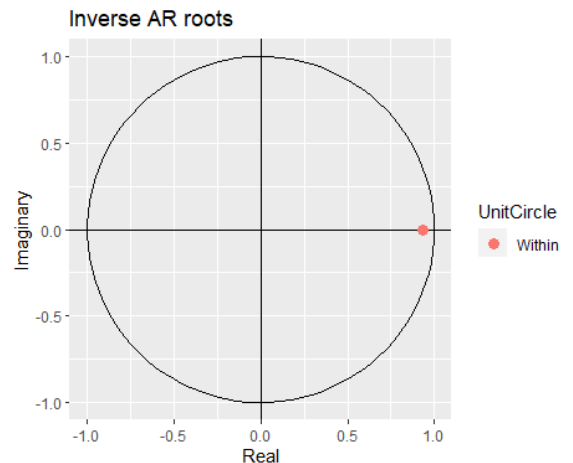
Slika 1.90. Testiranje autokorelacije procesa Y1 do zaključno 10. reda

Procijenjen je AR(1) model i izračunat je korijen inverzne karakteristične jednadžbe (slika 1.91.), gdje se uočava da je korijen $1,08 > 1$ pa je proces stacionaran, a po definiciji je proces invertibilan. Sam proces pišemo na sljedeći način: $y_t = 34,97 + 0,93y_{t-1} + \varepsilon_t$.

```

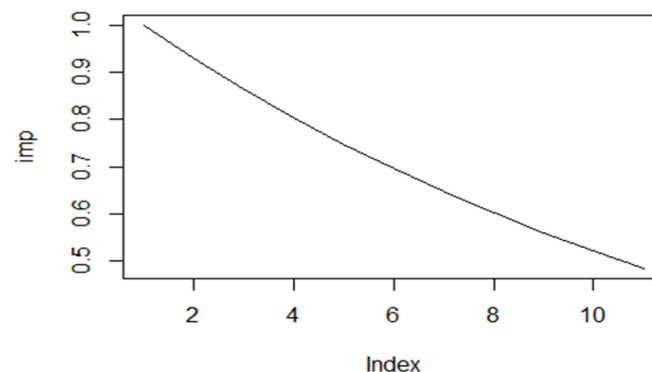
m1<-auto.arima(y1,d=0)
summary(m1)
## Series: y1
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##      ar1      mean
##  0.9302  34.9706
## s.e.  0.0251  0.2691
##
## sigma^2 = 0.0803: log likelihood = -31.59
## AIC=69.18  AICc=69.31  BIC=79.08
##
korijeni<-autoplot(m1)
plot(korijeni)
korijeni$data$roots
## [1] 1.075056+0i

```



Slika 1.91. Procjena AR(1) procesa za Y1 i izračun inverznog korijena karakteristične jednadžbe

Konačno, slika 1.92. predočava IRF funkciju za 10 razdoblja. Uočava se da šok od jedne standardne devijacije u slučajnoj varijabli pozitivno utječe na proces Y1, pri čemu je najjači učinak u prvome razdoblju i onda se on smanjuje do kraja promatranog razdoblja. Učinak je dosta jak kroz vrijeme, s obzirom da je koeficijent AR procesa 0,9302.



Slika 1.92. IRF funkcija za proces Y1

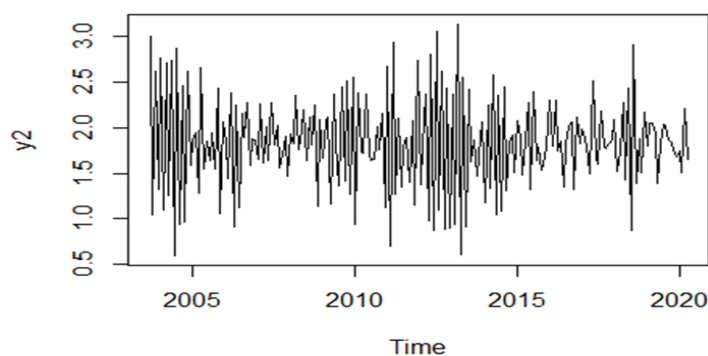
Slike 1.93. do 1.97. prikazuju rezultate postupaka a) do e) za niz „Y2“. Na slici 1.93. uočava se da određena (negativna) autokorelacija procesa postoji, s obzirom da povećanje vrijednosti slijede prethodna smanjenja, i obrnuto. Nadalje, slika 1.94. predočuje korelograme, na kojima se uočava da se koeficijenti ACF funkcije smanjuju sinusoidalno, a koeficijenti PACF funkcije iščezavaju nakon prvog pomaka, što bi moglo upućivati na AR(1) proces. S obzirom na prvih 10 vrijednosti ACF funkcije, očekuje se koeficijent uz pomak varijable y_t da će biti relativno velik (bliži -1 nego nuli).

Kako korelogrami upućuju na AR(1) proces, očekuje se da će testiranje autokorelacije do zaključno 10. reda ukazivati na odbacivanje nulte hipoteze. Hipoteze testa su sljedeće:

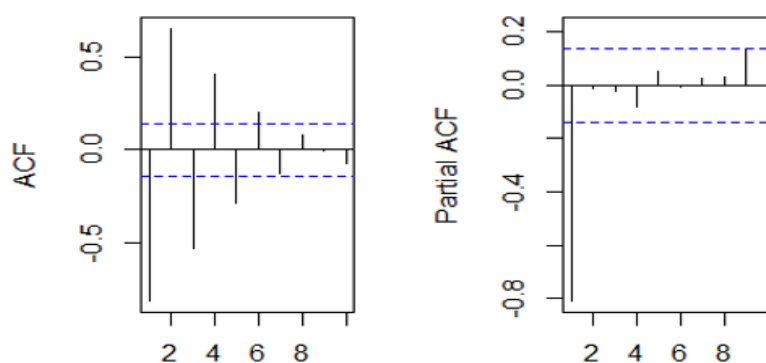
$$H_0 : \rho_1 = \dots = \rho_{10} = 0$$

$$H_1 : \exists \rho_j \neq 0, j \in \{1, \dots, 10\}$$

s test veličinama $Q(10) = 332,09$, p -vrijednost $\approx 0 < \alpha$ za Box-Pierceov test te $Q^*(10) = 339,33$, p -vrijednost $\approx 0 < \alpha$ za Ljung-Boxov test (slika 1.95.), što znači da se odbacuje hipoteza o nepostojanju autokorelacije među članovima procesa Y2 do zaključno 10. reda. Procijenjen je AR(1) model i izračunat je korijen inverzne karakteristične jednačbe (slika 1.96.), gdje se uočava da je korijen $|-1,21| > 1$ pa je proces stacionaran, a po definiciji je proces invertibilan. Sam proces pišemo na sljedeći način: $y_t = 1,83 - 0,83y_{t-1} + \varepsilon_t$. Konačno, slika 1.97. predočava IRF funkciju za 10 razdoblja. Uočava se da šok od jedne standardne devijacije u slučajnoj varijabli najprije pozitivno utječe na proces Y2, pri čemu je najjači učinak u prvome razdoblju i onda učinak ima sinusoidalno ponašanje do kraja promatranog razdoblja.



Slika 1.93. Grafički prikaz niza Y2



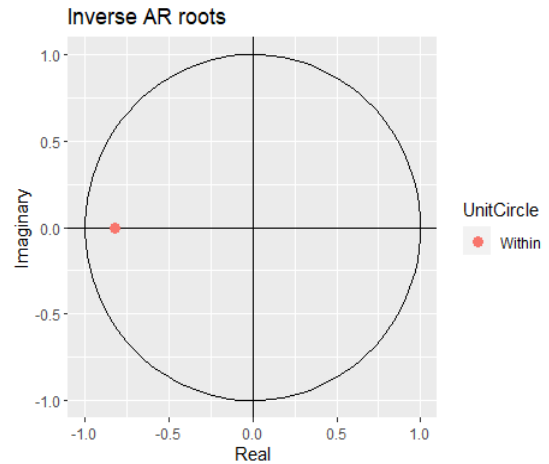
Slika 1.94. Korelogrami ACF i PACF funkcija za Y2

```
Box.test(y2,lag=10,type="Box-Pierce")
##
## Box-Pierce test
##
## data: y2
## X-squared = 332.09, df = 10, p-value < 2.2e-16

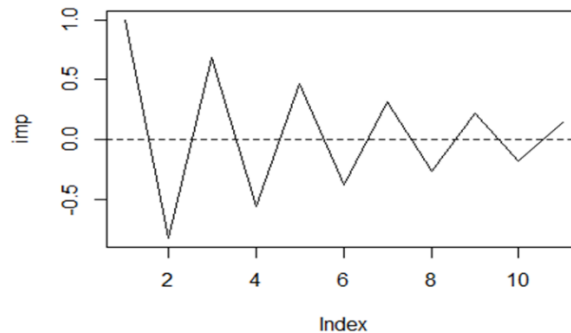
Box.test(y2,lag=10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: y2
## X-squared = 339.33, df = 10, p-value < 2.2e-16
```

Slika 1.95. Testiranje autokorelacije procesa Y2 do zaključno 10. reda

```
m1<-auto.arima(y2,d=0)
summary(m1)
## Series: y2
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##      ar1      mean
##    -0.8258  1.8328
## s.e.   0.0406  0.0115
##
## sigma^2 = 0.08876: log likelihood = -41.17
## AIC=88.35  AICC=88.47  BIC=98.24
##
korijeni<-autoplot(m1)
plot(korijeni)
korijeni$data$roots
## [1] -1.210896+0i
```



Slika 1.96. Procjena AR(1) procesa za Y2 i izračun korijena inverzne karakteristične jednačbe



Slika 1.97. IRF funkcija za proces Y2

Slike 1.98. do 1.102. prikazuju rezultate postupaka a) do e) za niz „Y3“. Na slici 1.98. uočava se da je autokorelacija procesa nešto slabija u odnosu na prethodna dva procesa, što može upućivati na MA proces. Nadalje, slika 1.99. predočuje korelograme, na kojima se uočava da se koeficijenti ACF funkcije eksponencijalno smanjuju i to veoma brzo, a koeficijenti PACF funkcije iščezavaju nakon drugog pomaka, što upućuje na AR(2) proces. Već se ovdje uočava da je putem korelograma nekad teško odrediti o kojem procesu bi se radilo.

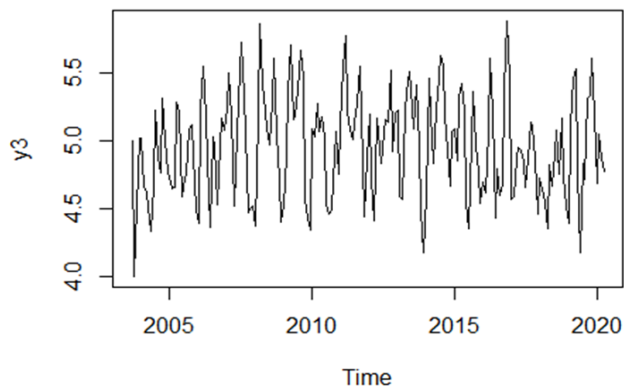
Hipoteze testova autokorelacije su sljedeće:

$$H_0 : \rho_1 = \dots = \rho_{10} = 0$$

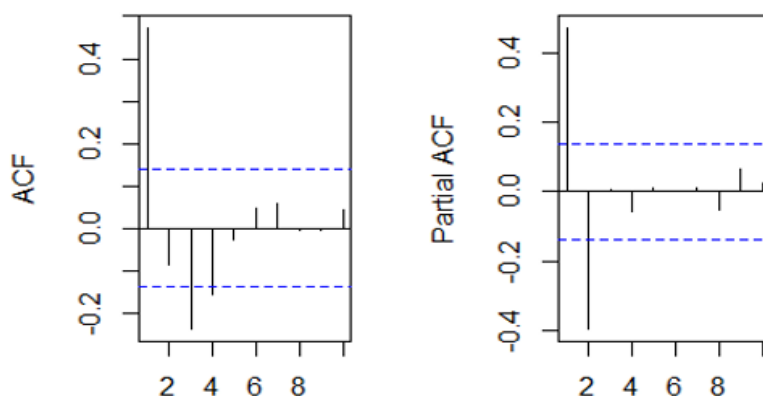
$$H_1 : \exists \rho_j \neq 0, j \in \{1, \dots, 10\}$$

s test veličinama $Q(10) = 63,64$, p -vrijednost $\approx 0 < \alpha$ za Box-Pierceov test te $Q^*(10) = 64,85$, p -vrijednost $\approx 0 < \alpha$ za Ljung-Boxov test (slika 1.100.), što znači da se odbacuje hipoteza o nepostojanju autokorelacije među članovima procesa Y3 do zaključno 10. reda. Procijenjen je AR(2) model i izračunati su korijeni inverzne karakteristične jednačbe (slika 1.101.), gdje se uočava da su korijeni po apsolutnoj vrijednosti veći od 1 pa je proces stacionaran, a po definiciji je proces invertibilan. Sam proces pišemo na sljedeći način: $y_t = 4,97 + 0,67y_{t-1} - 0,41y_{t-2} + \varepsilon_t$. Konačno, slika 1.102. predočava IRF funkciju za 10 razdoblja. Uočava se da šok od jedne

standardne devijacije u slučajnoj varijabli najprije pozitivno utječe na proces Y3, pri čemu je najjači učinak u prvome razdoblju i onda učinak ima sinusoidalno ponašanje do kraja promatranog razdoblja.



Slika 1.98. Grafički prikaz niza Y3



Slika 1.99. Korelogrami ACF i PACF funkcija za Y3

```
Box.test(y3,lag=10,type="Box-Pierce")
##
## Box-Pierce test
##
## data: y3
## X-squared = 63.64, df = 10, p-value = 7.371e-10

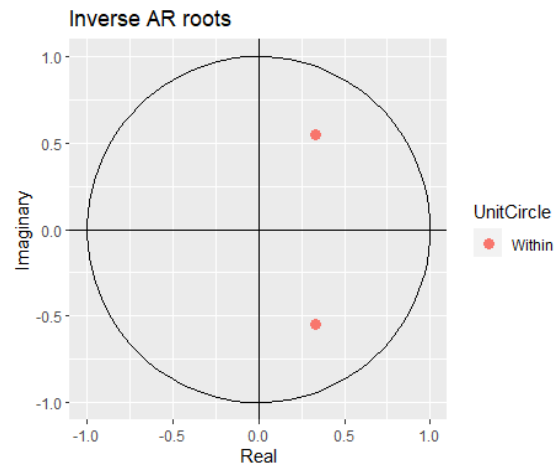
Box.test(y3,lag=10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: y3
## X-squared = 64.851, df = 10, p-value = 4.328e-10
```

Slika 1.100. Testiranje autokorelacije procesa Y3 do zaključno 10. reda

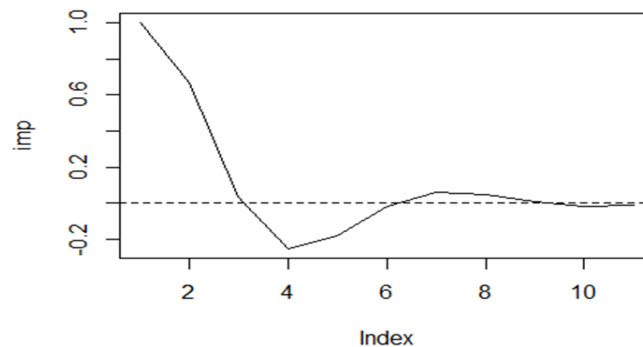

```

m1<-auto.arima(y3,d=0)
summary(m1)
## Series: y3
## ARIMA(2,0,0) with non-zero mean
##
## Coefficients:
##      ar1      ar2      mean
##      0.6661 -0.4107  4.9674
## s.e.  0.0649  0.0657  0.0285
##
## sigma^2 = 0.09111: log likelihood = -43.02
## AIC=94.03  AICc=94.24  BIC=107.23
##
korijeni<-autoplot(m1)
plot(korijeni)
korijeni$data$roots
## [1] 0.81109+1.333151i 0.81109-1.333151i

```



Slika 1.101. Procjena AR(2) procesa za Y3 i izračun korijena inverzne karakteristične jednadžbe



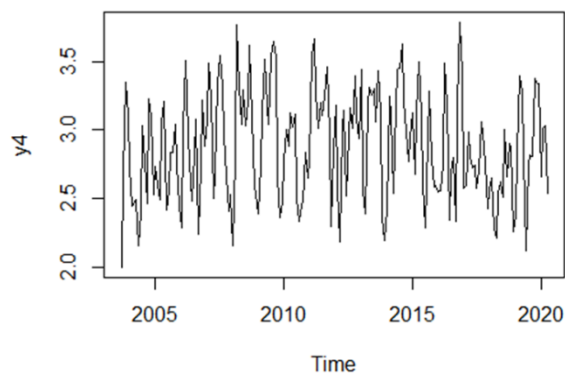
Slika 1.102. IRF funkcija za proces Y3

Slike 1.103. do 1.107. prikazuju rezultate postupaka a) do e) za niz „Y4“. Na slici 1.103. uočava se da je stacionaran proces. Nadalje, slika 1.104. predočuje korelograme, na kojima se uočava da koeficijenti ACF funkcije iščezavaju nakon prvog pomaka, a koeficijenti PACF funkcije se sinusoidalno smanjuju, što upućuje na MA(1) proces. Hipoteze testova autokorelacije su sljedeće:

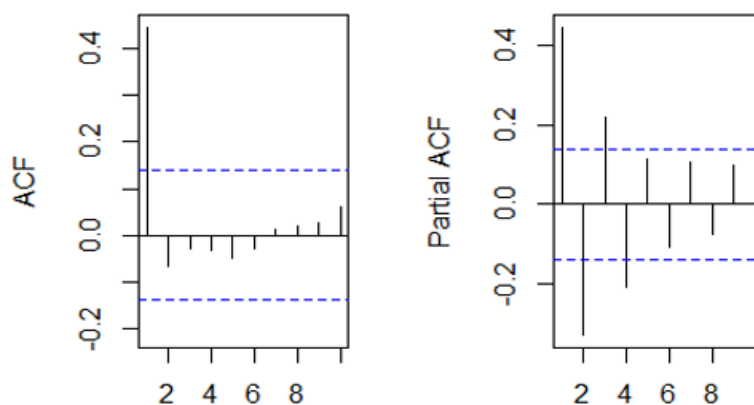
$$H_0 : \rho_1 = \dots = \rho_{10} = 0$$

$$H_1 : \exists \rho_j \neq 0, j \in \{1, \dots, 10\}$$

s test veličinama $Q(10) = 42,33$, p -vrijednost $\approx 0 < \alpha$ za Box-Pierceov test te $Q^*(10) = 43,03$, p -vrijednost $\approx 0 < \alpha$ za Ljung-Boxov test (slika 1.105.), što znači da se odbacuje hipoteza o nepostojanju autokorelacije među članovima procesa Y4 do zaključno 10. reda. Procijenjen je MA(1) model i izračunat je korijen inverzne karakteristične jednadžbe (slika 1.106.), gdje se uočava da je po apsolutnoj vrijednosti veći od 1 pa je proces invertibilan, a po definiciji je proces stacionaran. Sam proces pišemo na sljedeći način: $y_t = 2,88 + 0,88\varepsilon_{t-1} + \varepsilon_t$. Konačno, slika 1.107. predočava IRF funkciju za 10 razdoblja. Uočava se da šok od jedne standardne devijacije u slučajnoj varijabli najprije pozitivno utječe na proces Y4, pri čemu je najjači učinak u prvome razdoblju i učinak je još prisutan samo u idućem razdoblju, nakon čega nestaje.



Slika 1.103. Grafički prikaz niza Y4



Slika 1.104. Korelogrami ACF i PACF funkcija za Y4

```
Box.test(y4,lag=10,type="Box-Pierce")
##
## Box-Pierce test
##
## data: y4
## X-squared = 42.326, df = 10, p-value = 6.56e-06

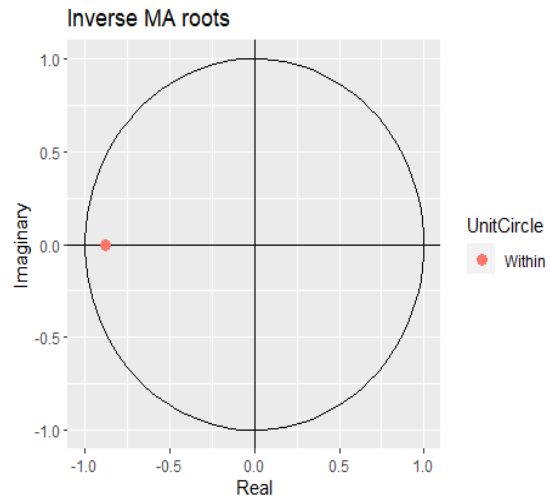
Box.test(y4,lag=10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: y4
## X-squared = 43.032, df = 10, p-value = 4.908e-06
```

Slika 1.105. Testiranje autokorelacije procesa Y4 do zaključno 10. reda

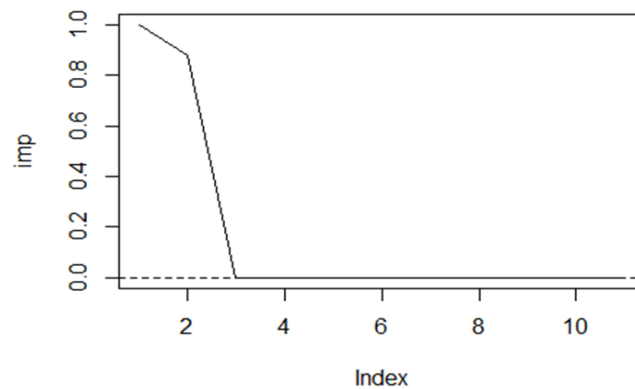
```

m1<-arima(y4,order=c(0,0,1))
summary(m1)
##
## Call:
## arima(x = y4, order = c(0, 0, 1))
##
## Coefficients:
##          ma1  intercept
##      0.8799    2.8817
## s.e. 0.0324    0.0394
##
## sigma^2 estimated as 0.08813:  log likel
## ihood = -41.64,  aic = 89.28
##
korijeni<-autoplot(m1)
plot(korijeni)
korijeni$data$roots
## [1] -1.136501+0i

```



Slika 1.106. Procjena⁷ MA(1) procesa za Y4 i izračun korijena inverzne karakteristične jednadžbe

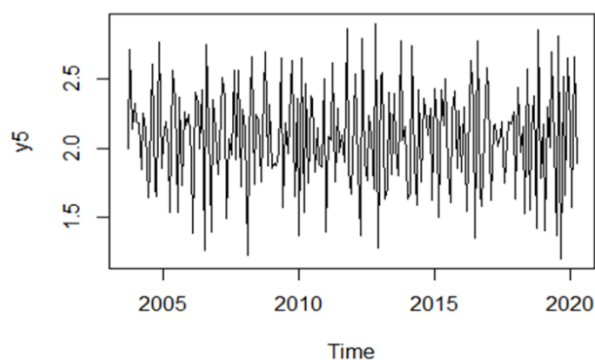


Slika 1.107. IRF funkcija za proces Y4

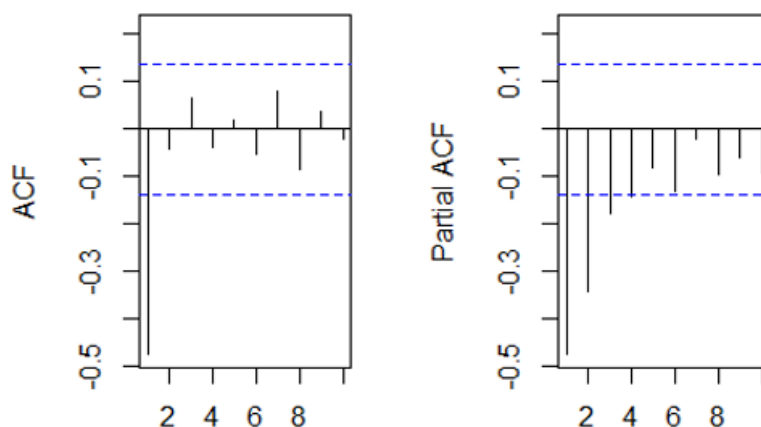
Slike 1.108. do 1.112. prikazuju rezultate postupaka a) do e) za niz „Y5“. Na slici 1.108. uočava se da je stacionaran proces, pri čemu postoji negativna autokorelacija (slično tumačenje kao za niz „Y2“). Nadalje, slika 1.109. predočuje korelograme, na kojima se uočava da koeficijenti ACF funkcije iščezavaju nakon prvog pomaka, a koeficijenti PACF funkcije se smanjuju eksponencijalno, što upućuje na MA(1) proces. Hipoteze testova autokorelacije jednake kao i do sada, s test veličinama $Q(10) = 50,07$, p -vrijednost $\approx 0 < \alpha$ za Box-Pierceov test te $Q^*(10) = 50,97$, p -vrijednost $\approx 0 < \alpha$ za Ljung-Boxov test (slika 1.110.), što znači da se odbacuje hipoteza o nepostojanju autokorelacije među članovima procesa Y5 do zaključno 10. reda. Procijenjen je MA(1) model i izračunat je korijen inverzne karakteristične jednadžbe (slika 1.111.), gdje se uočava da je po apsolutnoj vrijednosti veći od 1 pa je proces invertibilan, a po definiciji je proces stacionaran. Sam proces pišemo na sljedeći način: $y_t = 2,07 - 0,87\varepsilon_{t-1} + \varepsilon_t$. Konačno, slika 1.112. predočava IRF funkciju za 10 razdoblja. Uočava se da šok od jedne standardne

⁷ **Napomena:** Iako se koristila naredba `auto.arima(...)` koja automatski traži najbolji model po određenim kriterijima, niz Y4 je generiran kao MA(1) proces te ova naredba prema tim kriterijima rezultira s procesom koji nije MA(1). Za svrhu ovog primjera se želi procijeniti MA(1) proces, stoga se koristi naredba u kojoj korisnik sam definira o kakvom ARMA procesu se radi, na način da odabire pomak p , q , te je li potrebno ili ne diferencirati niz (d).

devijacije u slučajnoj varijabli najprije pozitivno utječe na proces Y5, pri čemu je najjači učinak u prvome razdoblju i učinak je još prisutan samo u idućem razdoblju pri čemu je negativan, nakon čega nestaje.



Slika 1.108. Grafički prikaz niza Y5



Slika 1.109. Korelogrami ACF i PACF funkcija za Y5

```
Box.test(y5,lag=10,type="Box-Pierce")
##
## Box-Pierce test
##
## data: y5
## X-squared = 50.068, df = 10, p-value = 2.593e-07

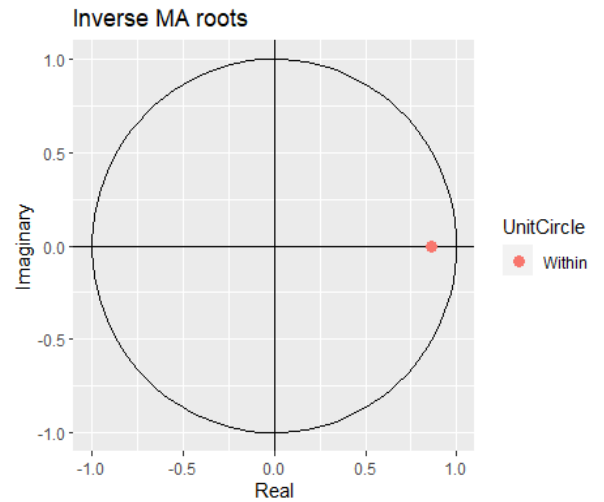
Box.test(y5,lag=10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: y5
## X-squared = 50.967, df = 10, p-value = 1.771e-07
```

Slika 1.110. Testiranje autokorelacije procesa Y5 do zaključno 10. reda

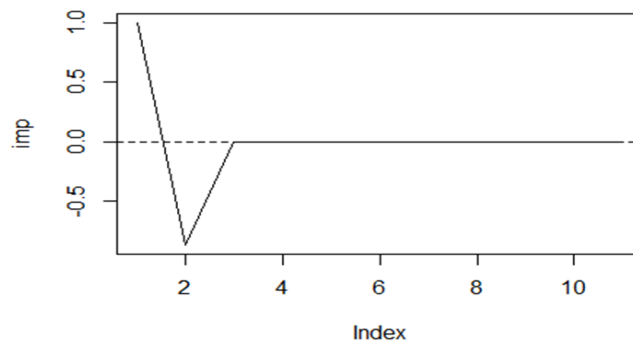
```

m1<-arima(y5,order=c(0,0,1))
summary(m1)
##
## Call:
## arima(x = y5, order = c(0, 0, 1))
##
## Coefficients:
##          ma1  intercept
##      -0.8636   2.0734
## s.e.   0.0384   0.0030
##
## sigma^2 estimated as 0.09202:  log likelihood =
## -45.9,  aic = 97.8
##
korijeni<-autoplot(m1)
plot(korijeni)
korijeni$data$roots
## [1] 1.157941+0i

```

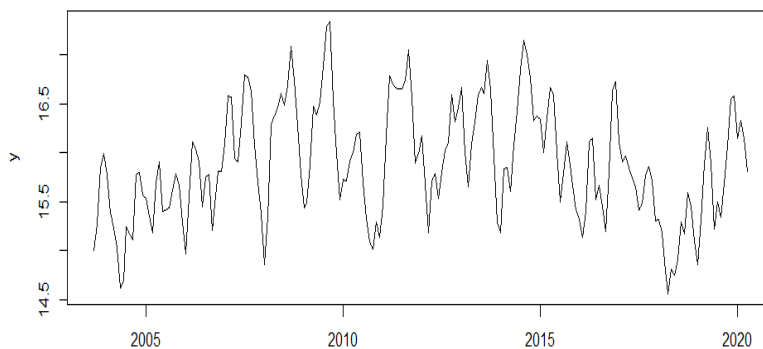


Slika 1.111. Procjena MA(1) procesa za Y5 i izračun korijena inverzne karakteristične jednadžbe

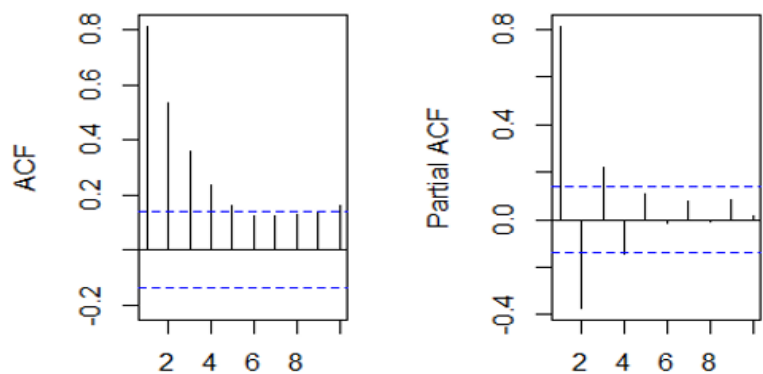


Slika 1.112. IRF funkcija za proces Y5

Slike 1.113. do 1.117. prikazuju rezultate postupaka a) do e) za niz „Y6“. Na slici 1.113. uočava se da je stacionaran proces. Nadalje, slika 1.114. predočuje korelograme, na kojima se uočava da se koeficijenti ACF funkcije eksponencijalno smanjuju nakon prvog pomaka, a koeficijenti PACF funkcije se smanjuju sinusoidalno nakon prvog pomaka također, što upućuje na ARMA(1,1) proces. Hipoteze testova autokorelacije jednake kao i do sada, s test veličinama $Q(10) = 248,79$, p -vrijednost $\approx 0 < \alpha$ za Box-Pierceov test te $Q^*(10) = 254,08$, p -vrijednost $\approx 0 < \alpha$ za Ljung-Boxov test (slika 1.115.), što znači da se odbacuje hipoteza o nepostojanju autokorelacije među članovima procesa Y6 do zaključno 10. reda. Procijenjen je ARMA(1,1) model i izračunati su korijeni inverzne karakteristične jednadžbe (slika 1.116.), gdje se uočava da su oba po apsolutnoj vrijednosti veći od 1 pa je proces invertibilan i stacionaran. Sam proces pišemo na sljedeći način: $y_t = 15,87 + 0,65y_{t-1} + 0,65\varepsilon_{t-1} + \varepsilon_t$. Konačno, slika 1.117. predočava IRF funkciju za 10 razdoblja. Uočava se da šok od jedne standardne devijacije u slučajnoj varijabli najprije pozitivno utječe na proces Y6, pri čemu je najjači učinak u drugome razdoblju i nakon toga se brzo smanjuje u idućim razdobljima.



Slika 1.113. Grafički prikaz niza Y6



Slika 1.114. Korelogrami ACF i PACF funkcija za Y6

```
Box.test(y6,lag=10,type="Box-Pierce")
##
## Box-Pierce test
##
## data: y6
## X-squared = 248.79, df = 10, p-value < 2.2e-16

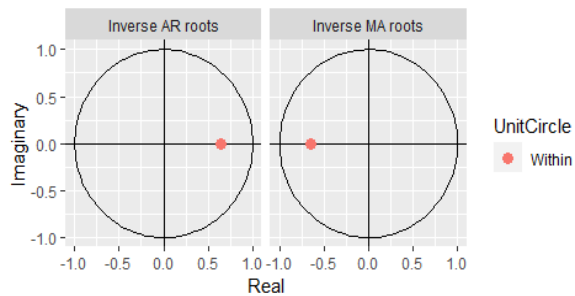
Box.test(y6,lag=10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: y6
## X-squared = 254.08, df = 10, p-value < 2.2e-16
```

Slika 1.115. Testiranje autokorelacije procesa Y6 do zaključno 10. reda

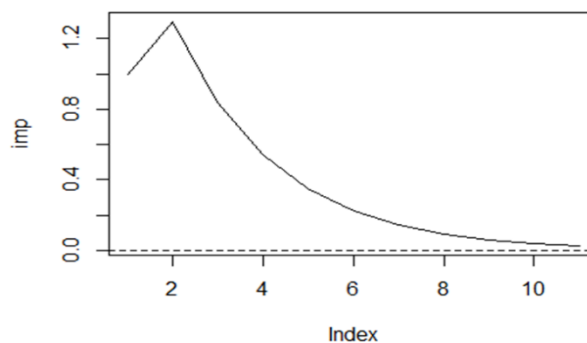
```

m1<-arima(y6,order=c(1,0,1))
summary(m1)
##
## Call:
## arima(x = y6, order = c(1, 0, 1))
##
## Coefficients:
##          ar1          ma1  intercept
##          0.6462  0.6499   15.8726
## s.e.      0.0601  0.0653    0.0961
##
## sigma^2 estimated as 0.08672:  log like
## likelihood = -40.17,  aic = 88.34
##
korijeni<-autoplot(m1)
plot(korijeni)
korijeni$data$roots
## [1] 1.547603+0i -1.538605+0i

```

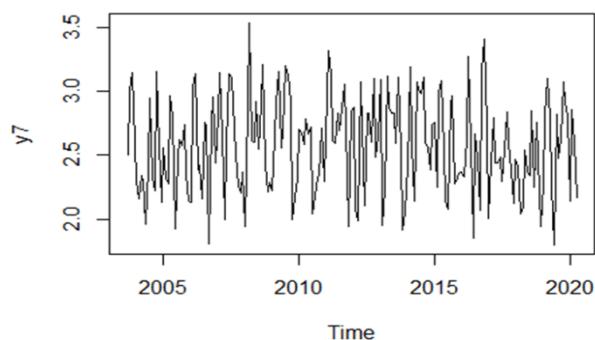


Slika 1.116. Procjena ARMA(1,1) procesa za Y6 i izračun korijena inverzne karakteristične jednačbe

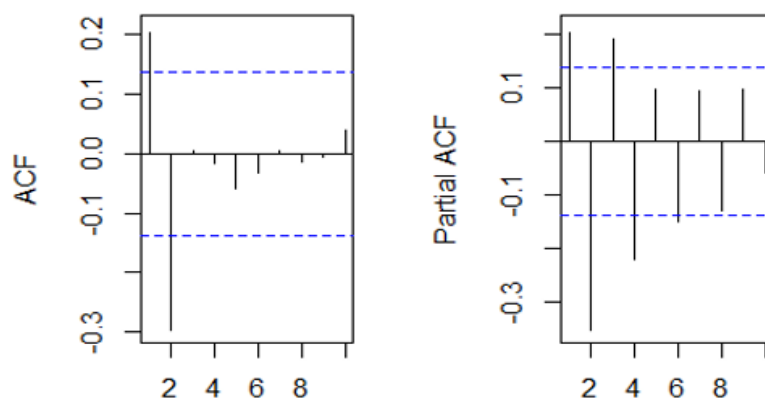


Slika 1.117. IRF funkcija za proces Y6

Slike 1.118. do 1.122. prikazuju rezultate postupaka a) do e) za niz „Y7“. Na slici 1.118. uočava se da je stacionaran proces, moguća je negativna autokorelacija s obzirom na „ponašanje“ procesa. Nadalje, slika 1.119. predočuje korelograme, na kojima se uočava da se koeficijenti ACF funkcije eksponencijalno iščezavaju nakon drugog pomaka, a koeficijenti PACF funkcije se smanjuju sinusoidalno nakon prvog pomaka također, što upućuje na MA(2) proces. Hipoteze testova autokorelacije jednake kao i do sada, s test veličinama $Q(10) = 25,67$, p -vrijednost $\approx 0 < \alpha$ za Box-Pierceov test te $Q^*(10) = 26,19$, p -vrijednost $\approx 0 < \alpha$ za Ljung-Boxov test (slika 1.120.), što znači da se odbacuje hipoteza o nepostojanju autokorelacije među članovima procesa Y7 do zaključno 10. reda. Procijenjen je MA(2) model i izračunati su korijeni inverzne karakteristične jednačbe (slika 1.121.), gdje se uočava da je jedan veći od 1, dok je drugi po apsolutnoj vrijednosti jednak 1. To znači da je narušeno svojstvo invertibilnosti. Više o daljnjim postupcima analize ovakvog procesa bit će u odjeljku 1.9.3.4. Sam proces za sada pišemo na sljedeći način: $y_t = 2,58 + 0,53y_{t-1} - 0,47\varepsilon_{t-1} + \varepsilon_t$. Konačno, slika 1.122. predočava IRF funkciju za 10 razdoblja. Uočava se da šok od jedne standardne devijacije u slučajnoj varijabli najprije u prva dva razdoblja pozitivno utječe na proces Y7, pri čemu je najjači učinak u drugome razdoblju i nakon toga u trećem razdoblju je učinak negativan, nakon čega nestaje nadalje.



Slika 1.118. Grafički prikaz niza Y7



Slika 1.119. Korelogrami ACF i PACF funkcija za Y7

```
Box.test(y7,lag=10,type="Box-Pierce")
##
## Box-Pierce test
##
## data: y7
## X-squared = 27.308, df = 10, p-value = 0.002328

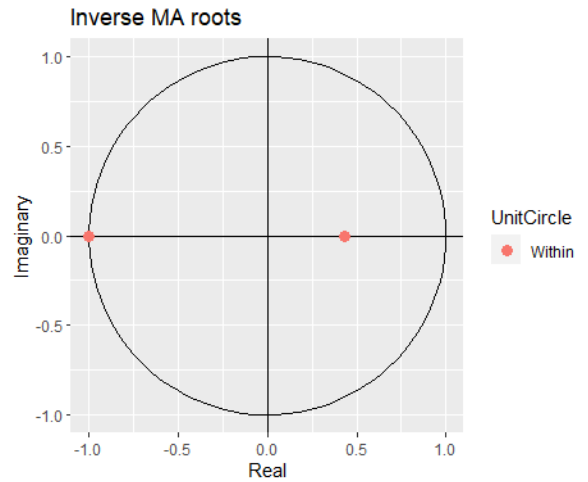
Box.test(y7,lag=10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: y7
## X-squared = 27.848, df = 10, p-value = 0.001909
```

Slika 1.120. Testiranje autokorelacije procesa Y7 do zaključno 10. reda

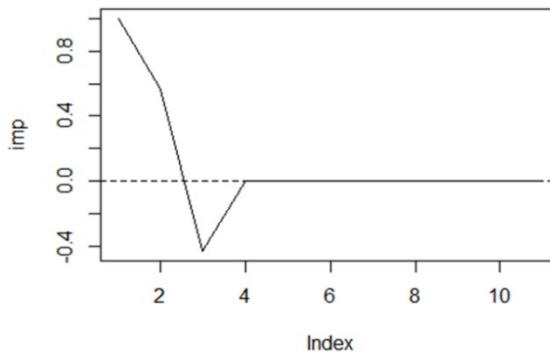

```

m1<-arima(y7,order=c(0,0,2))
summary(m1)
##
## Call:
## arima(x = y7, order = c(0, 0, 2))
##
## Coefficients:
##      ma1      ma2  intercept
##  0.5679 -0.4321   2.5787
## s.e.  0.0688  0.0677   0.0236
##
## sigma^2 estimated as 0.08607:  log likelihood
## = -40.93,  aic = 89.85
##
korijeni<-autoplot(m1)
plot(korijeni)
korijeni$data$roots
## [1] -1.00000-0i  2.31421+0i

```



Slika 1.121. Procjena MA(2) procesa za Y6 i izračun korijena inverzne karakteristične jednačbe



Slika 1.122. IRF funkcija za proces Y7

Napomena: uočimo da smo za sve procese kod testiranja autokorelacije zaključili da odbacujemo nultu hipotezu o nekoreliranosti članova procesa do zaključno nekog reda. Jedino za slučaj čistog slučajnog procesa ne možemo odbaciti nultu hipotezu (vidjeti definiciju i svojstva čistog slučajnog procesa).

1.9.10. Pitanja za ponavljanje

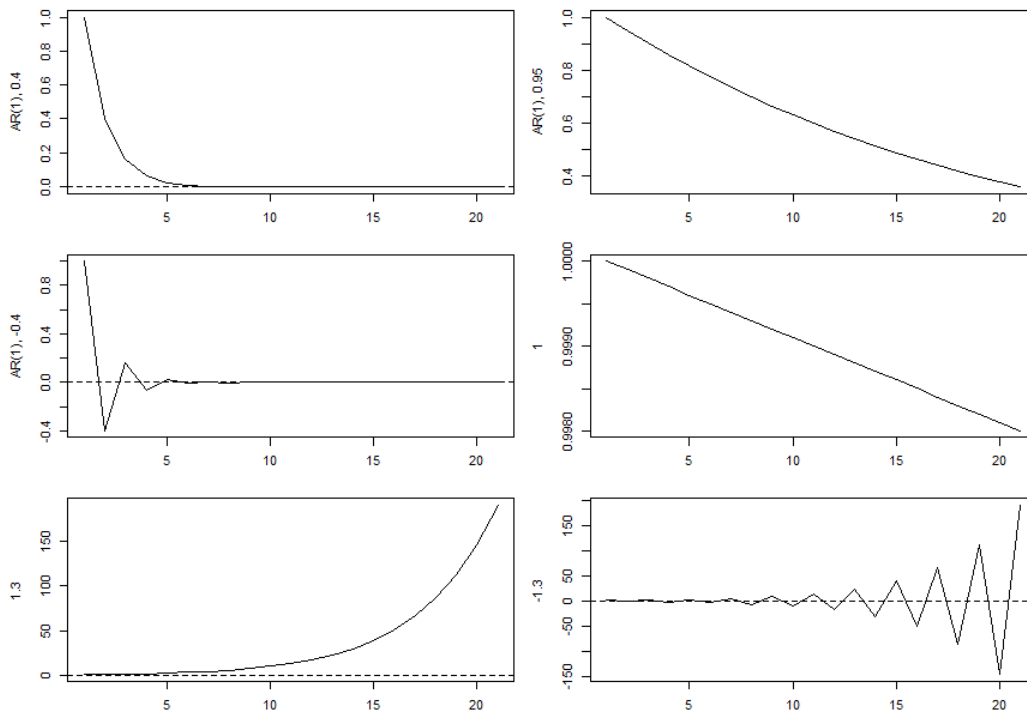
1. Definirajte stacionaran vremenski niz.
2. Definirajte striktno stacionaran vremenski niz.
3. Definirajte slabo stacionaran vremenski niz.
4. Zapišite simbolički tri uvjeta slabo stacionarnih vremenskih nizova.
5. Što je to autokorelacijska funkcija?
6. Što je to parcijalna autokorelacijska funkcija?
7. Što je to korelogram?
8. Kako se provodi test značajnosti koeficijenta autokorelacije vremenskog niza?
9. Zapišite nultu hipotezu testa značajnosti koeficijenta autokorelacije nekog niza do zaključno 8. reda.
10. Što je to čisti slučajni proces? Koje su njegove karakteristike?
11. Kako se ponašaju autokorelacijska i parcijalna autokorelacijska funkcija čistog slučajnog procesa?
12. Što je to autoregresijski proces reda p ?
13. Što je to operator pomaka L ?
14. Koji je uvjet koji mora vrijediti da bi autoregresijski proces bio stacionaran?
15. Što je to autoregresijski proces reda 1? Zapišite ga simbolički.
16. Koja su svojstva autokorelacijske i parcijalne autokorelacijske funkcije autoregresijskog procesa reda 1?
17. Što je to autoregresijski proces reda 2? Zapišite ga simbolički.
18. Koja su svojstva autokorelacijske i parcijalne autokorelacijske funkcije autoregresijskog procesa reda 2?
19. Zapišite simbolički autoregresijski proces reda p .
20. Koja su svojstva autokorelacijske i parcijalne autokorelacijske funkcije autoregresijskog procesa reda p ?
21. Iskažite Woldov teorem dekompozicije.
22. Što znači da je stohastički proces invertibilan?
23. Što je to proces pomičnih prosjeka?
24. Koji je uvjet koji mora vrijediti da bi proces pomičnih prosjeka bio invertibilan?
25. Što je to proces pomičnih prosjeka reda 1? Zapišite ga simbolički.
26. Koja su svojstva autokorelacijske i parcijalne autokorelacijske funkcije procesa pomičnih prosjeka reda 1?
27. Što je to proces pomičnih prosjeka reda 2? Zapišite ga simbolički.
28. Koja su svojstva autokorelacijske i parcijalne autokorelacijske funkcije procesa pomičnih prosjeka reda 2?
29. Zapišite simbolički proces pomičnih prosjeka reda q .
30. Koja su svojstva autokorelacijske i parcijalne autokorelacijske funkcije procesa pomičnih prosjeka reda q ?
31. Što je to autoregresijski proces pomičnih prosjeka ARMA(p,q)? Zapišite ga simbolički.
32. Zapišite simbolički ARMA(1,1) proces. Koji uvjeti moraju vrijediti da bi bio stacionaran i da bi bio invertibilan?
33. Definirajte funkciju impulsnog odziva za ARMA procese. Što se njome predočuje? Što se sve interpretira kada se promatra funkcija impulsnog odziva?
34. Klasificirajte koji od sljedećih procesa je AR(p), MA(q), ARMA(p,q) te odredite p i q :

a) $y_t = 2 + 0,3y_{t-1} + 0,2\varepsilon_{t-1} + \varepsilon_t$	d) $y_t = -0,3y_{t-1} + 0,8y_{t-2} + \varepsilon_t$
b) $y_t = -0,9\varepsilon_{t-1} + \varepsilon_t$	e) $y_t = 0,1\varepsilon_{t-1} + 0,4\varepsilon_{t-2} + \varepsilon_t$
c) $y_t = -0,3y_{t-1} + 0,8y_{t-2} + 0,4\varepsilon_{t-1} + \varepsilon_t$	f) $y_t = 0,1y_{t-1} + 0,4\varepsilon_{t-1} + 0,4\varepsilon_{t-2} + \varepsilon_t$

35. U RStudio učitajte datoteku „vremenski2.txt“. Datoteka sadrži 200 opservacija za 4 različita procesa (Y1 do Y4 nazivi). Definirajte za svaki proces da se radi o vremenskom nizu, s početnom vrijednošću u rujnu 2003. godine, podaci su mjesečni te potom definirajte novih 7 nizova na način da ćete izračunati recipročnu vrijednost svakog niza.

- Grafički predočite svaki niz i komentirajte njegovu vremensku dinamiku.
- Predočite korelograme za ACF i PACF funkcije i komentirajte o kakvom procesu se radi temeljem ponašanja koeficijenata obiju funkcija.
- Provedite Box-Pierceov i Ljung-Boxov test autokorelacije do zaključno 10. reda, uz razinu značajnosti od 5%.

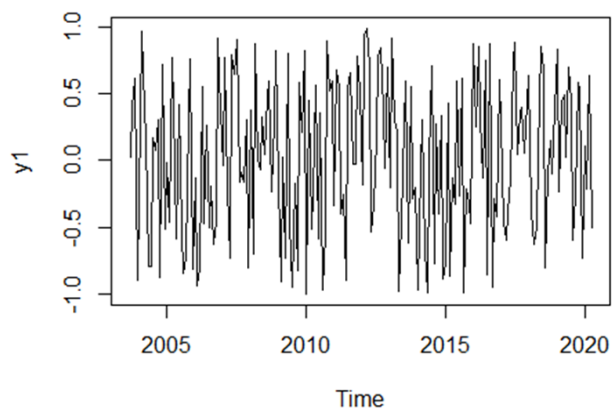
36. Komentirajte funkcije impulsnog odziva za sljedeće procese:



Rješenja

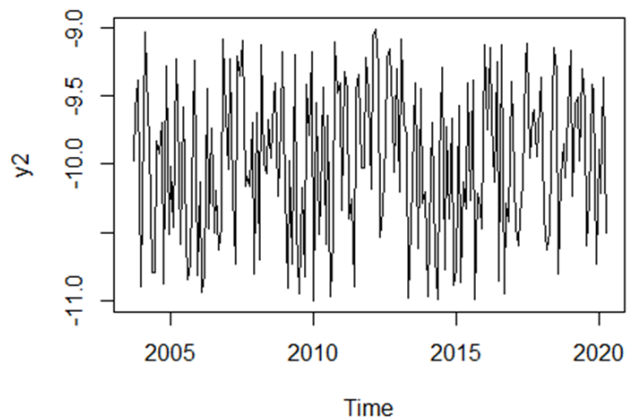
Zadatak 35)

a)



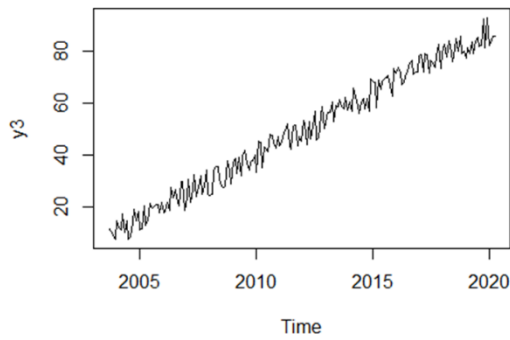
Niz Y1:

Niz Y1 izgleda kao stacionaran proces, moguće određeni MA proces, ili čisti slučajni proces.



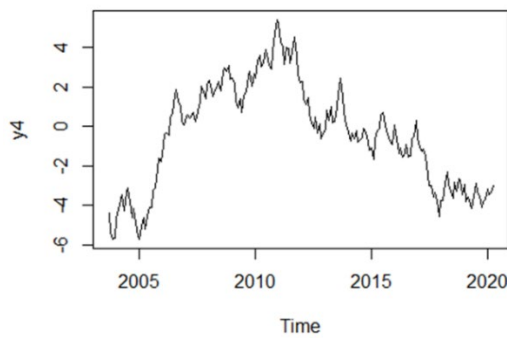
Niz Y2:

Niz Y2 izgleda kao stacionaran proces, moguće određeni MA proces, ili čisti slučajni proces



Niz Y3:

Niz Y3 izgleda kao nestacionaran proces, s linearnim trendom.

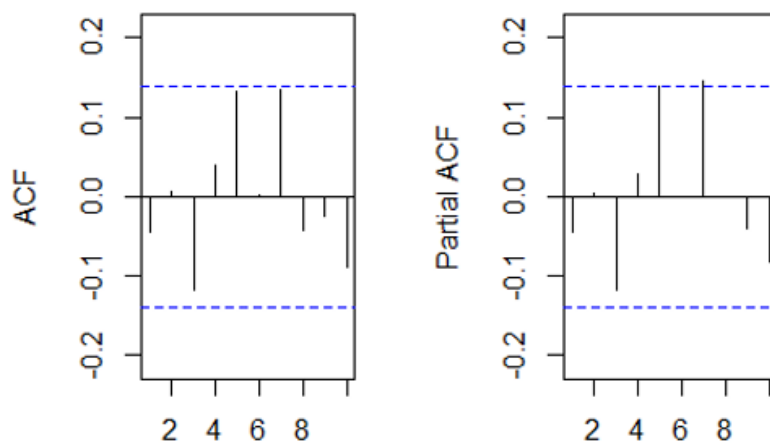


Niz Y4:

Niz Y4 izgleda kao nestacionaran proces.

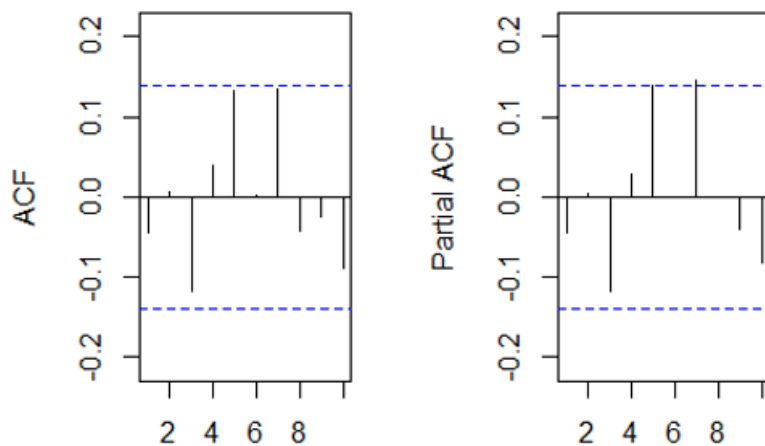
b)

Niz Y1:



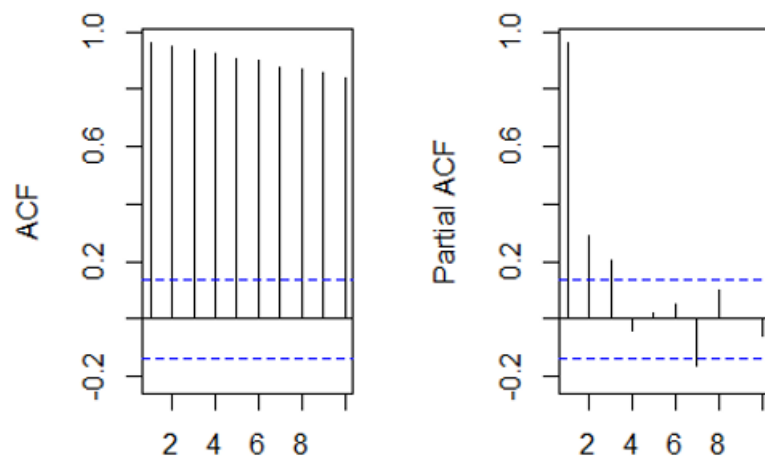
Obje funkcije imaju neznačajne koeficijente, i ACF i PACF. Proces izgleda kao čisti slučajni proces.

Niz Y2:



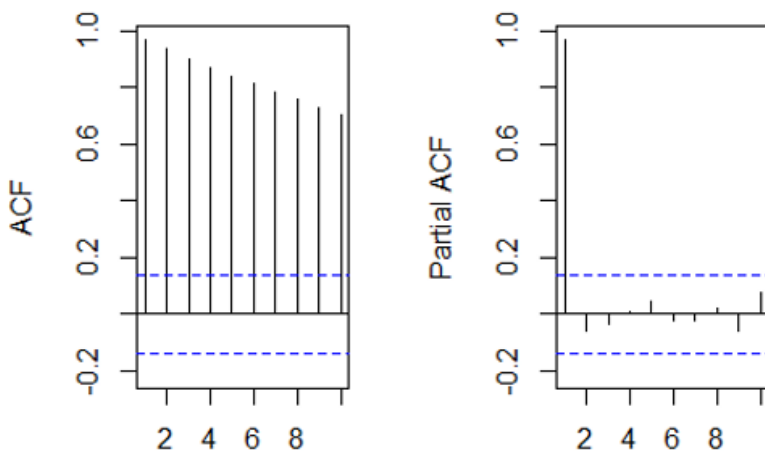
Obje funkcije imaju neznačajne koeficijente, i ACF i PACF. Proces izgleda kao čisti slučajni proces.

Niz Y3:



Autokorelacijska funkcija pada jako sporo, što nije karakteristika stacionarnih procesa.

Niz Y4:



Autokorelacijska funkcija pada jako sporo, što nije karakteristika stacionarnih procesa.

c)

Niz Y1:

```
Box.test(y1,lag=10,type="Box-Pierce")

##
## Box-Pierce test
##
## data: y1
## X-squared = 12.592, df = 10, p-value = 0.2474

Box.test(y1,lag=10,type="Ljung-Box")

##
## Box-Ljung test
##
## data: y1
## X-squared = 13.096, df = 10, p-value = 0.2184
```

Uz razinu značajnosti od 5%, ne odbacujemo nulte hipoteze oba testa da ne postoji autokorelacija procesa Y1 do zaključno 10. reda.

Niz Y2:

```
Box.test(y2,lag=10,type="Box-Pierce")

##
## Box-Pierce test
##
## data: y2
## X-squared = 12.592, df = 10, p-value = 0.2474

Box.test(y2,lag=10,type="Ljung-Box")

##
## Box-Ljung test
##
## data: y2
## X-squared = 13.096, df = 10, p-value = 0.2184
```

Uz razinu značajnosti od 5%, ne odbacujemo nulte hipoteze oba testa da ne postoji autokorelacija procesa Y2 do zaključno 10. reda.

Niz Y3:

```
Box.test(y3,lag=10,type="Box-Pierce")

##
## Box-Pierce test
##
## data: y3
## X-squared = 1629.9, df = 10, p-value < 2.2e-16

Box.test(y3,lag=10,type="Ljung-Box")

##
## Box-Ljung test
##
## data: y3
## X-squared = 1690.9, df = 10, p-value < 2.2e-16
```

Uz razinu značajnosti od 5%, odbacujemo nulte hipoteze oba testa da ne postoji autokorelacija procesa Y3 do zaključno 10. reda.

Niz Y4:

```
Box.test(y4, lag=10, type="Box-Pierce")

##
## Box-Pierce test
##
## data: y4
## X-squared = 1396.5, df = 10, p-value < 2.2e-16

Box.test(y4, lag=10, type="Ljung-Box")

##
## Box-Ljung test
##
## data: y4
## X-squared = 1446.3, df = 10, p-value < 2.2e-16
```

Uz razinu značajnosti od 5%, odbacujemo nulte hipoteze oba testa da ne postoji autokorelacija procesa Y4 do zaključno 10. reda.

Zadatak 36)

Panel prvi redak i prvi stupac: uočavamo da je učinak pozitivan, najjači u prvom razdoblju, te vrlo brzo iščezava (nakon petog pomaka je učinak na nuli).

Panel prvi redak i drugi stupac: učinak je pozitivan, najjači u prvom razdoblju, te je puno jači u odnosu na prvi panel, s obzirom da veoma sporo iščezava sam učinak šoka.

Panel drugi redak i prvi stupac: učinak varira, tj. sinusoidalno, najjači je u prvom razdoblju i nakon toga brzo iščezava.

Panel drugi redak i drugi stupac: učinak je pozitivan, i veoma jak, čak i u 20. razdoblju, vjerojatno se radi o nestacionarnom procesu s obzirom da je i u 20. razdoblju učinak gotovo blizu jedinične vrijednosti.

Panel treći redak i prvi stupac: radi se o nestacionarnom procesu, jer je šok sve jači i jači kroz vrijeme.

Panel treći redak i drugi stupac: također se radi o nestacionarnom procesu, pri čemu šok alternira u predznacima i divergira kroz vrijeme.

1.10. Nestacionarni vremenski nizovi

Nestacionarni vremenski nizovi su oni čija je **promjenjiva očekivana vrijednost ili varijanca tijekom vremena**. Ekonomski nizovi u pravilu su nestacionarni, stoga ih je prije formalne analize potrebno provesti u stacionarni oblik. Pritom je potrebno razlikovati **dvije vrste nestacionarnosti**, koje proizlaze iz vrste trenda koji je prisutan u samome nizu: deterministički i stohastički trend. Svaki trend se detaljnije obrađuje u idućim odjeljcima. Inicijalna indikacija da se radi o nestacionarnom nizu jest korelogram autokorelacijske funkcije, gdje je moguće uočiti veoma **sporo smanjivanje** koeficijenata te funkcije. Međutim, postoje formalni testovi koji se provode kako bi se testirala nestacionarnost nekog procesa, a koji će se detaljnije obraditi u odjeljku 1.10.4.

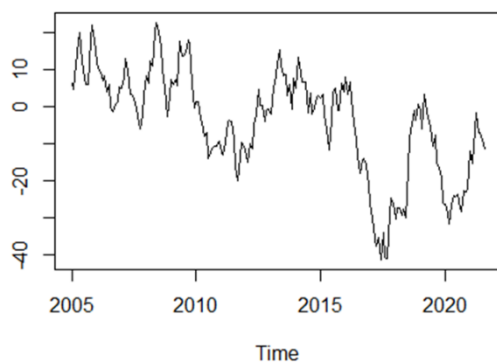
Problem koji se javlja u regresijskoj analizi ako se koriste nestacionarni nizovi naziva se problem **privedne regresije**, te se obrađuje u odjeljku 1.10.6. Dodatan problem koji se javlja jest da uobičajene pretpostavke regresijskog modela za asimptotsku analizu nisu validne, što znači da t -test, F -test i druga već spominjana inferencijalna statistička analiza (u prvome dijelu ovih publikacija) nije pouzdana za korištenje.

Primjer 1.21.

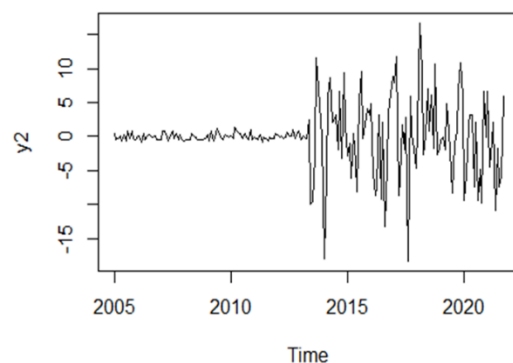
Slika 1.123. predočava nekoliko nestacionarnih procesa: panel a) predočava niz kojemu se očekivana vrijednost mijenja, panel b) predočava niz čija se varijanca mijenja tijekom vremena. Korelogrami ACF funkcija za oba niza su predočeni na slici 1.124., gdje se uočava da niz kojemu se očekivana vrijednost mijenja tijekom vremena ima karakteristiku veoma sporog smanjivanja vrijednosti koeficijenata ACF funkcije. S druge strane, za niz kojemu se varijanca mijenja tijekom vremena to nije uočljivo, i zato je potrebno provesti formalne testove nestacionarnosti kako bi se utvrdila (ne)stacionarnost nekog niza.

Konačno, za oba niza su provedeni Box-Pierceov i Ljung-Boxov test autokorelacije do zaključno 10. pomaka, te su rezultati predočeni na slici 1.125., gdje se uočava da uz razinu značajnosti od 10% odbacujemo hipotezu o nepostojanju autokorelacije među članovima oba procesa do zaključno 10. pomaka.

(a) Niz kojemu se očekivana vrijednost mijenja



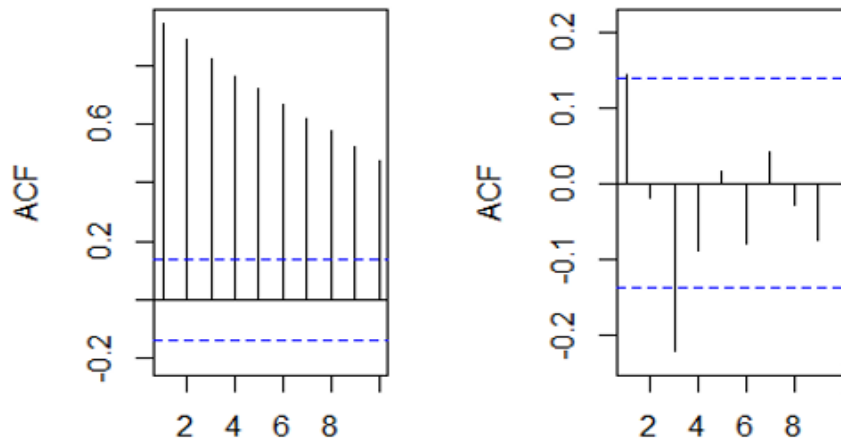
(b) Niz kojemu se varijanca mijenja tijekom vremena



Slika 1.123. Nestacionarni procesi

(a) Niz kojemu se očekivana vrijednost mijenja

(b) Niz kojemu se varijanca mijenja tijekom vremena



Slika 1.124. Korelogrami ACF funkcija nestacionarnih procesa

```
## Box.test(y1,lag=10,type="Box-Pierce")
##
## Box-Pierce test
##
## data: y1
## X-squared = 1025, df = 10, p-value < 2.2e-
16
Box.test(y1,lag=10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: y1
## X-squared = 1058.4, df = 10, p-value <
2.2e-16

## Box.test(y2,lag=10,type="Box-Pierce")
##
## Box-Pierce test
##
## data: y2
## X-squared = 29.6, df = 10, p-value =
0.0009957
Box.test(y2,lag=10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: y2
## X-squared = 30.796, df = 10, p-value =
0.000634
```

Slika 1.125. Testiranje autokorelacije do 10. pomaka za nizove sa slike 1.123.

1.10.1. Deterministički trend

Najčešći uzrok nestacionarnosti vremenskih nizova jest postojanje trenda. Trend može biti deterministički ili stohastički. Deterministički trend je prisutan ako je vremenski niz moguće zapisati kao funkciju vremena t :

$$y_t = f(t) + \varepsilon_t, \quad (1.120)$$

gdje funkcijski oblik može biti bilo koji. Najčešće se radi o nekome polinomu:

$$y_t = \sum_{j=0}^J \delta_j t^j + \varepsilon_t. \quad (1.121)$$

Ako se razmatraju linearni, kvadratni i kubni trend, model (1.121) pišemo na sljedeće načine:

$$y_t = \delta_0 + \delta_1 t + \varepsilon_t, \quad (1.122)$$

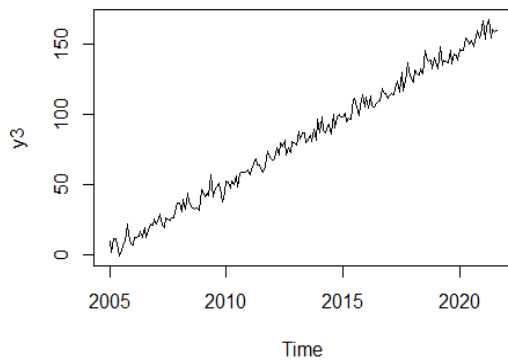
$$y_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \varepsilon_t, \quad (1.123)$$

$$y_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \varepsilon_t, \quad (1.124)$$

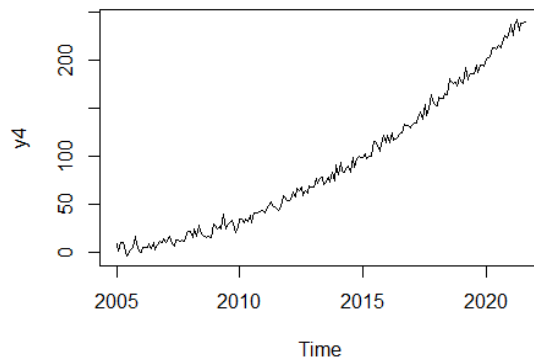
Primjer 1.22.

Grafički prikazi za modele (1.122) i (1.123) predloženi na slici 1.126., dok su pripadajući korelogrami ACF funkcija predloženi na slici 1.127., gdje se uočava izuzetno sporo smanjivanje koeficijenta spomenute funkcije što upućuje na nestacionarne procese.

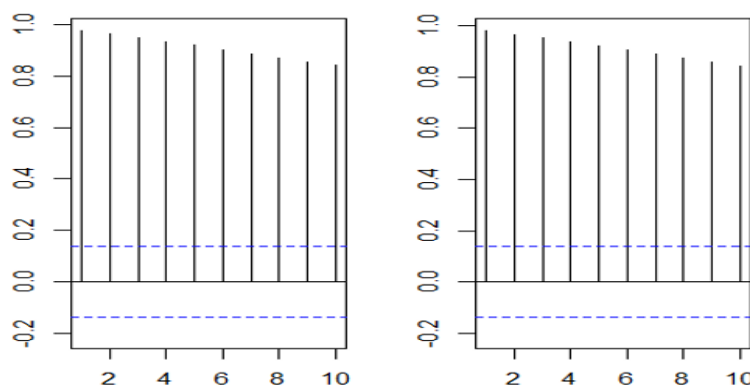
(a) linearni trend



(b) kvadratni trend



Slika 1.126. Linearni i kvadratni trend



Slika 1.127. Korelogrami ACF funkcija za procese sa slike 1.126.

Nizovi koje je moguće opisati modelom (1.120) nazivaju se **trend-stacionarni procesi**, jer se vrijednosti y_t u dugome roku kreću oko promjenjive očekivane vrijednosti. Očekivane vrijednosti linearnog (1.122), kvadratnog (1.123) i kubnog (1.124) modela iznose:

$$E(y_t) = \delta_0 + \delta_1 t, \quad (1.125)$$

$$E(y_t) = \delta_0 + \delta_1 t + \delta_2 t^2, \quad (1.126)$$

$$E(y_t) = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3, \quad (1.127)$$

stoga nije zadovoljeno svojstvo konstantnog očekivanja stacionarnih procesa. Općenito je očekivanje procesa (1.121) jednako $E(y_t) = \sum_{j=1}^J \delta_j t^j$.

Uklanjanje determinističkog trenda (detrendiranje) provodi se na sljedeći način. Najprije se procijeni model (1.121), prikupe se procijenjene vrijednosti procesa y_t :

$$\hat{y}_t = \sum_{j=0}^J \hat{\delta}_j t^j, \quad (1.1287)$$

te se procijene rezidualna odstupanja:

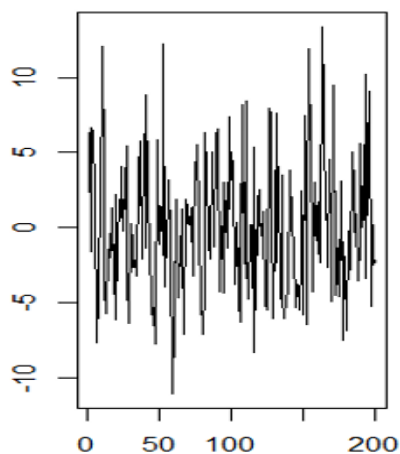
$$\hat{\varepsilon}_t = y_t - \hat{y}_t = y_t - \sum_{j=0}^J \hat{\delta}_j t^j, \quad (1.129)$$

koja su stacionarna i daljnja analiza vrši se nad tim odstupanjima.

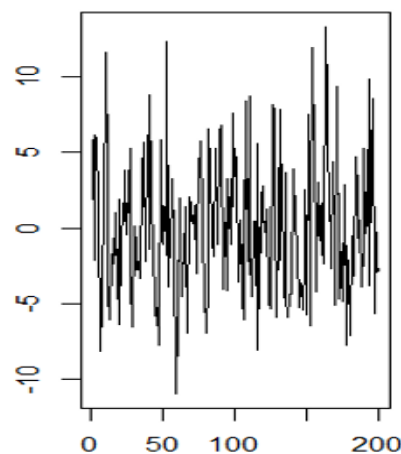
Primjer 1.23.

Slika 1.128. predočava rezidualne iz oba modela, gdje se uočava odsustvo linearnog, odnosno paraboličnog trenda. Korelogrami oba reziduala predočeni su na slici 1.129., gdje se uočava da su koeficijenti obiju funkcija neznačajni, što upućuje na čisti slučajni proces. Konačno, Ljung-Boxov test proveden je za oba niza reziduala do zaključno 10. pomaka na slici 1.130. gdje se uočava da se ne može odbaciti nulta hipoteza o nepostojanju autokorelacije među članovima procesa reziduala na uobičajenim razinama značajnosti.

(a) linearni trend

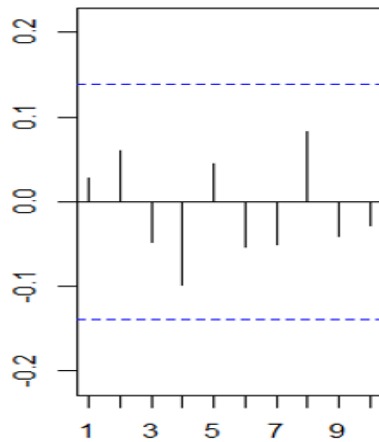


(b) kvadratni trend

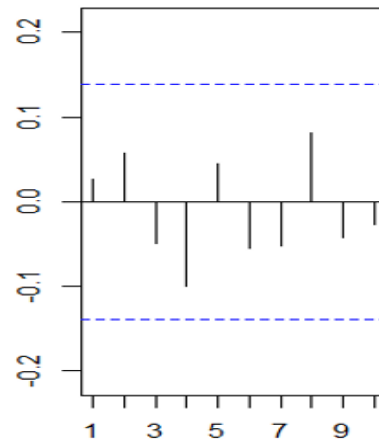


Slika 1.128. Reziduali dobiveni detrendiranjem modela (1.122) i (1.123)

(a) linearni trend



(b) kvadratni trend



Slike 1.129. Korelogrami ACF funkcija za rezidualne

```
Box.test(rez1,lag = 10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: rez1
## X-squared = 6.8487, df = 10, p-value = 0.7397
Box.test(rez2,lag = 10,type="Ljung-Box")
##
## Box-Ljung test
##
## data: rez2
```

Slika 1.130. Ljung-Boxov test nad rezidualima

1.10.2. Stohastički trend

Ako proces karakterizira stohastičko mijenjanje očekivane razine same pojave, u tom slučaju govorimo o stohastičkom trendu. Općenito se može zapisati sljedeći oblik procesa sa stohastičkim trendom:

$$y_t = \phi_0 + \sum_{\tau=1}^t \varepsilon_{\tau}, \quad (1.130)$$

čija očekivana vrijednost iznosi $E(y_t) = \phi_0$ (konstantna je), a varijanca:

$$\text{Var}(y_t) = \sigma^2 + \sigma^2 + \dots + \sigma^2 = t\sigma^2, \quad (1.131)$$

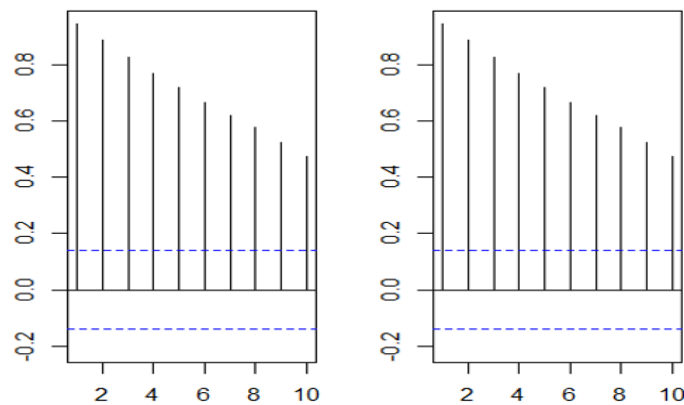
pa je narušeno svojstvo nepromjenjivosti varijance stacionarnog procesa.

Primjer 1.24.

Slika 1.131. predočava dva procesa koji sadržavaju stohastički trend, dok su njihovi korelogrami ACF funkcija predočeni na slici 1.132., čije sporo smanjivanje upućuje na nestacionarne procese. Sukladno tome, testiranje autokorelacije pomoću Box-Pierceova ili Ljung-Boxova testa bi ukazivalo na odbacivanje nulte hipoteze za ovakve nizove.



Slika 1.131. Dva niza stohastičkih trendova



Slika 1.132. Korelogrami ACF funkcija procesa na slici 1.131.

Uklanjanje stohastičkog trenda provodi se **diferenciranjem**. Oznaka za prvu diferenciju nekog niza je Δ , te se prva diferencija računa ovako:

$$\Delta y_t = y_t - y_{t-1}, \quad (1.132)$$

a druga diferencija, u oznaci Δ^2 :

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = y_t - y_{t-1} - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}, \quad (1.133)$$

diferencija reda d , u oznaci Δ^d , $d \in \mathbb{N}$, računa se izrazom:

$$\Delta^d y_t = \Delta^{d-1} y_t - \Delta^{d-1} y_{t-1}. \quad (1.134)$$

Stohastički niz se definira kao **integriran reda d** , ako ga je potrebno d puta diferencirati da bi postao stacionaran, u oznaci $y_t \sim I(d)$. Uočimo da je stacionaran proces integriran reda 0 jer ga je potrebno 0 puta diferencirati (tj. nije ga potrebno diferencirati) da bi postao stacionaran.

Ako se razmatra proces (1.130), te se pomakne za razdoblje $t-1$:

$$y_{t-1} = \phi_0 + \sum_{\tau=0}^{t-1} \varepsilon_{\tau}, \quad (1.135)$$

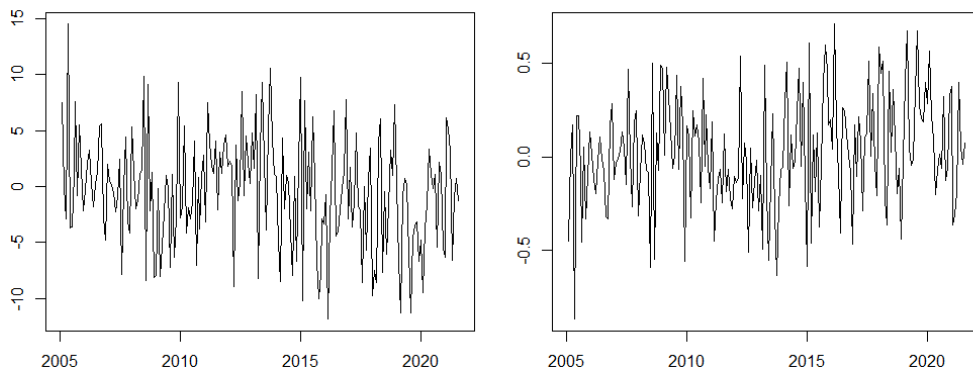
i izračuna se prva diferencija $\Delta y_t = y_t - y_{t-1}$:

$$\begin{aligned} \Delta y_t = y_t - y_{t-1} &= \phi_0 + \sum_{\tau=1}^t \varepsilon_{\tau} - \phi_0 - \sum_{\tau=0}^{t-1} \varepsilon_{\tau} = \\ &= \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_{t-1} + \varepsilon_t - \varepsilon_0 - \varepsilon_1 - \dots - \varepsilon_{t-1} = \varepsilon_t - \varepsilon_0 \end{aligned}, \quad (1.136)$$

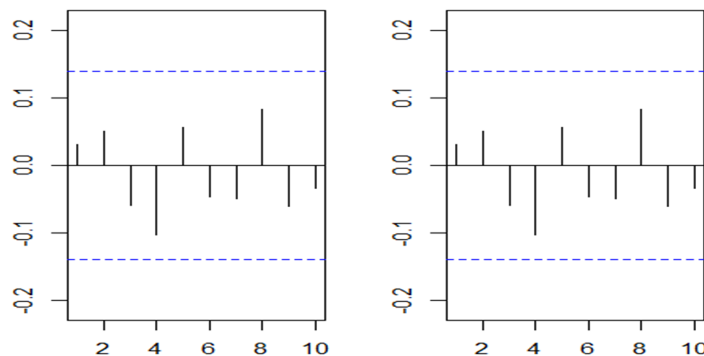
očekivana vrijednost izraza (1.136) jednaka je 0, što zadovoljava svojstvo konstantnog očekivanja stacionarnog procesa.

Primjer 1.24.

Promotrimo na slici 1.133. diferencirane nizove sa slike 1.131., gdje su oba niza diferencirani jedanput. Uočava se da prve diferencije za svaki niz zadovoljavaju definiciju stacionarnog procesa. Nadalje, korelogrami ACF funkcija na slici 1.134. također upućuju na stacionarnost oba procesa. Sukladno tome, testiranje autokorelacije pomoću Box-Pierceova ili Ljung-Boxova testa bi ukazivalo na ne odbacivanje nulte hipoteze za ovakve nizove.



Slika 1.133. Diferencirani nizovi sa slike 1.131.



Slika 1.134. Korelogrami ACF funkcija procesa na slici 1.132.

1.10.3. Pogrešno uklanjanje trenda

1.10.3.1. Detrendiranje stohastičkog trenda

Može se postaviti pitanje što ako se koristi kriva metoda za uklanjanje određenog trenda. Primjerice, ako je u nizu prisutan stohastički trend koji se uklanja diferenciranjem, a primijeni se detrendiranje determinističkog trenda.

Razmotrimo u tu svrhu proces sa stohastičkim trendom:

$$y_t = \phi_0 + \sum_{\tau=1}^t \varepsilon_\tau, \quad (1.137)$$

te procijenimo model kojim se pretpostavlja postojanje linearnog trenda:

$$\hat{y}_t = \hat{\delta}_0 + \hat{\delta}_1 t \quad (1.138)$$

i procijenimo rezidualne kao razliku (1.137) i (1.138):

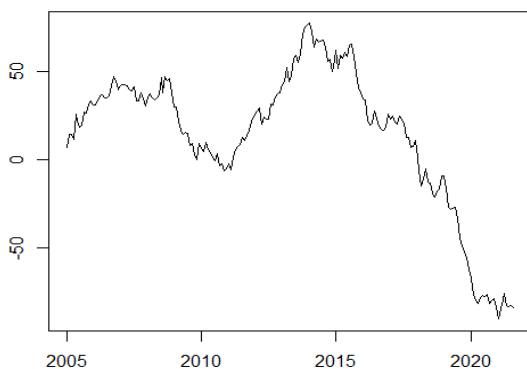
$$y_t - \hat{y}_t = \phi_0 + \sum_{\tau=1}^t \varepsilon_\tau - \hat{\delta}_0 - \hat{\delta}_1 t = (\phi_0 - \hat{\delta}_0) - \hat{\delta}_1 t + \sum_{\tau=1}^t \varepsilon_\tau. \quad (1.139)$$

Uočava se da stohastički trend i dalje nije uklonjen, jer je ostao izraz $\sum_{\tau=1}^t \varepsilon_\tau$.

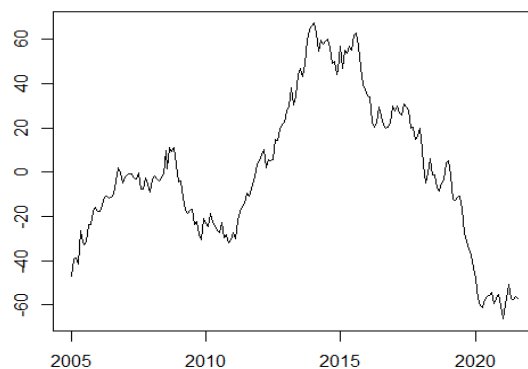
Primjer 1.25.

Slika 1.135. predočava usporedbu niza s lijevog panela slike 1.131. i rezidualne dobivene iz detrendiranja tog niza uz pretpostavku linearnog trenda. Uočava se da su reziduali također nestacionarni jer nije korektno uklonjen trend.

(a) Originalni nestacionarni niz



(b) Reziduali dobiveni detrendiranjem



Slika 1.135. Pogrešno uklanjanje trenda, niz je diferencijabilno stacionaran, trend uklanjan detrendiranjem

1.10.3.2. Diferenciranje determinističkog trenda

S druge strane, ako se radi o determinističkom trendu koji je prisutan u vremenskom nizu, te se pokušava ukloniti diferenciranjem, dolazi se do sljedećeg zaključka. Razmotrimo sljedeći niz, koji je definiran kao polinom stupnja J u varijabli t :

$$y_t = \sum_{j=0}^J \delta_j t^j + \varepsilon_t, \quad (1.140)$$

te potom pomaknemo u $t-1$:

$$y_{t-1} = \sum_{j=0}^J \delta_j (t-1)^j + \varepsilon_{t-1}, \quad (1.141)$$

kako bismo izračunali prvu diferenciju:

$$\begin{aligned} y_t - y_{t-1} &= \sum_{j=0}^J \delta_j t^j + \varepsilon_t - \sum_{j=0}^J \delta_j (t-1)^j - \varepsilon_{t-1} \\ &= \delta_0 + \delta_1 t + \dots + \delta_J t^J + \varepsilon_t - \delta_0 - \delta_1 (t-1) - \dots - \delta_J (t-1)^J - \varepsilon_{t-1} \quad , \\ &= \delta_1 (t - (t-1)) + \dots + \delta_J (t^J - (t-1)^J) + \varepsilon_t - \varepsilon_{t-1} \end{aligned} \quad (1.142)$$

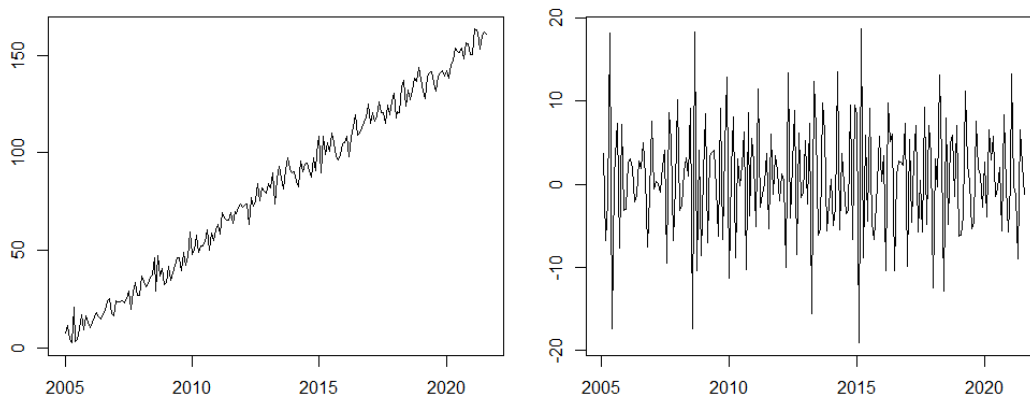
što općenito možemo pisati na sljedeći način:

$$y_t - y_{t-1} = \sum_{j=0}^J \delta_j (t^j - (t-1)^j) + \varepsilon_t - \varepsilon_{t-1}, \quad (1.143)$$

gdje se uočava da će prva diferencija biti definirana dijelom kao MA(1) proces, s obzirom na izraz $\varepsilon_t - \varepsilon_{t-1}$. Računanje drugih diferencija polučilo bi sličan ishod, pri čemu bi se radilo o MA(2) dijelu procesa vezan uz slučajnu varijablu. Stoga reziduali ne bi zadovoljavali pretpostavku čistog slučajnog procesa.

Primjer 1.26.

Slika 1.136. predočava niz koji sadrži linearni trend, sa slike 1.126. (lijevi panel) te njegovu prvu diferenciju (desni panel), koja izgleda kao stacionaran proces. No, procjenom modela (1.143) za analizirani niz, utvrđena je značajnost pomaka ε_{t-1} , što znači da prva diferencija u (1.143) slijedi MA(1) proces.



Slika 1.136. Pogrešno uklanjanje trenda, niz je trend-stacionaran, trend uklanjan diferenciranjem

1.10.4. Model slučajnog pomaka

Model slučajnog pomaka (engl. *random walk process*, RW) koristi se u području financija za testiranje efikasnosti tržišta kapitala. Najprije će se obraditi model slučajnog pomaka bez konstante, a potom s konstantom i nakon toga će se predočiti u odjeljku 1.10.4.3. hipoteza efikasnog tržišta u okviru koje se primjenjuje model slučajnog pomaka.

1.10.4.1. Model slučajnog pomaka bez konstante

Ako se razmatra proces koji je integriran reda 1, tj. $y_t \sim I(1)$ definiran izrazom:

$$y_t = y_{t-1} + \varepsilon_t, t \in \{\dots, -2, -1, 0, 1, 2, \dots\} \quad (1.144)$$

radi se o modelu slučajnog pomaka bez konstante. Uočimo da proces definiran u (1.144) podsjeća na AR(1) proces, no koeficijent uz pomak varijable y_t jednak je 1, što narušava svojstvo stacionarnosti AR(1) procesa.

Ako se razmotri prva diferencija ovakvog procesa:

$$\Delta y_t = y_t - y_{t-1} = \varepsilon_t, \quad (1.145)$$

uočava se da diferencija postaje čisti slučajni proces. Kako se proces u (1.144) može zapisati na sljedeći način:

$$\begin{aligned} y_t &= y_{t-1} + \varepsilon_t = y_{t-2} + \varepsilon_{t-1} + \varepsilon_t = y_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t \\ &= \dots = y_0 + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_{t-1} + \varepsilon_t, \\ &= y_0 + \sum_{j=1}^t \varepsilon_j \end{aligned} \quad (1.146)$$

očekivana razina pojave je konstantna jer vrijedi $E(y_t) = y_0$, no varijanca je promjenjiva jer je $Var(y_t) = t\sigma^2$, a slično se računa i kovarijanca među članovima procesa y_t i y_{t-k} , koja je jednaka $Cov(y_t, y_{t-k}) = t\sigma^2$, a pripadajuća korelacija jednaka je

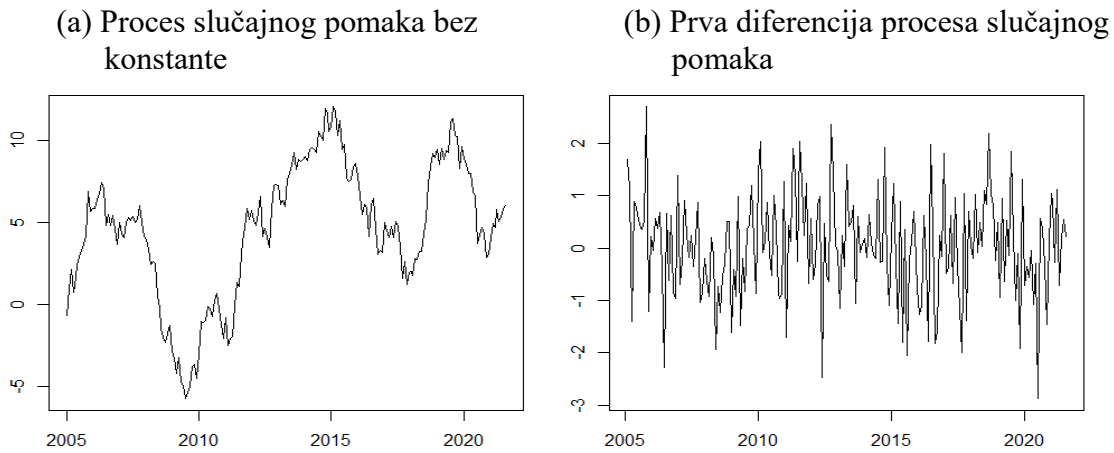
$$Corr(y_t, y_{t-k}) = \frac{t\sigma^2}{\sqrt{t\sigma^2(t+k)\sigma^2}} = \sqrt{\frac{t}{t+k}}. \quad (1.147)$$

Stoga će korelogram ACF funkcije upućivati na veoma sporo smanjivanje vrijednosti koeficijenata te funkcije.

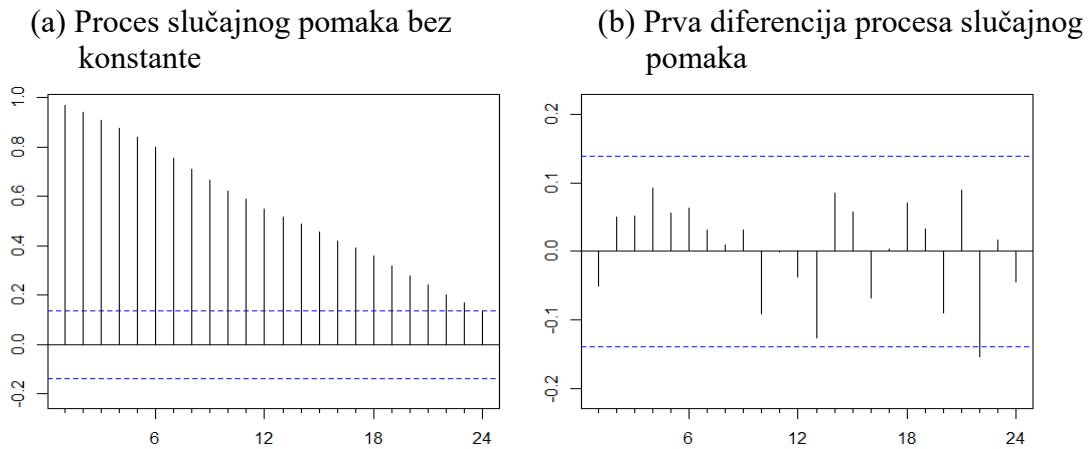
Primjer 1.27.

Slika 1.137. prikazuje proces slučajnog pomaka bez konstante (lijevi panel) i njegovu prvu diferenciju (desni panel). Uočava se kako diferenciranjem novi niz izgleda kao stacionaran proces. Slika 1.138. uspoređuje korelograme ACF funkcija za originalni niz i njegovu diferenciju. Dok se koeficijenti ACF funkcije za originalni niz smanjuju veoma sporo, korelogram procesa u prvim diferencijama upućuje na mogućnost slučajnog procesa.

Što se događa s funkcijom impulsnog odziva ako se razmatra ovakav proces? Šok u slučajnoj varijabli procesa $y_t = y_{t-1} + \varepsilon_t$ za t jednak je $\frac{dy_t}{d\varepsilon_t} = 1$, za $t+1$ i proces $y_{t+1} = y_t + \varepsilon_{t+1} = y_{t-1} + \varepsilon_t + \varepsilon_{t+1}$ je ponovno jednak 1, $t+2$ i proces $y_{t+2} = y_{t+1} + \varepsilon_{t+2} = y_t + \varepsilon_t + \varepsilon_{t+1} + \varepsilon_{t+2}$ je ponovno jednak 1, itd. **Dakle, šok ne nestaje, već ostaje trajan.**



Slika 1.137. Proces slučajnog pomaka bez konstante i prva diferencija procesa



Slika 1.138. Korelogram ACF funkcija slučajnog pomaka bez konstante i prve diferencije procesa

1.10.4.2. Model slučajnog pomaka s konstantom

Ako se razmatra proces koji je integriran reda 1, tj. $y_t \sim I(1)$ definiran izrazom:

$$y_t = \phi_0 + y_{t-1} + \varepsilon_t, t \in \{\dots-2, -1, 0, 1, 2, \dots\}, \quad (1.148)$$

radi se o modelu slučajnog pomaka s konstantom. Uočimo da proces definiran u (1.148) podsjeća na AR(1) proces, no koeficijent uz pomak varijable y_t jednak je 1, što narušava svojstvo stacionarnosti AR(1) procesa.

Ako se razmotri prva diferencija ovakvog procesa:

$$\Delta y_t = y_t - y_{t-1} = \phi_0 + \varepsilon_t, \quad (1.149)$$

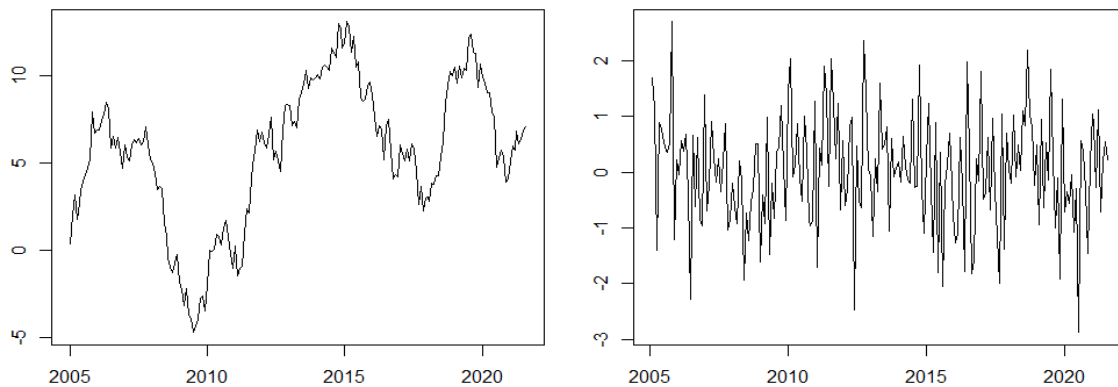
uočava se da diferencija postaje čisti slučajni proces uvećan za deterministički dio, konstantu ϕ_0 . Kako se proces u (1.149) može zapisati na sljedeći način:

$$\begin{aligned} y_t &= \phi_0 + y_{t-1} + \varepsilon_t = \phi_0 + \phi_0 + y_{t-2} + \varepsilon_{t-1} + \varepsilon_t \\ &= \phi_0 + \phi_0 + \phi_0 + y_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t \\ &= \dots = y_0 + t\phi_0 + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_{t-1} + \varepsilon_t, \quad (1.150) \\ &= y_0 + t\phi_0 + \sum_{j=1}^t \varepsilon_j \end{aligned}$$

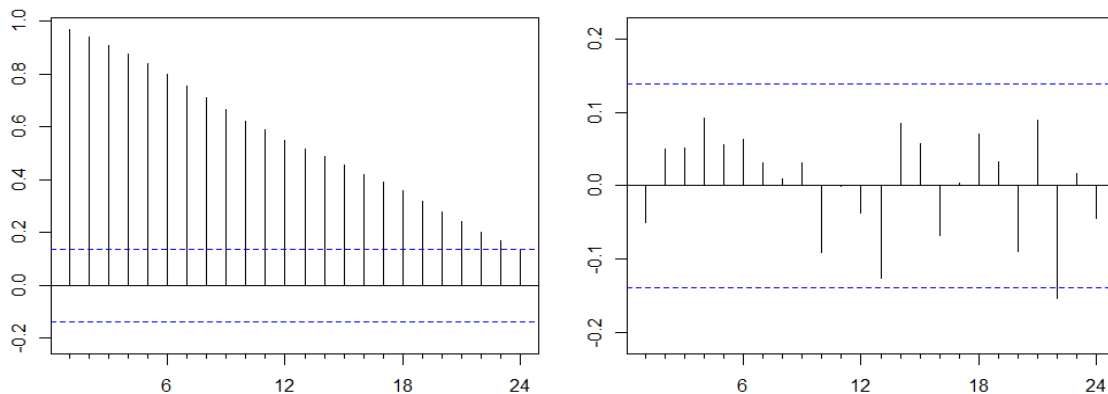
očekivana razina pojave je promjenjiva jer vrijedi $E(y_t) = y_0 + t\phi_0$, stoga je narušeno svojstvo konstantnog očekivanja stacionarnih procesa.

Primjer 1.28.

Slika 1.139. predočava proces slučajnog pomaka s konstantom (lijevi panel) i njegovu prvu diferenciju (desni panel), dok slika 1.140. prikazuje korelograme ACF funkcija za oba procesa. Zaključci su slični kao za prethodni proces u prethodnom odjeljku.



Slika 1.139. Proces slučajnog pomaka s konstantom i prva diferencija procesa



Slika 1.140. Korelogram ACF funkcija slučajnog pomaka s konstantom i prve diferencije procesa

1.10.4.3. Hipoteza efikasnog tržišta

Od druge polovice 20. stoljeća javlja se sve veći interes članova akademske zajednice i investitora za istraživanjem tržišta kapitala. Sve veći broj ekonomista počeo se zanimati za izučavanje tržišta kapitala, prinosa vrijednosnica i oblikovanja portfelja. Postavio se izazov je li moguće i ako da, kako modelirati i predviđati prinose dionica u svrhu „pobjeđivanja“ tržišta. Neke od najpoznatijih prekretnica učinili su u 50-im, 60-im i 70-im godinama prošlog stoljeća Eugene F. Fama i Maurice G. Kendall, formiranjem hipoteze efikasnog tržišta (engl. *Efficient Market Hypothesis*) i hipoteze slučajnog hoda (engl. *Random Walk Hypothesis*). Ukratko, pretpostavlja se kako su tržišta kapitala efikasna, što znači da su sve prošle informacije uključene u trenutnu cijenu dionica na tržištu, a nove se trenutno inkorporiraju u cijene. Stoga nije moguće ostvariti nadprosječne prinose trgujući na tim tržištima. Hipoteza efikasnog tržišta sugerira kako se sve važne informacije u potpunosti i trenutno uključuju u cijene dionica i zbog toga nije moguće ostvarivati nadprosječne prinose, već samo prosječne (vidjeti Fama 1970, 1991, 1995).

Hipoteza efikasnog tržišta (*Efficient Market Hypothesis*) može se sažeti na sljedeći način: na efikasnim tržištima sve važne informacije u potpunosti se i trenutno uključuju u cijene dionica i zbog toga nije moguće ostvarivati nadprosječne prinose, već samo prosječne. Kako je to sam Eugene Fama 1970. godine formulirao, sve prošle informacije već su odražene u cijenama, odnosno prinosisima dionica. Pri tom je efikasno tržište definirao kao „tržište gdje postoji velik broj racionalnih sudionika koji maksimiziraju profite, gdje svaki od njih pokušava predviđati buduće tržišne vrijednosti pojedinih vrijednosnica, i gdje su važne i trenutne informacije gotovo besplatno dostupne svim sudionicima. Na efikasnom tržištu, postojanje konkurencije među mnogim inteligentnim sudionicima vodi situaciji gdje, u bilo kojem vremenskom trenutku, stvarne cijene pojedinih vrijednosnica već odražavaju u sebi učinke informacija, temeljene na događajima koji su se već odvijali, kao i na onima za koje tržište očekuje da će se dogoditi u budućnosti. Drugim riječima, na efikasnom tržištu u svakom trenutku je stvarna cijena vrijednosnice dobra procjena njene intrinzične vrijednosti“ (Fama 1995:76).

Dakle, pretpostavlja se da je proces cijena (dionica) slučajan pomak bez konstante:

$$p_t = p_{t-1} + \varepsilon_t, \quad (1.151)$$

gdje je p_t vrijednost indeksa u danu/mjesecu t , p_{t-1} označava vrijednost indeksa u danu/mjesecu $t-1$, a ε_t slučajnu komponentu, odnosno element procesa inovacija. Modelom (1.151) definiran je proces slučajnog pomaka, s obzirom da je koeficijent uz p_{t-1} jednak 1. Ovakav proces potrebno je diferencirati jedanput da bi se postigla stacionarnost, što znači da je proces cijena integriran proces prvog reda, što kraće označavamo kao $p_t \sim I(1)$. Naime, ako se uvede operator pomaka L , vrijedi da je $Lp_t = p_{t-1}$. Zato se model (1.151) može zapisati na sljedeći način:

$$p_t = Lp_t + \varepsilon_t, \quad (1.152)$$

odnosno

$$(1-L)p_t = \varepsilon_t, \quad (1.153)$$

tj.

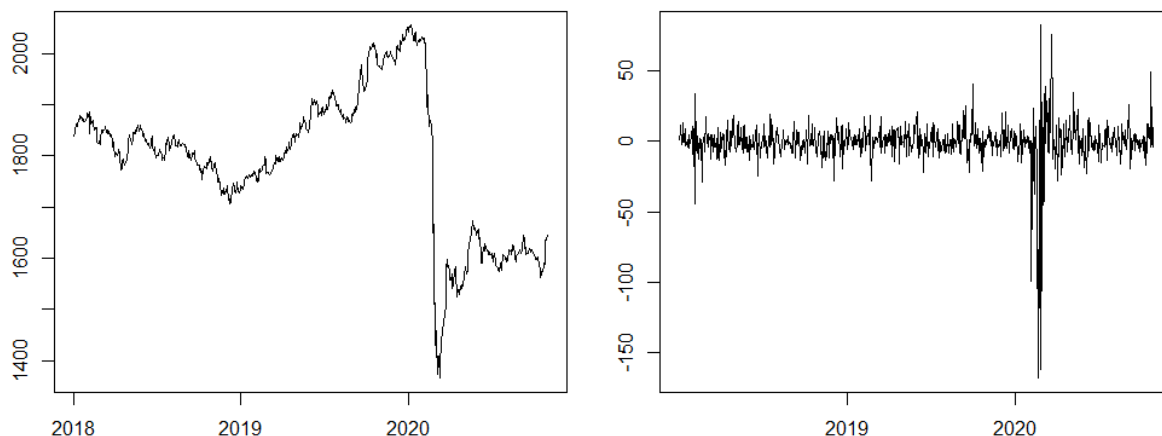
$$\Delta p_t = \varepsilon_t. \quad (1.154)$$

Posljednja jednakost upućuje da ako je proces cijena na tržištu proces slučajnog pomaka, tada će promjena cijena biti izvor inovacija. Na taj način nije moguće predviđati buduće promjene cijena, kao ni same cijene dionica na tržištu.

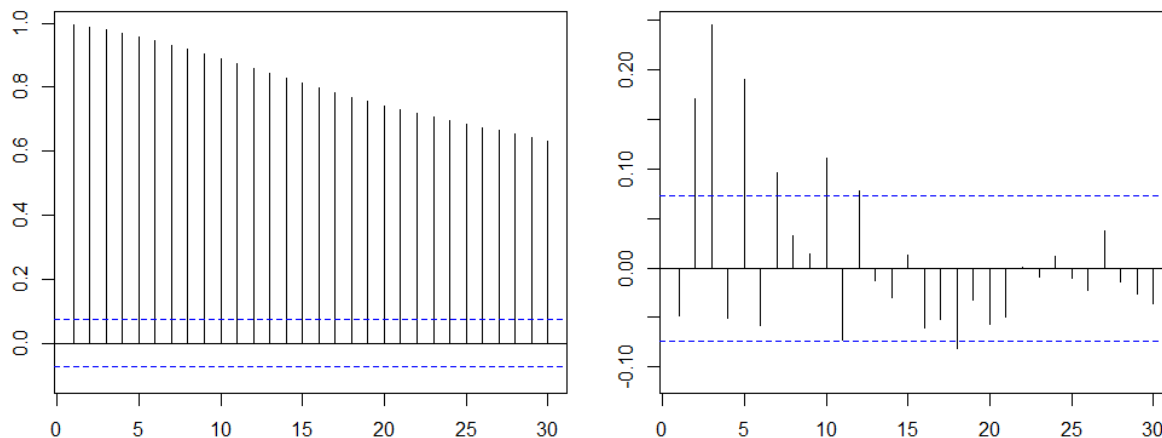
Primjer 1.29.

Sa stranica Investing.com (2020) preuzete su dnevne vrijednosti kretanja indeksa CROBEX Zagrebačke burze, za razdoblje 2. siječnja 2018. do 13. studenog 2020. godine. Predočimo grafički vremensku dinamiku indeksa CROBEX, te njegove prve diferencije, a potom i odgovarajućih korelograma ACF funkcija. Što zaključujemo?

Temeljem grafičkog prikaza na slici 1.141. uočava se da CROBEX nije stacionaran proces, već izgleda kao proces slučajnog pomaka (lijevi panel), dok njegova prva diferencija više podsjeća na stacionaran proces (desni panel). Utjecaj koronakrize uočljiv je na početku 2020. godine, kada i promjena, tj. prva diferencija CROBEX-a ima veću varijabilnost u odnosu na ostatak promatranog razdoblja. S druge strane, slika 1.142. upućuje da je CROBEX nestacionaran proces, jer se koeficijenti ACF funkcije smanjuju veoma sporo (lijevi panel); te iako se koeficijenti odgovarajuće ACF funkcije za prvu diferenciju (desni panel) smanjuju znatno brže, jedan dio koeficijenata je značajan (primjerice, pri pomacima 2, 3, 5, 7, itd.) što upućuje da se ne radi o čistom slučajnom procesu, tj. nije zadovoljeno svojstvo u (1.154). Stoga je u određenoj mjeri moguće predviđati promjenu cijena na Zagrebačkoj burzi.



Slika 1.141. Vremenska dinamika CROBEX-a i prva diferencija CROBEX-a

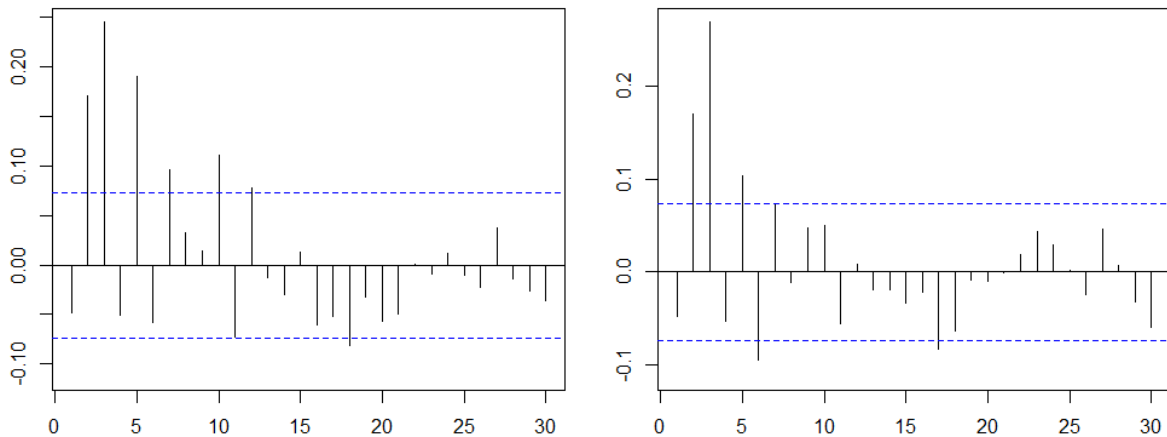


Slika 1.142. Korelogram ACF funkcija CROBEX-a i prve diferencije CROBEX-a

Dodatno su razmotreni korelogrami ACF i PACF funkcije za prvu diferenciju CROBEX-a, na slici 1.143., gdje se uočava da se radi o nekom ARMA(p,q) modelu. Točnije, radi se o sljedećem ARMA(3,2) modelu:

$$\Delta p_t = -0.44\Delta p_{t-1} + 0.38\Delta p_{t-2} + 0.35\Delta p_{t-3} + 0.38\varepsilon_{t-1} - 0.24\varepsilon_{t-2} + \varepsilon_t,$$

stoga bi se razmatrani model mogao koristiti za prognoziranje budućih vrijednosti promjene vrijednosti CROBEX-a (vidjeti odjeljak 1.9.5.).



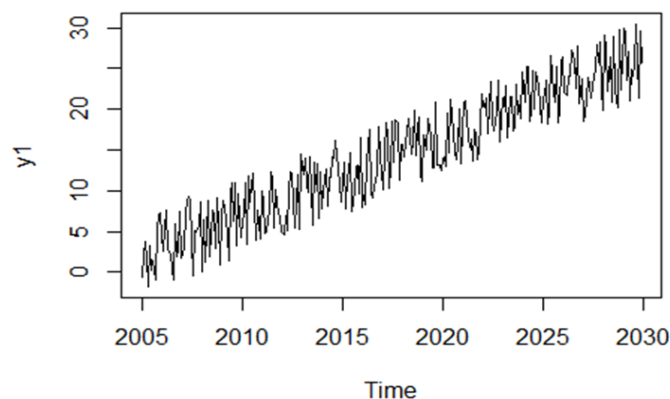
Slika 1.143. Korelogram ACF i PACF funkcija prve diferencije CROBEX-a

1.10.5. Primjer

Učitajmo datoteku „**nestac.txt**“ u RStudio. Datoteka sadrži podatke o kretanju triju nizova, y_1 , y_2 i y_3 . Definirajmo da se radi o vremenskim nizovima, s početnim datumom siječanj 2005., te da se radi o mjesečnim podacima i grafički predočimo sva tri niza i komentirajmo ih.

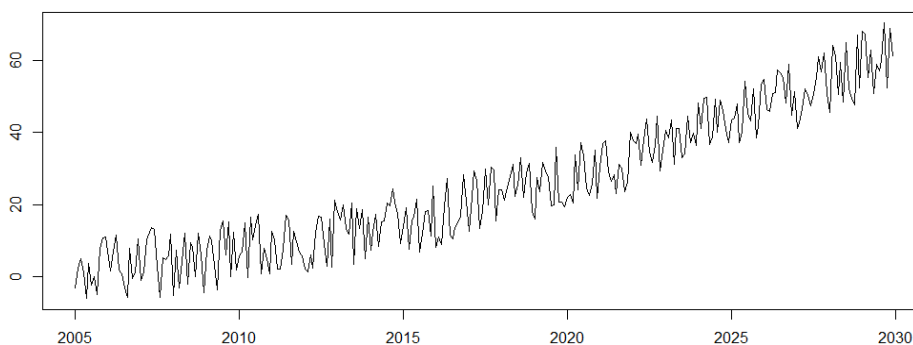
```
nestac<-read.table("nestac.txt",head=T,sep="\t")
y1<-ts(nestac$y1,start=c(2005,1),frequency = 12)
y2<-ts(nestac$y2,start=c(2005,1),frequency = 12)
y3<-ts(nestac$y3,start=c(2005,1),frequency = 12)
plot(y1)
```

Slika 1.144. Učitavanje datoteke u RStudio i grafičko predočavanje nizova

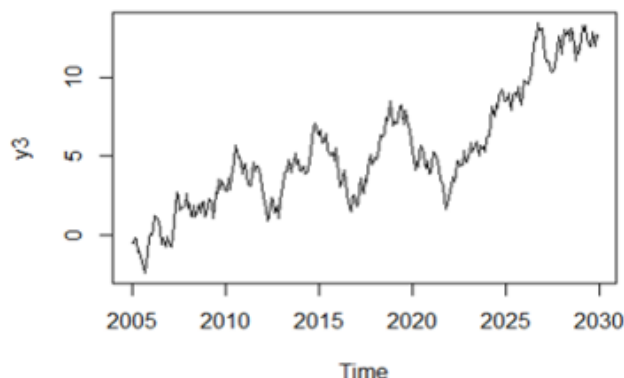


Slika 1.145. Grafički prikaz niza y_1

Uočava se da su sva tri niza nestacionarna, s obzirom da se očekivana razina pojave mijenja s vremenom. Prvi niz, y1 u sebi sadrži linearni trend, niz y2 zadrži kvadratni trend, dok posljednji, niz y3 izgleda kao proces slučajnog pomaka.



Slika 1.146. Grafički prikaz niza y2

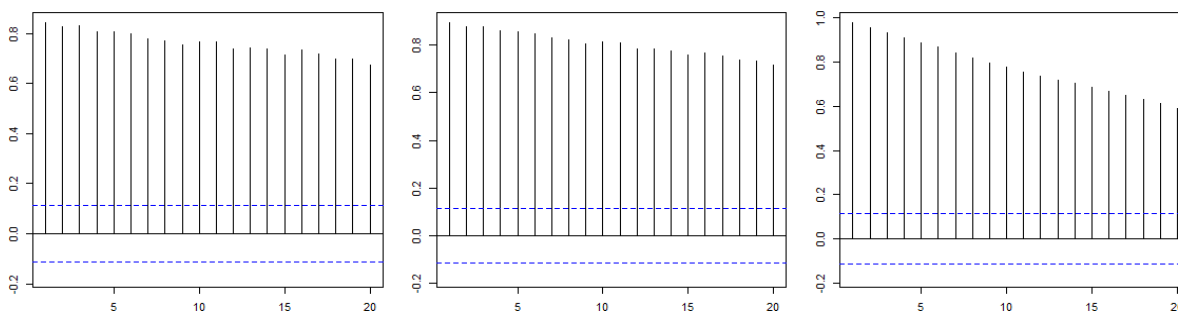


Slika 1.147. Grafički prikaz niza y3

Pređimo korelograme ACF funkcija za sva tri niza i komentirajmo ih. Na slici 1.148. pređene su naredbe za prikaz korelograma, dok slika 1.149. pređava redom korelograme: lijevi panel za y1, srednji panel za y2 i desni panel za niz y3. Koeficijenti svih ACF funkcija se smanjuju veoma sporo što upućuje na nestacionarne procese.

```
par(mfrow = c(1, 3),mar=c(2,2,1,2))
Acf(y1,20)
Acf(y2,20)
Acf(y3,20)
```

Slika 1.148. Naredbe za pređenje korelograma ACF funkcija sva tri niza



Slika 1.149. Korelogrami ACF funkcija sva tri niza

Provedimo postupak detrendiranja prva dva procesa na način da će se iz niza y_1 ukloniti linearni trend, iz niza y_2 kvadratni trend; dok za posljednji niz provedimo postupak diferenciranja, na način da se niz y_3 jedanput diferencira. Naredbe potrebne za provedbu ovih postupaka predočene su na slici 1.150. Uočimo da je najprije potrebno definirati varijablu trend naredbom `ts(..)` (engl. *time series*):

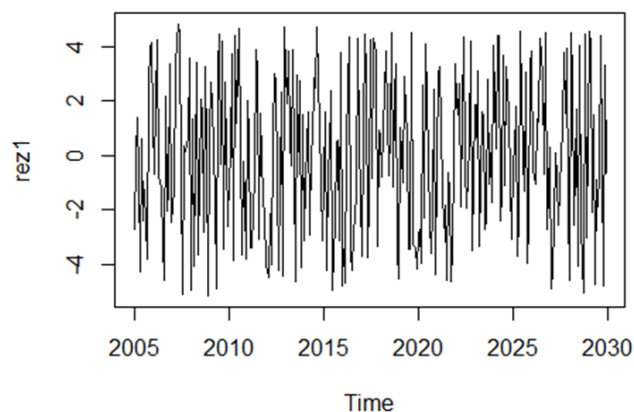
```
trend<-ts(1:length(y1),start=c(2005,1),frequency=12),
```

koja će biti duljine jednakoj duljini niza y_1 , i potom se definiraju reziduali iz modela gdje se niz y_1 regresira na trend, te potom niz y_2 na trend i njegov kvadrat. Konačno, naredbom `diff(...)` generira se prva diferencija niza y_3 . Potrebno je primijetiti da se naredbom diferenciranja izgubi jedan podatak, zbog pomaka za $t-1$, stoga se kod definiranja vremenskog niza za niz `dif` radi o početnom razdoblju veljača 2005. godine.

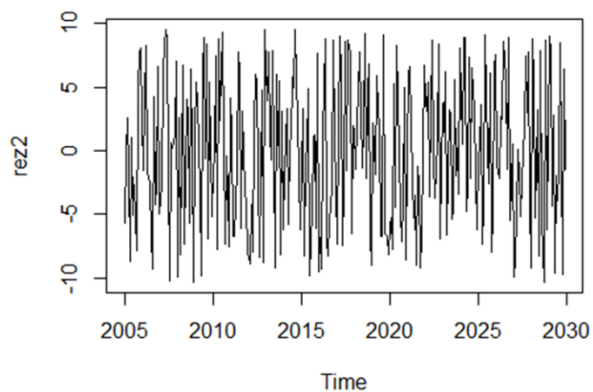
```
trend<-ts(1:length(y1),start=c(2005,1),frequency = 12)
rez1<-resid(lm(y1~trend))
rez2<-resid(lm(y2~trend+I(trend^2)))
dif<-diff(y3)
rez1<-ts(rez1,start=c(2005,1),frequency = 12)
rez2<-ts(rez2,start=c(2005,1),frequency = 12)
dif<-ts(dif,start=c(2005,2),frequency = 12)
plot(rez1)
```

Slika 1.150. Naredbe potrebne za detrendiranje i diferenciranje nizova

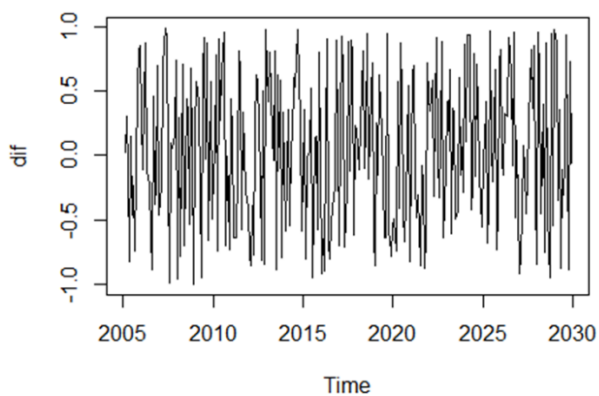
Konačno, predočimo detrendirane nizove y_1 i y_2 te diferencirani niz y_3 i potom korelograme odgovarajućih ACF funkcija za sva tri niza i komentirajmo ih. Slike 1.151., 1.152. i 1.153. redom predočavaju sva tri niza, gdje se uočava kako su postali stacionarni, što je dodatno potvrđeno na slici 1.154., gdje se uočava da su koeficijenti ACF funkcija unutar granica neznačajnosti, stoga upućuju na ponašanje za čisti slučajni proces.



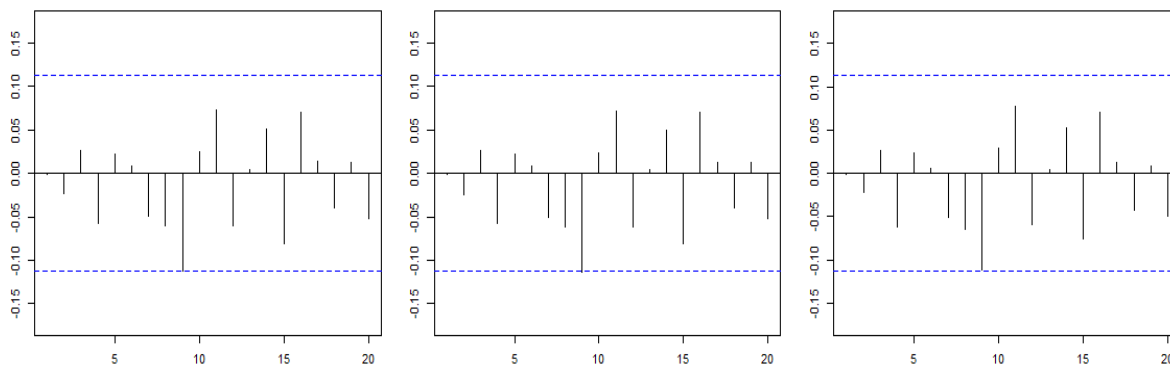
Slika 1.151. Grafički prikaz detrendiranog niza y_1



Slika 1.152. Grafički prikaz detrendiranog niza y2



Slika 1.153. Grafički prikaz diferenciranog niza y3



Slika 1.154. Korelogrami ACF funkcija za detrendirane i diferencirani niz

1.10.6. Napomena o problemu prividne regresije

Zašto je bitna stacionarnost u regresijskim modelima? Kako se ne bi pojavio problem prividne regresije (engl. *spurious regression*), na kojeg su prvi puta ukazali Granger i Newbold (1974). Naime, kada se razmatraju nestacionarni procesi u regresijskom modelu, on može ukazivati na lažnu povezanost između tih varijabli, iz razloga što imaju zajednički samo taj trend. A s druge strane, ta veza između varijabli zapravo ne postoji. Razmotrimo dva procesa koji slijede proces slučajnog pomaka, pri čemu je svaki generiran neovisno jedan o drugome:

$$y_t = y_{t-1} + e_t \quad (1.155)$$

i

$$x_t = x_{t-1} + u_t, \quad (1.156)$$

gdje su e_t i u_t čisti slučajni procesi, neovisni jedan o drugome. Ako se procijeni model u kojemu je proces y_t zavisna varijabla, a proces x_t nezavisna:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t, \quad (1.157)$$

ovaj model bi morao biti loš, u smislu da ne postoji učinak nezavisne na zavisnu varijablu (tj. varijabla x_t bi trebala biti neznajna). No, Granger i Newbold (1974) su proveli simulaciju u kojoj su pokazali da se čak 75% puta odbacuje nulta hipoteza o neznačajnosti varijable x_t u modelu (1.157) uz razinu značajnosti od 5%. Dodatno, regresijski modeli gdje se javlja problem lažne regresije također imaju velike koeficijente determinacije.

Dakle, ako postoji problem lažne regresije zbog prisutnosti trendova u varijablama, varijable će biti značajne u modelu iako u suštini nisu, bit će velika vrijednost R^2 , no prvi indikator da se radi o nekakvom problemu jest **vrijednost Durbin-Watsonove test veličine** za testiranje autokorelacije prvoga reda. Naime, kako je do sada više puta prikazano, nestacionarni nizovi imaju karakteristiku da ACF koeficijenti su veoma velike vrijednosti koje se veoma sporo smanjuju. Stoga će i autokorelacija reziduala takvih modela biti prisutna i značajna, što se odmah odražava u test veličini DW testa. Kako reziduali iz takvog modela nisu čisti slučajni proces, DW test upućuje na odbacivanje nulte hipoteze.

Kako bi prethodno navedeno bilo jasnije, razmotrimo slučajnu varijablu iz (1.157):

$$\varepsilon_t = y_t - \beta_0 - \beta_1 x_t, \quad (1.158)$$

te supstitucijom (1.155) i (1.156) u (1.158),

$$\varepsilon_t = y_{t-1} + e_t - \beta_0 - \beta_1 x_{t-1} - \beta_1 u_t, \quad (1.159)$$

$$\varepsilon_t = y_{t-2} + e_{t-1} + e_t - \beta_0 - \beta_1 x_{t-1} - \beta_1 u_{t-1} - \beta_1 u_t = \dots \quad (1.160)$$

$$\dots = y_0 - \beta_1 x_0 - \beta_0 + \sum_{i=1}^t e_i - \beta_1 \sum_{i=1}^t u_i, \quad (1.161)$$

gdje se uočava da \mathcal{E}_t nije slučajna varijabla, tj. čisti slučajni proces, nego proces čija se varijanca tijekom vremena povećava. S obzirom na promjenjivost varijance, više nisu pouzdani t -test, F -test, R^2 , intervalne procjene.

Primjer 1.30.

Generirana su dva niza prema modelima (1.155) i (1.156). Procijenimo regresijski model (1.157), provedimo t -test za varijablu x i komentirajmo ishod, te potom komentirajmo koeficijent determinacije i vrijednost Durbin-Watsonova testa.

Naredbe `cumsum(...)` koriste se za definiranje niza y i x , temeljem formule (1.146), pri čemu

$$\text{cumsum}(\text{rnorm}(200,0,1))$$

ima jedan argument: niz za koji se računa kumulativna suma, dok je naredba `rnorm(200,0,1)` ona kojom se definira niz čije su sljedeće karakteristike: prvi argument je broj opservacija (200), drugi argument je očekivana vrijednost niza (0) i posljednji argument je standardna devijacija niza (1). Naziv funkcije `rnorm` se odnosi na: r – *random* – slučajna varijabla, $norm$ – normalno distribuirana.

Na slici 1.155. predočeno je generiranje obje varijable te potrebni ispisi za rješavanje primjera. Uočava se da je p -vrijednost uz varijablu x manja od uobičajenih razina značajnosti, što upućuje na odbacivanje nulte hipoteze o neznačajnosti te varijable. Dodatno, koeficijent determinacije iznosi čak 20,72%, što bi se interpretiralo da je odabrani x opisao čak toliko varijacija varijable y , iako u suštini je y generiran neovisno o x . No, vrijednost DW test veličine iznosi 0.18, odnosno procijenjeni koeficijent autokorelacije reziduala prvog reda iznosi 0,91, što upućuje na jaku pozitivnu autokorelaciju reziduala prvog reda.

Dodatno, promotrimo kako izgledaju reziduali iz procijenjenog modela, razmotrimo njihov korelogram ACF funkcije i usporedimo s karakteristikama čistog slučajnog procesa. Slika 1.156., lijevi panel predočava vremensku dinamiku reziduala, gdje se uočava da se očekivana razina pojave mijenja, što narušava pretpostavku konstantnog očekivanja stacionarnog procesa, dok desni panel predočava korelogram ACF funkcije, na kojem se uočava sporo smanjivanje koeficijenata, što također upućuje na nestacionarnost procesa.

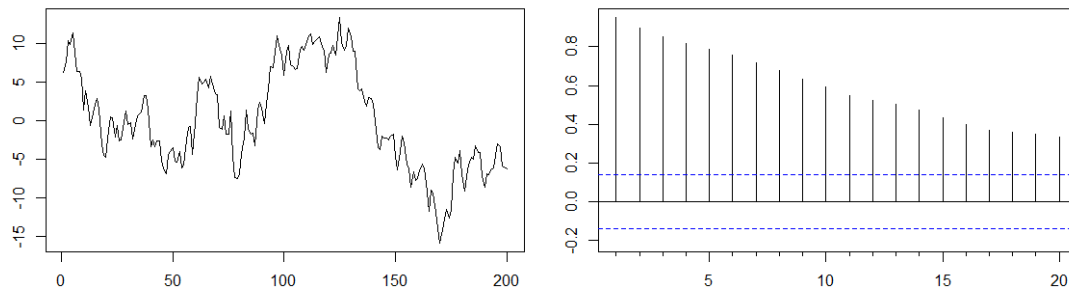
```
y<-cumsum(rnorm(200,0,1))
x<-cumsum(rnorm(200,0,1))
summary(lm(y~x))

## Call:
## lm(formula = y ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.5347 -1.4539  0.2548  1.8674  4.9780
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.17610     0.45048   4.831 2.72e-06 ***
## x           -0.50974     0.07002  -7.280 7.69e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.532 on 198 degrees of freedom
## Multiple R-squared:  0.2111, Adjusted R-squared:  0.2072
## F-statistic:   53 on 1 and 198 DF,  p-value: 7.686e-12
```

```
library(car)
durbinWatsonTest(lm(y~x))

## lag Autocorrelation D-W Statistic p-value
## 1 0.9106073 0.1765369 0
## Alternative hypothesis: rho != 0
```

Slika 1.155. Analiza problema prividne regresije



Slika 1.156. Grafički prikaz reziduala i korelograma ACF funkcije reziduala iz modela prividne regresije

1.10.7. Testovi jediničnog korijena

U prethodnim odjeljcima, kada se razmatrao korelogram ACF funkcija za nestacionarne procese, komentirano je kako se vrijednosti odgovarajućih koeficijenata smanjuju veoma sporo, za razliku od koeficijenata stacionarnih procesa koji se smanjuju brže. Kako se kod inverznog karakterističnog polinoma AR(1) procesa razmatra korijen, za kojeg smo pokazali da mora vrijediti $|\phi_1| < 1$ da bi proces bio stacionaran, upravo se u testovima jediničnog korijena razmatra vrijednost ϕ_1 u odnosu na jediničnu vrijednost, i od tuda slijedi **naziv test jediničnog korijena**. Kada bi se temeljem empirijskih podataka procijenio AR(1) proces, $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$, kako bi se procijenila vrijednost koeficijenta autokorelacije prvog reda za konstrukciju korelograma, $\text{Corr}(y_t, y_{t-1}) = \phi_1$, ta vrijednost će po apsolutnoj vrijednosti biti manja od 1, stoga se iz te informacije ne može raspoznati radi li se doista o stacionarnom procesu $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$, gdje je $|\phi_1| < 1$, ili pak o procesu slučajnog pomaka $y_t = y_{t-1} + \varepsilon_t$.

1.10.7.1. Dickey-Fullerov test i prošireni Dickey-Fullerov test

Dickey-Fullerov test

Stoga su Dickey i Fuller (1979) proveli Monte Carlo simulaciju, gdje su generirali mnogo procesa slučajnog pomaka, procijenili koeficijent ϕ_1 u modelu $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$ i razmatrali koliko je tih procijenjenih koeficijenata manje od 1. Rezultati su uputili da:

- 10% procijenjenih vrijednosti koeficijenta ϕ_1 je manje od 2.58 standardnih pogrešaka ispod jedinične vrijednosti;
- 5% procijenjenih vrijednosti koeficijenta ϕ_1 je manje od 2.89 standardnih pogrešaka ispod jedinične vrijednosti;

- 1% procijenjenih vrijednosti koeficijenta ϕ_1 je manje od 3.51 standardnih pogrešaka ispod jedinične vrijednosti.

Ove kritične granice su generirane uz pretpostavku da procesi slijede slučajan pomak, stoga se u **nultoj hipotezi Dickey-Fullerova testa pretpostavlja nestacionarnost procesa** koji se razmatra. Pa je zato Dickey-Fullerov test prvi formalan test za provedbu testiranja nestacionarnosti procesa.

Razmotrimo sljedeći model za proces y_t za koji se u nultoj hipotezi pretpostavlja da je nestacionaran:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t. \quad (1.162)$$

Zato je za postizanje stacionarnosti potrebno razmotriti prvu diferenciju takvog procesa:

$$y_t - y_{t-1} = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t - y_{t-1}, \quad (1.163)$$

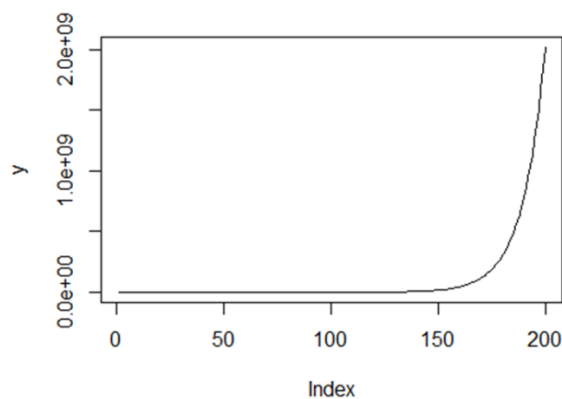
$$\Delta y_t = \phi_0 + (\phi_1 - 1)y_{t-1} + \varepsilon_t, \quad (1.164)$$

$$\Delta y_t = \phi_0 + \lambda y_{t-1} + \varepsilon_t. \quad (1.165)$$

Kako se u nultoj hipotezi pretpostavlja nestacionarnost procesa, tj. $\phi_1 = 1$, to znači da se u modelu (1.165) testira hipoteza da je $\lambda = 0$, jer je $\lambda = \phi_1 - 1$, dok je alternativna hipoteza ona koja pretpostavlja stacionaran proces:

$$\begin{aligned} H_0 : \lambda = 0 \quad (\phi_1 = 1) \\ H_1 : \lambda < 0 \quad (\phi_1 < 1) \end{aligned} \quad (1.166)$$

Uočimo da se ne razmatra slučaj $\lambda > 0$, tj. ($\phi_1 > 1$), jer bi se radilo o eksplozivnom procesu (vidjeti sliku 1.157., takvi procesi u ekonomiji ne postoje). Nadalje, jednačba definirana u (1.165) naziva se **DF (Dickey-Fullerova) jednačba**.



Slika 1.157. Proces $y_t = 1 + 1,1y_{t-1} + \varepsilon_t$

Stoga se procijeni model (1.165), formira se empirijski *DF* omjer:

$$\frac{\hat{\lambda}}{SE(\hat{\lambda})} \sim DF, \quad (1.167)$$

koji se uspoređuje sa spomenutim kritičnim granicama, tj. njihovim negativnim vrijednostima jer se radi o testu na donju granicu: $-2,58$, $-2,89$ ili $-3,51$, za model u kojemu je uključena konstanta. Ostale kritične granice (u slučaju isključivanja konstante, ili pak uključivanja i determinističkog trenda, te također ovise i o veličini niza koji se razmatra) dio su standardnog ispisa programske podrške za ekonometrijske analize. Dobiveni DF omjer se uspoređuje s kritičnim granicama i ako je veći od granice, ne odbacujemo nultu hipotezu, dok za vrijednost manju od kritične granice ne možemo odbaciti nultu hipotezu.

Dakle, uz model (1.165), gdje je uključena samo konstanta:

$$\Delta y_t = \phi_0 + \lambda y_{t-1} + \varepsilon_t, \quad (1.168)$$

možemo razmatrati i DF test bez konstante:

$$\Delta y_t = \lambda y_{t-1} + \varepsilon_t \quad (1.169)$$

i test s uključenom konstantom i linearnim trendom:

$$\Delta y_t = \phi_0 + \lambda y_{t-1} + \phi_2 t + \varepsilon_t. \quad (1.170)$$

Ako se utvrdi neodbacivanje nulte hipoteze, tj. da ne možemo odbaciti hipotezu da se radi o nestacionarnom nizu, kažemo da **proces sadrži jedinični korijen** (engl. *unit root*) i stoga se ovaj test često naziva **test jediničnog korijena**.

Napomena: uočimo da dobivena test veličina u (1.167) slijedi DF distribuciju i pogrešno bi bilo primijeniti standardni t -test, tj. uspoređivati dobivenu veličinu s kritičnim granicama Studentove distribucije jer standardna inferencijalna statistička analiza u takvom modelu nije valjana kada je zadovoljena pretpostavka o nestacionarnosti stohastičkog procesa koji se razmatra.

Prošireni Dickey-Fullerov test

U slučaju da se u modelu (1.168)-(1.170) javi problem da ε_t nije slučajan proces, već postoji autokorelacija među njegovim članovima, umjesto $AR(1)$ modela koji se razmatrao do sada, razmatra se $AR(p)$ model kako bi se uklonila spomenuta autokorelacija. U tom slučaju se radi o **proširenom Dickey-Fullerovom testu** (engl. *augmented DF test, ADF*):

$$\Delta y_t = \phi_0 + \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + \lambda y_{t-1} + \varepsilon_t, \quad (1.171)$$

$$\Delta y_t = \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + \lambda y_{t-1} + \varepsilon_t, \quad (1.172)$$

$$\Delta y_t = \phi_0 + \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + \lambda y_{t-1} + \phi_1 t + \varepsilon_t. \quad (1.173)$$

Kao i do sada, u proširenim testovima se testira ista nulta hipoteza i formira se ista test veličina u (1.167), koja se uspoređuje s kritičnim granicama DF distribucije.

Napomena: Uočimo da za originalni model, kada se razmotri AR(1) proces $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$, da DF jednadžbu $\Delta y_t = \phi_0 + \lambda y_{t-1} + \varepsilon_t$ možemo pisati kao AR(0) proces jer nema pomaka zavisne varijable Δy_t ; dok za AR(p) procese za koje se provode testovi jednadžbama (1.171)-(1.173), ADF jednadžbe se predočavaju kao AR($p-1$) procesi. Naime, jedan pomak se uvijek „izgubi“ zbog diferenciranja.

Odabir **DF ili ADF testa** u smislu uključivanja konstante ili isključivanja konstante ili pak uključivanja i konstante i trenda ovisi o iskustvu istraživača i grafičkom predočavanju vremenskog niza. Ako se na grafičkom prikazu uočava prisustvo konstante i linearnog trenda, uključuju se u sam test. Drugi pristup jest provedba od općeg pristupa prema specifičnom: na način da se najprije provede DF ili ADF test uz uključene i konstante i linearnog trenda, pa ako samo trend nije značajan u modelu, provesti (A)DF test samo uz konstantu uključenu u (A)DF jednadžbu, te ako i u tom slučaju nije značajna konstanta, na kraju provesti (A)DF test bez uključene konstante i trenda.

Nadalje, postavlja se još pitanje na koji način **odrediti adekvatan broj pomaka p** za AR(p) model, tj. $p-1$ za odgovarajuću ADF jednadžbu. S obzirom da se može javiti problem da se ne postigne slučajni proces za rezidualne nekog od modela ako se uključi premali pomak ili se pak s druge strane može odabrati prevelik pomak p pa se smanjuje snaga samog testa. Postoje dva pristupa. Prvi pristup je da se odabere velik pomak p te se u procijenjenom modelu analizira značajnost pomaka zavisne varijable i postepeno se smanjuje broj pomaka, u ovisnosti o ishodu testiranja značajnosti. Drugi pristup je korištenje informacijskih kriterija (vidjeti opširnije odjeljak 2.6.2.), koji su mjere adekvatnosti modela. Radi se o mjerama koje istovremeno uzimaju u obzir optimalnu vrijednost funkcije cilja dobivene u samoj optimizaciji za procjenu samog modela koji se razmatra, a istovremeno penalizira prevelik broj varijabli (ili njihovih pomaka) koje su nepotrebno uključene u model.

Spomenimo i da su Dickey i dr. (1986) pokazali da se DF test može primijeniti i u slučaju postojanja sezonskih komponenata u stohastičkom procesu, stoga se može razmotriti i sljedeća DF jednadžba:

$$\Delta y_t = \phi_0 + \sum_{s=1}^{S-1} \delta_s D_s + \lambda y_{t-1} + \phi_1 t + \varepsilon_t, \quad (1.174)$$

gdje su D_s binarne varijable definirane za slučaj mjesečnih ili tromjesečnih podataka, pri čemu je uključeno $S-1$ binarnih varijabli zbog uključene konstante.

Problemi s ADF testovima

Neki od problema vezani uz ADF testove su sljedeći:

- Spomenut je već odabir pomaka p u AR(p) modelu za inicijalni proces koji se razmatra, tj. hoće li se koristiti DF ili ADF test.
- Proces čija nestacionarnost se testira može biti definiran i MA(q) dijelom, pa se postavlja pitanje na koji način modelirati taj dio. Kako je u odjeljku 1.9.6. obrađeno,

svaki $MA(q)$ proces se može prikazati kao $AR(\infty)$ proces. Stoga se u DF jednadžbu uključi dovoljan broj pomaka zavisne varijable kako bi se aproksimirao $AR(\infty)$ proces.

- Također, već je spomenut problem kako odabrati adekvatnu specifikaciju testa u smislu uključivanja i trenda i konstante, samo konstante ili bez determinističkih dijelova uopće. Dva pristupa su ponuđena kao moguća rješenja.
- Ako se razmatra $AR(p)$ proces, koji ima ukupno p korijena karakteristične jednadžbe, postavlja se pitanje kako testirati više jediničnih korijena, jer (A)DF test testira samo jedan jedinični korijen. U praksi se proces diferencira onaj broj puta dok dođemo do odbacivanja nulte hipoteze (A)DF testa.

Vežano uz posljednji problem, u praksi se najprije u slučaju ne odbacivanja nulte hipoteze za niz u razinama, sam niz diferencira jedanput i potom se (A)DF test provodi nad prvim diferencijama testa i ako se u tom slučaju odbacuje nulta hipoteza testa, zaključuje se da je niz u prvim diferencijama stacionaran. Ako se pak nulta hipoteza opet ne odbacuje, niz se ponovno diferencira i nad drugim diferencijama se provodi (A)DF test, itd. Međutim, većina ekonomskih procesa su integrirani reda 1 ili 2. Veće vrijednosti mogu ukazivati na krivo primjenjivanje (A)DF testa, koji se ne odnosi na, primjerice, postojanje sezonskih jediničnih korijena.

Primjer 1.31.

Provedimo prošireni Dickey-Fullerov test za dva niza generirana primjeru 1.30. Pritom odaberimo tri specifikacije testa za svaki niz: najprije s uključenom konstantom i trendom, potom samo s uključenom konstantom, te potom bez konstante i bez trenda, kako slijedi:

$$\Delta y_t = \phi_0 + \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + \lambda y_{t-1} + \phi t + \varepsilon_t,$$

$$\Delta y_t = \phi_0 + \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + \lambda y_{t-1} + \varepsilon_t,$$

$$\Delta y_t = \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + \lambda y_{t-1} + \varepsilon_t.$$

U samome odabiru pomaka $p-1$, koristit ćemo Schwartzov informacijski kriterij (vidjeti odjeljak 1.9.8.). Zapišimo sve tri jednadžbe za oba procesa, test veličine i usporedimo s kritičnim granicama DF distribucije i donesimo zaključak o ishodu testova.

Slika 1.158. predočava naredbe za provođenje testa u okviru urca paketa. Naredba `ur.df(...)` (ur znači *unit root* te `df` – Dickey-Fuller) provodi upravo ADF test:

```
ur.df(x,type="trend",selectlags="BIC")
```

pri čemu kao prvi argument pišemo naziv niza za koji se provodi, za argument `type` navodimo „trend“ ako želimo uključiti i konstantu i trend kao determinističke regresore, „drift“ samo za konstantu i „none“ ako razmatramo model bez konstante i trenda. Posljednji argument `selectlags` se odnosi na kriterij temeljem kojeg se odabire optimalan broj pomaka p , pri čemu je odabran BIC, tj. Schwartzov kriterij.

```
library(urca)
summary(ur.df(x,type="trend",selectlags="BIC"))
summary(ur.df(x,type="drift",selectlags="BIC"))
summary(ur.df(x,type="none",selectlags="BIC"))
```

```
summary(ur.df(y,type="trend",selectlags="BIC"))
summary(ur.df(y,type="drift",selectlags="BIC"))
summary(ur.df(y,type="none",selectlags="BIC"))
```

Slike 1.158. Naredbe za provođenje ADF testa

Ako najprije razmotrimo niz x , te zapišemo prvu ADF jednadžbu nad kojom se test provodi, ona je sljedećeg oblika (slika 1.149.):

$$\Delta x_t = 0,68 - 0,20x_{t-1} + 0,006t + 0,11 \Delta x_{t-1} + \varepsilon_t,$$

s obzirom da u ispisu oznake treba pratiti kako slijedi:

- Ako piše z.lag.1 – to znači da je varijabla z prethodna (lagirana) za jedan pomak (z_{t-1})
- Ako piše z.diff.lag – to znači da je diferencija od z lagirana za jedan pomak (Δz_{t-1})
- Ako piše z.diff – to znači da promatramo diferenciju od z (Δz_t)
- Ako piše tt – to znači da je uključen linearni trend, t .

Dakle, temeljem BIC informacijskog kriterija bilo je potrebno uključiti jedan pomak zavisne varijable u model. Hipoteze testa su sljedeće:

$H_0: \lambda = 0$
 $H_1: \lambda < 0$, pri čemu je test veličina $DF = (-0,20)/0,05 = -4,93$, a kritične granice DF distribucije s kojima se uspoređuje empirijski DF omjer: $-3,99$, $-3,43$ i $-3,13$ za 10%, 5% i 1% (vidjeti zadnji dio ispisa na slici 1.159.). Kako je vrijednost $-4,93$ veća od svih triju vrijednosti, stoga se pri uobičajenim razinama značajnosti ne odbacuje hipoteza da je niz x nestacionaran.

Napomena: spomenuto je kako dobiveni omjer slijedi DF distribuciju, stoga se u ispisu ovoga testa ne razmatra posljednji stupac koji se odnosi na p -vrijednosti računane temeljem Studentove distribucije.

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.65595 -0.61460  0.03191  0.58374  3.09116
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.676577    0.173333   3.903 0.000131 ***
## z.lag.1      -0.201783    0.040909  -4.932 1.74e-06 ***
## tt           0.005588    0.001786   3.129 0.002025 **
## z.diff.lag   0.107765    0.070253   1.534 0.126670
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9464 on 194 degrees of freedom
## Multiple R-squared:  0.1147, Adjusted R-squared:  0.101
## F-statistic: 8.381 on 3 and 194 DF,  p-value: 2.883e-05
##
##
## Value of test-statistic is: -4.9324 8.5128 12.5478
##
## Critical values for test statistics:
##      1pct  5pct 10pct
```

```
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

Slika 1.159. Ispis ADF testa za niz x , uključeni konstanta i trend

Nadalje, ispisi ADF testova uz uključenu samo konstantu (gornji panel) i bez konstante i trenda (donji panel) predočeni su na slici 1.160., te su pripadajuće DF jednadžbe sljedeće:

$$\Delta x_t = 0,669 - 0,11x_{t-1} + 0,05 \Delta x_{t-1} + \varepsilon_t$$

i

$$\Delta x_t = -0,01x_{t-1} + 0,02 \Delta x_{t-1} + \varepsilon_t,$$

uz istu nultu hipotezu. No, sada su empirijski DF omjeri sljedeći: $-0,11 / 0,03 = -3,827$ i $0,05 / 0,07 = 0,785$, a odgovarajuće kritične granice za gornji panel: $-3,369$, $-2,88$ i $-2,57$ za 10%, 5% i 1% te donji: $-2,58$, $-1,95$ i $-1,63$ za 10%, 5% i 1%. Kako je za slučaj testa samo s konstantom empirijski DF omjer veći od kritičnih granica, a za slučaj testa bez konstante i trenda također empirijski DF omjer veći od kritičnih granica, za oba testa zaključujemo da se pri uobičajenim razinama značajnosti ne odbacuje hipoteza da je niz x nestacionaran.

Uočimo da je u prvome testu, gdje su bili uključeni i konstanta i trend, da oboje nisu značajni u modelu, stoga se nakon toga mogao odmah primijeniti test na donjem panelu slike 1.160. Čitatelju ostaje za vježbu provesti sve testove za niz y na slikama 1.160, 1.161. i 1.162.

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.9615 -0.6274  0.0243  0.5414  3.4159
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.66962    0.17718   3.779 0.000209 ***
## z.lag.1      -0.10559    0.02759  -3.827 0.000175 ***
## z.diff.lag   0.05473    0.06970   0.785 0.433282
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9675 on 195 degrees of freedom
## Multiple R-squared:  0.07005,    Adjusted R-squared:  0.06052
## F-statistic: 7.345 on 2 and 195 DF,  p-value: 0.0008406
##
##
## Value of test-statistic is: -3.8269 7.5344
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8499 -0.5282  0.0952  0.6787  3.3282
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## z.lag.1 -0.009498 0.011073 -0.858 0.392
## z.diff.lag 0.022307 0.071472 0.312 0.755
##
## Residual standard error: 0.9997 on 196 degrees of freedom
## Multiple R-squared: 0.004009, Adjusted R-squared: -0.006154
## F-statistic: 0.3945 on 2 and 196 DF, p-value: 0.6746
##
##
## Value of test-statistic is: -0.8578
##
## Critical values for test statistics:
##      1pct 5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.160. Ispis ADF testa za niz x , uključena konstanta (gornji panel); bez konstante i trenda (donji panel)

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2074 -0.7644 -0.0966  0.6268  2.7003
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)  1.002e+00  1.411e-01  7.098e+00 2.29e-11 ***
## z.lag.1      1.000e-01  2.630e-10  3.803e+08 < 2e-16 ***
## tt          -3.103e-04  1.363e-03 -2.280e-01  0.82
## z.diff.lag      NA           NA         NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9274 on 195 degrees of freedom
## Multiple R-squared: 1, Adjusted R-squared: 1
## F-statistic: 1.01e+17 on 2 and 195 DF, p-value: < 2.2e-16
##
##
## Value of test-statistic is: 380271892 91.5597 183.1128
##
## Critical values for test statistics:
##      1pct 5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2 6.22 4.75 4.07
## phi3 8.43 6.49 5.47
```

Slika 1.161. Ispis ADF testa za niz y , uključeni konstanta i trend

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.19751 -0.76893 -0.09193  0.61871  2.67970
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)  9.736e-01  6.954e-02  14 <2e-16 ***
## z.lag.1      1.000e-01  2.219e-10 450635062 <2e-16 ***
## z.diff.lag      NA           NA         NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9251 on 196 degrees of freedom
```

```
## Multiple R-squared:      1, Adjusted R-squared:      1
## F-statistic: 2.031e+17 on 1 and 196 DF, p-value: < 2.2e-16
##
## Value of test-statistic is: 450635062 183.9576
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5973  0.1248  0.7742  1.5287  3.5986
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error  t value Pr(>|t|)
## z.lag.1    1.00e-01  2.96e-10 337862204 <2e-16 ***
## z.diff.lag      NA           NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.305 on 197 degrees of freedom
## Multiple R-squared:      1, Adjusted R-squared:      1
## F-statistic: 1.142e+17 on 1 and 197 DF, p-value: < 2.2e-16
##
## Value of test-statistic is: 337862204
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.162. Ispis ADF testa za niz y , uključena konstanta (gornji panel); bez konstante i trenda (donji panel)

Primjer 1.32.

S obzirom na zaključke u prethodnom primjeru, provedimo ADF test nad oba niza u prvim diferencijama.

Prethodno smo razmatrali oba niza u razinama, jer nismo vršili nikakvu transformaciju nad njima. Sada ćemo najprije svaki niz diferencirati (vidjeti naredbu na slici 1.163.) pomoću naredbe `diff`, a potom nad diferenciranim nizovima provesti ADF test, bez uključene konstante i trenda, jer su oboje bili neznačajni kod testova varijabli u razinama, pa diferenciranjem nema smisla njih uključivati u jednadžbu.

```
dx<-diff(x)
dy<-diff(y)
summary(ur.df(dx,type="none",selectlags="BIC"))
summary(ur.df(dy,type="none",selectlags="BIC"))
```

Slika 1.163. Diferenciranje nizova x i y te ADF testovi nad diferencijama

Sada su sljedeće ADF jednadžbe (ispis na slici 1.164., x gornji panel, y donji panel):

$$\Delta^2 x_t = -1,04\Delta x_{t-1} + 0,06\Delta^2 x_{t-1} + \varepsilon_t \text{ i } \Delta^2 y_t = -1,15\Delta y_{t-1} + 0,05\Delta^2 y_{t-1} + \varepsilon_t.$$

Obratimo pažnju da s obzirom na diferencije originalnih nizova, ADF test se provodi nad diferencijama pa se diferenciranjem prvih diferencija dobiva druga diferencija.

Naime, za originalni niz i test smo pisali DF jednadžbu $\Delta y_t = \lambda y_{t-1} + \varepsilon_t$, jer smo ju dobili u postupcima (1.163)-(1.165) (za slučaj s konstantom, slično je bez konstante), a sada nad inicijalnim AR(1) procesom $\Delta y_t = \phi \Delta y_{t-1} + \varepsilon_t$ provodimo test pa je potrebno od lijeve i desne strane jednakosti oduzeti Δy_{t-1} pa se dobije: $\Delta y_t - \Delta y_{t-1} = \phi \Delta y_{t-1} - \Delta y_{t-1} + \varepsilon_t$, tj. $\Delta^2 y_t = (\phi - 1) \Delta y_{t-1} + \varepsilon_t$.

```
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8009 -0.6274  0.0211  0.6133  3.2388
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -1.04118    0.09972  -10.441  <2e-16 ***
## z.diff.lag   0.06449    0.07103   0.908    0.365
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9978 on 195 degrees of freedom
## Multiple R-squared:  0.4931, Adjusted R-squared:  0.4879
## F-statistic: 94.83 on 2 and 195 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic is: -10.4412
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

```
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.45363 -0.78393 -0.06921  0.70155  3.08627
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -1.15088    0.106   -10.83  < 2e-16 ***
## z.diff.lag   0.04962    0.0059   8.299  1.71e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.105 on 195 degrees of freedom
## Multiple R-squared:  1, Adjusted R-squared:  1
## F-statistic: 6.575e+14 on 2 and 195 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic is: -10.830
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.164. Ispis ADF testa za niz x (gornji panel) i y (donji panel), bez konstante i trenda

Odgovarajuće DF test veličine iznose $-10,44$ i $-10,83$, što je manje od kritičnih granica na dnu ispisa svakog testa, stoga se odbacuje nulta hipoteza o nestacionarnosti pojedinog niza.

Dakle, svaki niz smo jedanput diferencirali da bi postao stacionaran, stoga je svaki niz integriran prvog reda, tj. $x_t \sim I(1)$ i $y_t \sim I(1)$.

Primjer 1.33.

Iz sveobuhvatnog primjera 1.10.5., gdje smo učitali datoteku „**nestac.txt**“ u RStudio, s tri niza y_1 , y_2 i y_3 , y_1 smo razmatrali s uključenim determinističkim linearnim trendom. Provedimo ADF test nad tim nizom na način da u ADF jednadžbu uključimo trend i konstantu, a potom isključimo determinističke komponente te usporedimo ishode testova i zaključimo o kakvom se nizu radi.

Slika 1.165. predočava potrebne naredbe za provedbu oba testa, dok ispisi na slici 1.166. prikazuju ADF jednadžbe oba testa (uključena konstanta i trend – gornji panel, bez determinističkih regresora – donji panel). ADF jednadžbe za oba testa su:

$$\Delta y_{1t} = 2,27 - 1,02y_{1t-1} + 0,08t + 0,02 \Delta y_{1t-1} + \varepsilon_t \quad \Delta y_{1t} = -0,01y_{1t-1} - 0,02 \Delta y_{1t-1} + \varepsilon_t$$

Test veličine su redom: $-12,42$ i $-0,77$. Kako je test veličina u slučaju testa s konstantom i trendom ($-12,42$) manja od triju kritičnih granica ($-3,98, -3,42, -3,13$), odbacuje se nulta hipoteza o nestacionarnosti procesa. To znači da je prisutan linearni trend u razmatranoj seriji, oko kojeg se kreću vrijednosti same serije y_1 . S druge strane, test veličina u slučaju testa bez determinističkih regresora ($-0,77$) je veća od triju kritičnih granica ($-2,58, -1,95, -1,62$), u ovome testu se ne odbacuje nulta hipoteza o nestacionarnosti procesa. Izbacivanjem linearnog trenda iz testne jednadžbe dolazimo do drugačijeg zaključka, da se radi o nestacionarnom procesu. Kako je niz stacionaran uz uključivanje trenda u ADF jednadžbu, a nestacionaran kada se trend izbaci, zaključujemo u takvom slučaju **da se ovdje radi o trend-stacionarnom procesu.**

```
summary(ur.df(y1,type="trend",selectlags="BIC"))
summary(ur.df(y1,type="none",selectlags="BIC"))
```

Slika 1.165. Naredbe za ADF test, uključeni trend i konstanta te bez konstante i trenda

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.1056 -2.5411 -0.0808  2.7330  4.8772
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.268763   0.384441   5.901 9.91e-09 ***
## z.lag.1      -1.024805   0.082496 -12.422 < 2e-16 ***
## tt           0.082545   0.006936  11.901 < 2e-16 ***
## z.diff.lag   0.023907   0.058352   0.410  0.682
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -12.4224 51.6259 77.1581
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.98 -3.42 -3.13
```

```
## phi2 6.15 4.71 4.05
## phi3 8.34 6.30 5.36
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.3099 -2.5599  0.8519  2.7192  9.3164
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -0.01014    0.01316  -0.770   0.442
## z.diff.lag  -0.48219    0.05131  -9.398 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -0.7704
##
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.166. Ispis ADF testa za niz y_1 , uključeni trend i konstanta (gornji panel) i bez konstante i trenda (donji panel)

Uklonimo sada trend iz niza y_1 detrendiranjem, te potom nad rezidualima provedimo ADF test, bez konstante i trenda. Ako se prisjetimo definicije i karakteristika čistog slučajnog procesa (odjeljak 1.9.2.), reziduali dobiveni postupkom detrendiranja bi trebali zadovoljavati upravo tu definiciju, jer je odabran ispravan način izbacivanja nestacionarnosti (vidjeti odjeljak 1.10.3.). Provedbom ADF testa bi trebali dobiti zaključak kako se radi o stacionarnom procesu. Kako je očekivana vrijednost slučajne varijable jednaka 0, kada nad njome, kao i nad rezidualima, provodimo ADF test, uvijek razmatramo **test bez determinističkih komponenti**.

ADF jednadžba nad rezidualima je sljedeća (slika 1.167.):

$\Delta \varepsilon_t = -1,02\varepsilon_{t-1} + 0,02\Delta \varepsilon_{t-1} + u_t$, uz DF test veličinu $-12,464$ što je manje od kritičnih granica (u tablici 1.2., vidjeti napomene u odjeljku 1.14.), stoga se odbacuje nulta hipoteza da su reziduali nestacionarni.

```
summary(ur.df(resid(lm((y1~trend))),type="none",selectlags="BIC"))

## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.1217 -2.5501 -0.0828  2.7466  4.9085
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -1.02475    0.08222 -12.464 <2e-16 ***
## z.diff.lag   0.02389    0.05815  0.411  0.682
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Value of test-statistic is: -12.464
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.167. Ispis ADF testa za niz reziduala

Primjer 1.34.

Generirajmo dva čista slučajna procesa, koji će svaki imati ukupno 200 opservacija. Prvi proces ima očekivanje 0 i standardnu devijaciju 1, dok drugi ima očekivanje 5 i standardnu devijaciju također 1. Provedimo ADF test za oba procesa. Koji deterministički član ćemo uključiti u ADF jednadžbe i zašto?

Za prvi proces, čije je očekivanje 0, ne uključujemo niti trend niti konstantu, dok za drugi proces, čije je očekivanje 5 uključujemo samo konstantu (slika 1.168.).

```
epsilon<-rnorm(200,0,1)
summary(ur.df(epsilon,type="none",selectlags="BIC"))
epsilon2<-rnorm(200,5,1)
summary(ur.df(epsilon2,type="drift",selectlags="BIC"))
```

Slika 1.168. Generiranje nizova i naredbe za provedbu ADF testova

```
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2552 -0.7553 -0.1481  0.6124  2.6545
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -0.91123    0.09925  -9.181  <2e-16 ***
## z.diff.lag  -0.05897    0.07102  -0.830   0.407
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -9.1814
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62

## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8353 -0.6715 -0.0273  0.5740  3.1952
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.32901    0.50699  10.511  <2e-16 ***
## z.lag.1     -1.05637    0.09965 -10.600  <2e-16 ***
## z.diff.lag   0.07237    0.07107   1.018   0.31
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -10.6004 56.1901
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

Slika 1.169. Ispis ADF testa za niz epsilon (gornji panel) i epsilon2 (donji panel)

Za niz epsilon provedba je testa kako slijedi. ADF jednadžba je:

$\Delta\varepsilon_t = -0,91\varepsilon_{t-1} - 0,06\Delta\varepsilon_{t-1} + u_t$, *DF* test veličina je $-0,91/0,099 = -9,181$, te je vrijednost manja od kritičnih granica $(-2,58, -1,95, -1,62)$, pa se odbacuje nulta hipoteza o nestacionarnosti procesa na uobičajenim razinama značajnosti.

Za niz epsilon2 se također odbacuje nulta hipoteza o nestacionarnosti procesa na uobičajenim razinama značajnosti.

Primjer 1.35.

U sveobuhvatnom primjeru 1.10.5. analizirali smo datoteku „gen_nizovi.txt“, koja je sadržavala AR i MA nizove. Provedimo ADF test za niz y2 te za niz y4, uz uključenu konstantu.

```
y2<-ts(nizovi$y2,start=c(2003,9),frequency=12)
summary(ur.df(y2,type="drift",selectlags="BIC"))
y4<-ts(nizovi$y4,start=c(2003,9),frequency=12)
summary(ur.df(y4,type="drift",selectlags="BIC"))
```

Slika 1.170. Naredbe za provedbu ADF testova za nizove y2 i y4

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.04286 -0.49022 -0.00813  0.52685  1.00958
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -10.37264    1.03287  -10.042  <2e-16 ***
## z.lag.1      -1.03947    0.10345  -10.048  <2e-16 ***
## z.diff.lag   -0.00403    0.07188   -0.056    0.955
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -10.0477 50.4805
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.96311 -0.46007  0.02978  0.42309  0.99787
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0003547  0.0403394  -0.009  0.9930
## z.lag.1     -0.0287624  0.0150640  -1.909  0.0577 .
## z.diff.lag   0.0013526  0.0708127   0.019  0.9848
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -1.9093 1.8685
##
## Critical values for test statistics:
##      1pct  5pct 10pct
```

```
## tau2 -3.46 -2.88 -2.57
## phi1 6.52 4.63 3.81
```

Slika 1.171. Ispis ADF testa za niz y_2 (gornji panel) i y_4 (donji panel)

Uočava se da su obje empirijske DF veličine manje od odgovarajućih kritičnih granica, stoga se odbacuje nulta hipoteza o nestacionarnosti pojedinog procesa. Naime, kako su i y_2 i y_4 stacionarni procesi, jer se radi o AR(1) i MA(1) procesima, ishodi ADF testova na to su trebali ukazati.

Primjer 1.36.

Paritet kupovne moći (PPP, engl. *purchasing power parity*) upućuje na vezu između domaće razine cijena i tečaja. Jednostavan oblik testiranja PPP-a odnosi se na razmatranje razlike između domaće i strane stope inflacije. Neka su p_t i p_t^f logaritmirane vrijednosti domaće razine cijena i strane razine cijena (engl. *foreign, f*), dok je e logaritam tečaja domaće valute i strane valute, PPP implicira da vrijedi:

$$e_t = p_t - p_t^f + d_t,$$

gdje d_t predstavlja odstupanje od PPP-a. Ekonometrijsko testiranje PPP-a se svodi na izračun: $r_t = e_t - p_t + p_t^f$, i testiranje stacionarnosti procesa r_t . Naime, ako vrijedi PPP između dvije zemlje, tečaj korigiran za domaće i strane cijene treba biti stacionaran.

Učitajmo datoteku „**ppp.txt**“ u RStudio. Datoteka sadrži mjesečne podatke⁸ o kretanju HICP-a Hrvatske (engl. *harmonized index of consumer prices*) i Austrije, te tečaj kune naspram eura. Podaci kreću s početnim datumom siječanj 1998. godine. Najprije desezonirajmo samu vrijednost tečaja pomoću metode pomičnih prosjeka, te potom logaritmirajmo sva tri niza i za niz r_t skicirajmo najprije sam niz i komentirajmo ga, predložimo korelograme ACF i PACF funkcija i na kraju provedimo samo testiranje nestacionarnosti ADF testom, pri čemu uključimo samo konstantu kao deterministički regresor (vidjeti Enders, 2004 o detaljima testa i zašto ne uključiti trend zbog PPP teorije.).

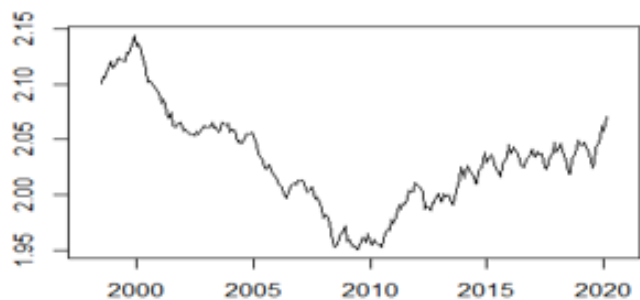
Slika 1.172. predočava učitavanje podataka u RStudio, definiranje varijabli i grafičko predočavanje niza r_t . Uočava se da sam niz nije stacionaran, s obzirom na njegovu vremensku dinamiku na slici 1.173., što se potvrđuje korelogramima na slici 1.174., gdje se uočava veoma sporo smanjivanje koeficijenata i ACF i PACF funkcije.

```
ppp<-read.table("ppp.txt",header=T,sep="\t")
tecaj<-ts(ppp$tecaj, start=c(1998,1),frequency = 12)
library(forecast)
tecaj<-ma(tecaj,12,centre = T)

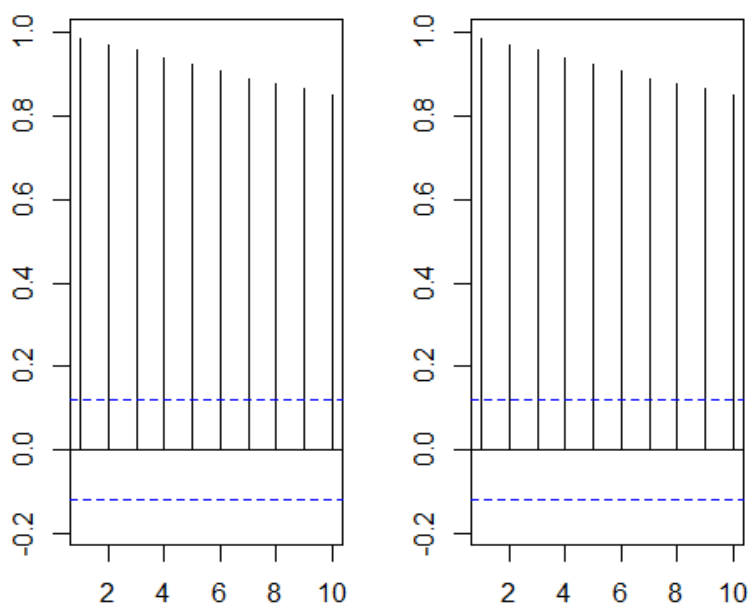
r<-log(tecaj)+log(ppp$aut)-log(ppp$cro)
plot(r,ylab=NA,xlab=NA)
```

Slika 1.172. Učitavanje podataka i definiranje varijable r_t

⁸ Podaci prikupljeni s Eurostata (2020).



Slika 1.173. Vremenska dinamika r_t



Slika 1.174. Korelogrami ACF i PACF funkcija za niz r_t

Nadalje, slika 1.175. predočava ispis ADF jednadžbe, temeljem koje se provodi ADF test kako slijedi. Test veličina $(-1,543)$ veća je od kritičnih granica, stoga se ne odbacuje nulta hipoteza o nestacionarnosti procesa koji se razmatra. Dakle, za dvije razmatrane zemlje ne vrijedi PPP.

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.02284    0.01489   1.534   0.126
## z.lag.1     -0.01131    0.00733  -1.543   0.124
## z.diff.lag   0.08533    0.06205   1.375   0.170
##
## Value of test-statistic is: -1.5433 1.2671
##
## Critical values for test statistics:
##           1pct  5pct 10pct
## tau2  -3.44 -2.87 -2.57
## phi1   6.47  4.61  3.79
```

Slika 1.175. ADF test za niz r_t

Primjer 1.37.

U odjeljku 1.10.4.3. obrađena je hipoteza efikasnog tržišta. Pretpostavilo se da je proces cijena (dionica) slučajan pomak bez konstante:

$$p_t = p_{t-1} + \varepsilon_t,$$

gdje je p_t vrijednost indeksa u danu/mjesecu t , p_{t-1} označava vrijednost indeksa u danu/mjesecu $t-1$, a ε_t slučajnu komponentu, odnosno element procesa inovacija. Zaključili smo da je cijena integriran proces prvog reda prema spomenutoj hipotezi, tj. $p_t \sim I(1)$, dok smo za promjenu cijene dobili rezultat:

$$\Delta p_t = \varepsilon_t,$$

što smo interpretirali da su promjene cijena čisti slučajni proces.

Pomoću ADF testa nad istim podacima iz spomenutog odjeljka, tj. nad podacima za tržišni indeks CROBEX, provedimo najprije sam ADF test u razinama s uključenom konstantom, te potom nad prvim diferencijama provedimo navedeni test bez uključenih determinističkih komponenti. Temeljem ishoda testova zaključimo jesu li cijene nestacionaran proces, a promjena cijena stacionaran.

Naredbe za provedbu ADF testova za oba niza predočene su na slici 1.176., dok su same ADF jednadžbe prikazane na slici 1.177. Gornji panel upućuje da je niz CROBEX nestacionaran u razinama, dok za prve diferencije samog niza odbacujemo hipotezu o nestacionarnosti. Dakle, iako je promjena cijena na Zagrebačkoj burzi stacionaran proces, vidjeli smo u odjeljku 1.10.4.3. da korelogrami upućuju na postojanje autokorelacije među članovima procesa promjena cijena, stoga je u nekoj mjeri moguće predviđanje budućih kretanja promjene cijena na hrvatskom tržištu kapitala.

```
cro<-read.table("crobox_dnevno.txt",sep="\t",head=T)
cro<-ts(cro$Price,start=c(2018,1,2),frequency = 252)
summary(ur.df(cro,type="drift",selectlags="BIC"))
summary(ur.df(diff(cro),type="none",selectlags="BIC"))
```

Slika 1.176. Naredbe za provedbu ADF testova

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Coefficients:
##          Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.354926   7.272842   1.149   0.251
## z.lag.1     -0.004831   0.004050  -1.193   0.233
## z.diff.lag  -0.045838   0.037543  -1.221   0.223
##
## Value of test-statistic is: -1.1928 0.8444
##
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau2  -3.43 -2.86 -2.57
## phi1   6.43  4.59  3.78
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Coefficients:
##          Estimate Std. Error t value Pr(>|t|)
```

```
## z.lag.1 -0.87031 0.05359 -16.240 < 2e-16 ***
## z.diff.lag -0.16976 0.03702 -4.585 5.35e-06 ***
##
## Value of test-statistic is: -16.2404
##
## Critical values for test statistics:
## 1pct 5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.177. Ispis ADF testa za niz CROBEX (gornji panel) i prva diferencija CROBEX-a (donji panel)

Primjer 1.38.

Iz primjera 1.36. provedimo testiranje nestacionarnosti niza desezonirani tečaj uz uključenu konstantu te potom konstantu i trend kao determinističke regresore, kako bismo ispitali je li stacionaran ili ne, i potom komentirajmo politiku HNB-a s obzirom na same rezultate.

Temeljem ispisa na slici 1.178., uočava se da se u oba testa odbacuje nulta hipoteza o nestacionarnosti niza desezonirani tečaj, stoga se zaključuje da je niz stacionaran, HNB uspješno brani tečaj kune naspram eura. Čitatelju za vježbu preostaje zapisati ADF jednadžbe i izračun test veličina.

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.970e-02 1.534e-02 3.891 0.000128 ***
## z.lag.1 -8.054e-03 2.058e-03 -3.914 0.000117 ***
## tt 2.874e-06 3.218e-06 0.893 0.372653
## z.diff.lag 9.240e-01 2.017e-02 45.808 < 2e-16 ***
##
## Value of test-statistic is: -3.9136 5.2892 7.9332
##
## Critical values for test statistics:
## 1pct 5pct 10pct
## tau3 -3.98 -3.42 -3.13
## phi2 6.15 4.71 4.05
## phi3 8.34 6.30 5.36

## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.059543 0.015337 3.882 0.000132 ***
## z.lag.1 -0.007983 0.002056 -3.883 0.000131 ***
## z.diff.lag 0.921364 0.019941 46.205 < 2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -3.8834 7.5409
##
## Critical values for test statistics:
## 1pct 5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phi1 6.47 4.61 3.79
```

Slika 1.178. Ispis ADF testa za niz tečaj, uključeni trend i konstanta (gornji panel) te samo konstanta (donji panel)

1.10.8. Pitanja za ponavljanje

1. Što su to nestacionarni vremenski nizovi?
2. Koje vrste trenda postoje?
3. Pojasnite deterministički trend u vremenskom nizu. Kako se utvrđuje? Kako se uklanja?
4. Što je to trend-stacionaran proces?
5. Pojasnite stohastički trend u vremenskom nizu. Kako se utvrđuje? Kako se uklanja?
6. Što je to diferenciranje vremenskog niza?
7. Što je to integriranost vremenskog niza?
8. Koja je posljedica ako se detrendira stohastički trend?
9. Koja je posljedica ako se diferencira deterministički trend?
10. Što je to model slučajnog pomaka? O kakvom se procesu radi?
11. Zapišite simbolički model slučajnog pomaka s konstantom i bez konstante.
12. Objasnite što je to problem prividne regresije.
13. Što su testovi jediničnog korijena?
14. Zapišite osnovnu jednadžbu Dickey-Fullerova testa, te hipoteze koje se testiraju.
15. Kada se koristi prošireni Dickey-Fullerov test?
16. Kako se određuje adekvatan broj pomaka zavisne varijable u Dickey-Fullerovom testu?
17. Koji se problemi javljaju kod provođenja Dickey-Fullerova testa?

1.11. Autoregresijski model distribuiranih pomaka, ARDL

Autoregresijski model distribuiranih pomaka (engl. ARDL – *autoregressive distributed lag*) je dinamički model u kojemu se razmatra veza između dviju varijabli, x i y , pri čemu je y zavisna, a x nezavisna varijabla. Pretpostavlja se da varijabla x utječe na varijablu y u različitim vremenskim pomacima, s obzirom da brojne ekonomske varijable vezane uz makroekonomske politike u pravilu ne reagiraju odmah na promjene u drugim varijablama. Nadalje, pretpostavlja se da su varijable stacionarne.

Ako se razmatra model s **beskonačnim brojem pomaka** (engl. *infinite lag model*) nezavisne varijable, razmatra se model:

$$y_t = \alpha + \sum_{j=0}^{\infty} \beta_j x_{t-j} + \varepsilon_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \varepsilon_t, \quad (1.175)$$

gdje promjena u varijabli x u nekom razdoblju t trajno ostaje u varijabli y . Ako su oba procesa u (1.175) stacionarna, jednadžba dugog roka, u kojoj se ostvaruje ravnoteža je:

$$E(y_t) = \alpha + \sum_{j=0}^{\infty} \beta_j E(x_{t-j}), \quad (1.176)$$

$$\mu_y = \alpha + \mu_x \sum_{j=0}^{\infty} \beta_j, \quad (1.177)$$

pa su **učinak x na y u dugom roku** $\sum_{j=0}^{\infty} \beta_j$ još i naziva multiplikator dugog roka. Da bi ta

vrijednost bila konačna, mora vrijediti: $\left| \sum_{j=0}^{\infty} \beta_j \right| < \infty$. Ako se razmotri trenutni učinak (učinak u

kratkome roku) varijable x na y , u svrhu konstrukcije funkcije impulsnog odziva (vidjeti odjeljak 1.9.7.), razmatra se parcijalna derivacija od (1.177) za vremensko razdoblje t :

$$\frac{\partial y_t}{\partial x_t} = \beta_0. \quad (1.178)$$

Greene (2002) definira β_0 kao **kratkoročni multiplikator**, tj. trenutni učinak promjene x na promjenu y , s obzirom da se radi o istom razdoblju t u kojemu se promijeni vrijednost x . Slično tome, vrijednost y za vremensko razdoblje $t+1$,

$$y_{t+1} = \alpha + \beta_0 x_{t+1} + \beta_1 x_t + \beta_2 x_{t-1} + \dots + \varepsilon_{t+1},$$

reagirat će na promjenu varijable x iz razdoblja t na sljedeći način:

$$\frac{\partial y_{t+1}}{\partial x_t} = \beta_1, \quad (1.179)$$

za vremensko razdoblje $t+2$,

$$y_{t+2} = \alpha + \beta_0 x_{t+2} + \beta_1 x_{t+1} + \beta_2 x_t + \dots + \varepsilon_{t+2},$$

$$\frac{\partial y_{t+2}}{\partial x_t} = \beta_2, \quad (1.180)$$

što se može poopćiti na sljedeći niz:

Razdoblje	t	$t+1$	$t+2$	$t+3$...	τ	...
Učinak	β_0	β_1	β_2	β_3	...	β_τ	...
Kumulativni učinak	β_0	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_2$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$...	$\beta_0 + \beta_1 + \beta_2 + \beta_3 + \dots + \beta_\tau$...

Kumulativni učinak promjene x na y u razdoblju τ iznosi: $\sum_{j=0}^{\tau} \beta_j$.

Općenito se **ARDL model s konačnim brojem pomaka** može pisati na sljedeći način:

$$y_t = \alpha + \sum_{j=1}^p \phi_j y_{t-j} + \sum_{j=0}^q \beta_j x_{t-j} + \varepsilon_t, \quad (1.181)$$

gdje se uočava da vrijednost varijable y ovisi o p svojih prethodnih vrijednosti, te q prethodnih vrijednosti varijable x . Parametri β_j su težine vezane uz pomake (engl. *lag weights*), te određuju uzorak učinka varijable x na varijablu y kroz vrijeme. Zato se govori o modelu distribuiranih pomaka, jer učinak varijable x na y nije samo trenutni, u razdoblju t , već se raspoređuje (distribuiraju) kroz vrijeme, tj. više razdoblja. Postavlja se pitanje određivanja pomaka p i posebno q , kao i što se postavlja pitanje stabilnosti procesa y , slično kao i za običan AR(p) model (korijeni inverznog karakterističnog korijena moraju po apsolutnoj vrijednosti biti veći od 1, vidjeti odjeljak 1.9.3.). Kako se model u (1.181) može zapisati pomoću operatora pomaka:

$$y_t - \sum_{j=1}^p \phi_j y_{t-j} = \alpha + \sum_{j=0}^q \beta_j x_{t-j} + \varepsilon_t, \quad (1.182)$$

$$(1 - \phi_1 L - \dots - \phi_p L^p) y_t = \alpha + (\beta_0 + \beta_1 L + \dots + \beta_q L^q) x_t + \varepsilon_t, \quad (1.183)$$

$$\phi(L) y_t = \alpha + \beta(L) x_t + \varepsilon_t, \quad (1.184)$$

model (1.184) možemo pisati i kao **racionalni** (Jorgenson, 1966; Kotzé, 2022; Wall, 1974) model distribuiranih pomaka:

$$y_t = \frac{\alpha}{\phi(L)} + \frac{\beta(L)}{\phi(L)} x_t + \frac{\varepsilon_t}{\phi(L)}, \quad (1.185)$$

tj.

$$y_t = \delta + \frac{\beta(L)}{\phi(L)} x_t + \eta_t. \quad (1.186)$$

Da bi proces bio stabilan, tj. stacionaran, korijeni karakteristične jednačbe $\phi(L) = 0$ moraju po apsolutnoj vrijednosti biti veći od 1.

Posebni slučajevi ARDL modela

Model (1.181) možemo općenito pisati kao $ARDL(p,q)$. U ovome odjeljku obradit će se nekoliko osnovnih slučajeva ARDL modela u ovisnosti o odabiru pomaka p i q .

ARDL(1,1) model oblika je:

$$y_t = \alpha + \phi_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t, \quad (1.187)$$

te kako su oba procesa stacionarni, očekivana vrijednost u (1.187) iznosi:

$$E(y_t) = \alpha + \phi_1 E(y_{t-1}) + \beta_0 E(x_t) + \beta_1 E(x_{t-1}) + 0, \quad (1.188)$$

$$\mu_y = \alpha + \phi_1 \mu_y + \beta_0 \mu_x + \beta_1 \mu_x \Rightarrow (1 - \phi_1) \mu_y = \alpha + (\beta_0 + \beta_1) \mu_x, \quad (1.189)$$

$$\mu_y = \frac{\alpha}{1 - \phi_1} + \frac{\beta_0 + \beta_1}{1 - \phi_1} \mu_x, \quad (1.190)$$

stoga je očekivani **učinak varijable x na varijablu y u dugom roku** jednak $\frac{\beta_0 + \beta_1}{1 - \phi_1}$. Uočimo

da mora vrijediti $1 - \phi_1 \neq 0$, tj. $1 \neq \phi_1$. Model (1.190) naziva se **model dugog roka**, ili jednačba dugog roka.

Ako se razmotri trenutni učinak (**učinak u kratkome roku**) varijable x na y , u svrhu konstrukcije funkcije impulsnog odziva (vidjeti odjeljak 1.9.4.3.), razmatra se parcijalna derivacija od (1.187) za vremensko razdoblje t :

$$\frac{\partial y_t}{\partial x_t} = \beta_0, \quad (1.191)$$

za vremensko razdoblje $t+1$,

$$\begin{aligned} y_{t+1} &= \alpha + \phi_1 \underbrace{y_t}_{=\alpha + \phi_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t} + \beta_0 x_{t+1} + \beta_1 x_t + \varepsilon_{t+1}, \\ &= (1 + \phi_1) \alpha + \phi_1^2 y_{t-1} + \phi_1 \beta_0 x_t + \beta_1 x_{t-1} + \phi_1 \varepsilon_t + \beta_0 x_{t+1} + \beta_1 x_t + \varepsilon_{t+1} \end{aligned}$$

$$\frac{\partial y_{t+1}}{\partial x_t} = \phi_1 \beta_0 + \beta_1, \quad (1.192)$$

za vremensko razdoblje $t+2$,

$$\frac{\partial y_{t+1}}{\partial x_t} = \phi_1 \beta_0 + \beta_1, \quad (1.192)$$

za vremensko razdoblje $t+2$,

$$\begin{aligned} y_{t+2} &= \alpha + \phi_1 \underbrace{y_{t+1}}_{=\alpha + \phi_1 y_t + \beta_0 x_{t+1} + \beta_1 x_t + \varepsilon_{t+1}} + \beta_0 x_{t+2} + \beta_1 x_{t+1} + \varepsilon_{t+2} \\ &= \alpha + \alpha \phi_1 + \alpha \phi_1^3 y_{t-1} + \phi_1^2 \beta_0 x_t + \phi_1^2 \beta_1 x_{t-1} + \phi_1^2 \varepsilon_t + \phi_1 \beta_0 x_{t+1} + \\ &\quad + \phi_1 \beta_1 x_t + \phi_1 \varepsilon_{t+1} + \beta_0 x_{t+2} + \beta_1 x_{t+1} + \varepsilon_{t+2} \end{aligned}$$

$$\frac{\partial y_{t+2}}{\partial x_t} = \phi_1^2 \beta_0 + \phi_1 \beta_1 = \phi_1 (\phi_1 \beta_0 + \beta_1), \quad (1.193)$$

što se može poopćiti na sljedeći niz:

Razdoblje	t	$t+1$	$t+2$	$t+3$...	τ
Učinak	β_0	$\phi_1 \beta_0 + \beta_1$	$\phi_1 (\phi_1 \beta_0 + \beta_1)$	$\phi_1^2 (\phi_1 \beta_0 + \beta_1)$...	$\phi_1^{\tau-1} (\phi_1 \beta_0 + \beta_1)$
Kumulativni učinak	β_0	$(1 + \phi_1) \beta_0 + \beta_1$	$(1 + \phi_1 + \phi_1^2) \beta_0 + (1 + \phi_1) \beta_1$	$(1 + \phi_1 + \phi_1^2 + \phi_1^3) \beta_0 + (1 + \phi_1 + \phi_1^2) \beta_1$...	$\beta_0 \sum_{t=0}^{\tau} \phi_1^t + \beta_1 \sum_{t=0}^{\tau-1} \phi_1^t$

Ako se razmatra dugi rok, kumulativni učinak $\beta_0 \sum_{t=0}^{\tau} \phi_1^t + \beta_1 \sum_{t=0}^{\tau-1} \phi_1^t$ može se računati kao granična vrijednost kada $\tau \rightarrow \infty$:

$$\lim_{\tau \rightarrow \infty} \left[\beta_0 \sum_{t=0}^{\tau} \phi_1^t + \beta_1 \sum_{t=0}^{\tau-1} \phi_1^t \right], \quad (1.194)$$

i ako vrijedi pretpostavka da je $|\phi_1| < 1$, slijedi:

$$\lim_{\tau \rightarrow \infty} \left[\beta_0 \sum_{t=0}^{\tau} \phi_1^t + \beta_1 \sum_{t=0}^{\tau-1} \phi_1^t \right] = \lim_{\tau \rightarrow \infty} \left[\beta_0 \frac{1 - \phi_1^{\tau}}{1 - \phi_1} + \beta_1 \frac{1 - \phi_1^{\tau-1}}{1 - \phi_1} \right] = \frac{\beta_0}{1 - \phi_1} + \frac{\beta_1}{1 - \phi_1} = \frac{\beta_0 + \beta_1}{1 - \phi_1}, \quad (1.195)$$

jer se radi o sumi geometrijskog niza u oba pribrojnika pod limesom, i upravo se dobije učinak u dugome roku kojeg smo već definirali u (1.190).

Primjer 1.39.

Učitajmo datoteku „ardl.txt“ u Rstudio, te definirajmo da se radi o vremenskim nizovima s početnim datumom siječanj 2004. godine, a frekvencija je mjesečna. Datoteka sadrži podatke o varijablama x , y , z i w . Procijenimo ARDL(1,1) model gdje je zavisna varijabla y , a nezavisna x , zapišimo sam model, izračunajmo učinak varijable x na y u kratkome roku, učinak u dugome roku, te formirajmo funkciju impulsnog odziva i komentirajmo rezultate.

```

ardl<-read.table("ardl.txt",header=T, sep="\t")
ardl<-ts(ardl,start=c(2004,1),frequency = 12)
library(dynlm)

dynlm(y~L(y)+x+L(x),data=ardl)

##
## Time series regression with "ts" data:
## Start = 2004(2), End = 2021(6)
##
## Call:
## dynlm(formula = y ~ L(y) + x + L(x), data = ardl)
##
## Coefficients:
## (Intercept)          L(y)              x          L(x)
##          20.0           0.7           0.8           0.4

```

Slika 1.179. Učitavanje podataka i procjena ARDL(1,1) modela

U okviru paketa dynlm koristi se istoimena naredba za dinamičke modele:

$$\text{dynlm}(y \sim L(y) + x + L(x), \text{data} = \text{ardl}),$$

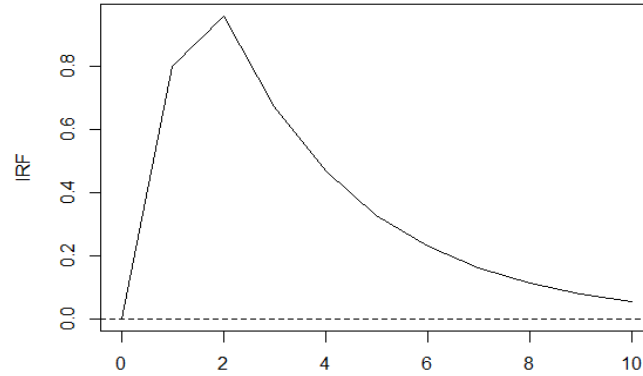
gdje je zapis veoma sličan linearnome modelu, tako da se s lijeve strane oznake \sim piše naziv zavisne varijable, a nakon desne strane imena varijabli koje su nezavisne. Oznaka $L(\dots)$ je naredba kojom se računa pomak varijable čije ime pišemo u zagradu, vidjeti sliku 1.179. Uočava se da je model sljedeći: $\hat{y}_t = 20 + 0,7y_{t-1} + 0,8x_t + 0,4x_{t-1}$. Učinak u kratkome roku iznosi 0,8, dok se učinak dugog roka iščitava iz izračuna jednadžbe dugog roka:

$$E(y_t) = 20 + 0,7E(y_{t-1}) + 0,8E(x_t) + 0,4E(x_{t-1}) \Rightarrow \mu_y - 0,7\mu_y = 20 + 0,8\mu_x + 0,4\mu_x,$$

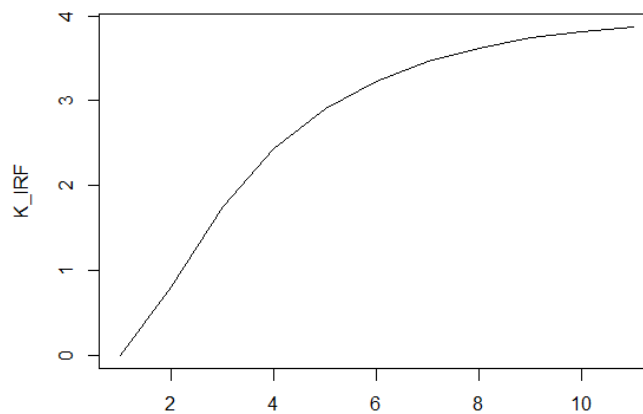
$$(1 - 0,7)\mu_y = 20 + (0,8 + 0,4)\mu_x \Rightarrow \mu_y = \frac{20}{1-0,7} + \frac{1,2}{1-0,7}\mu_x \Rightarrow \mu_y = 666,67 + 4\mu_x.$$

U dugom roku je učinak varijable x na varijablu y jednak 4. Ako se razmotri funkcija impulsnog odziva y na x , slika 1.180. predočava⁹ da je učinak povećanja x na y pozitivan, najjači je u drugom razdoblju, nakon čega se umjereno smanjuje prema vrijednosti nula. Ako razmotrimo kumulativni učinak, na slici 1.181., uočavamo da se nakon nekoliko razdoblja učinak približava vrijednosti 4, i već oko desetog pomaka kažemo da konvergira prema dugoročnom učinku, kojeg smo izračunali u jednadžbi dugog roka.

⁹ **Napomena:** U razdoblju pisanja ove publikacije, nije pronađen paket u kojem se vrši procjena funkcija impulsnog odziva za ARDL modele, koja bi uključivala samu procjenu, kao i intervalnu procjenu. Zato ako se pogledaju naredbe u pratećoj R-skripti za ovaj dio izračuna, radi se o dosta nezgrapnom zapisu. Međutim, naredbe su priređene na način da se razumije što je koji dio tih naredbi, s obzirom da je potrebno samo definirati koeficijente koji su procijenjeni u ARDL modelima. Zato grafički prikazi funkcija impulsnog odziva imaju samo procjenu tih odziva, bez intervalnih procjena.



Slika 1.180. Funkcija impulsnog odziva ARDL(1,1) modela



Slika 1.181. Kumulativna funkcija impulsnog odziva ARDL(1,1) modela

ARDL(1,0) model oblika je:

$$y_t = \alpha + \phi_1 y_{t-1} + \beta_0 x_t + \varepsilon_t, \quad (1.196)$$

te kako su oba procesa stacionarna, očekivana vrijednost u (1.196) iznosi:

$$E(y_t) = \alpha + \phi_1 E(y_{t-1}) + \beta_0 E(x_t) + 0, \quad (1.197)$$

$$\mu_y = \alpha + \phi_1 \mu_y + \beta_0 \mu_x \Rightarrow (1 - \phi_1) \mu_y = \alpha + \beta_0 \mu_x, \quad (1.198)$$

$$\mu_y = \frac{\alpha}{1 - \phi_1} + \frac{\beta_0}{1 - \phi_1} \mu_x, \quad (1.199)$$

stoga je očekivani učinak varijable x na varijablu y u dugom roku jednak $\frac{\beta_0}{1 - \phi_1}$. Uočimo da mora vrijediti $1 - \phi_1 \neq 0$, tj. $1 \neq \phi_1$.

Ako se razmotri trenutni učinak (učinak u kratkome roku) varijable x na y , u svrhu konstrukcije funkcije impulsnog odziva (vidjeti odjeljak 1.9.4.3.), razmatra se parcijalna derivacija od (1.196) za vremensko razdoblje t :

$$\frac{\partial y_t}{\partial x_t} = \beta_0, \quad (1.200)$$

za vremensko razdoblje $t+1$,

$$\begin{aligned} y_{t+1} &= \alpha + \phi_1 \underbrace{y_t}_{=\alpha + \phi_1 y_{t-1} + \beta_0 x_t + \varepsilon_t} + \beta_0 x_{t+1} + \varepsilon_{t+1}, \\ &= (1 + \phi_1)\alpha + \phi_1^2 y_{t-1} + \phi_1 \beta_0 x_t + \phi_1 \varepsilon_t + \beta_0 x_{t+1} + \varepsilon_{t+1} \end{aligned}$$

$$\frac{\partial y_{t+1}}{\partial x_t} = \phi_1 \beta_0, \quad (1.201)$$

za vremensko razdoblje $t+2$,

$$\begin{aligned} y_{t+2} &= \alpha + \phi_1 \underbrace{y_{t+1}}_{=\alpha + \phi_1 y_t + \beta_0 x_{t+1} + \varepsilon_{t+1}} + \beta_0 x_{t+2} + \varepsilon_{t+2} \\ &= \alpha + \phi_1 \alpha + \phi_1^2 \underbrace{y_t}_{=\alpha + \phi_1 y_{t-1} + \beta_0 x_t + \varepsilon_t} + \phi_1 \beta_0 x_{t+1} + \phi_1 \varepsilon_{t+1} + \beta_0 x_{t+2} + \varepsilon_{t+2}, \\ &= (1 + \phi_1)\alpha + \phi_1^2 (\alpha + \phi_1 y_{t-1} + \beta_0 x_t + \varepsilon_t) + \phi_1 \beta_0 x_{t+1} + \phi_1 \varepsilon_{t+1} + \beta_0 x_{t+2} + \varepsilon_{t+2} \end{aligned}$$

$$\frac{\partial y_{t+2}}{\partial x_t} = \phi_1^2 \beta_0, \quad (1.202)$$

što se može poopćiti na sljedeći niz:

Razdoblje	t	$t+1$	$t+2$	$t+3$...	τ
Učinak	β_0	$\phi_1 \beta_0$	$\phi_1^2 \beta_0$	$\phi_1^3 \beta_0$...	$\phi_1^\tau \beta_0$
Kumulativni učinak	β_0	$(1 + \phi_1)\beta_0$	$(1 + \phi_1 + \phi_1^2)\beta_0$	$(1 + \phi_1 + \phi_1^2 + \phi_1^3)\beta_0$...	$\beta_0 \sum_{t=0}^{\tau} \phi_1^t$

Ako se razmatra dugi rok, kumulativni učinak $\beta_0 \sum_{t=0}^{\tau} \phi_1^t$ može se računati kao granična vrijednost kada $\tau \rightarrow \infty$:

$$\lim_{\tau \rightarrow \infty} \left[\beta_0 \sum_{t=0}^{\tau} \phi_1^t \right], \quad (1.203)$$

i ako vrijedi pretpostavka da je $|\phi_1| < 1$, slijedi:

$$\lim_{\tau \rightarrow \infty} \left[\beta_0 \sum_{t=0}^{\tau} \phi_1^t \right] = \lim_{\tau \rightarrow \infty} \left[\beta_0 \frac{1 - \phi_1^\tau}{1 - \phi_1} \right] = \frac{\beta_0}{1 - \phi_1}, \quad (1.204)$$

jer se radi o sumi geometrijskog niza pod limesom, i upravo se dobije učinak u dugome roku kojeg smo već definirali u (1.199).

Primjer 1.40.

Učitajmo datoteku „ardl.txt“ u Rstudio, te definirajmo da se radi o vremenskim nizovima s početnim datumom siječanj 2004. godine, a frekvencija je mjesečna. Datoteka sadrži podatke o varijablama x , y , z i w . Procijenimo ARDL(1,0) model gdje je zavisna varijabla z , a nezavisna x , zapišimo sam model, izračunajmo učinak varijable x na z u kratkome roku, učinak u dugome roku, te formirajmo funkciju impulsnog odziva i komentirajmo rezultate.

```
dynlm(z~L(z)+x,data=ardl)

##
## Time series regression with "ts" data:
## Start = 2004(2), End = 2021(6)
##
## Call:
## dynlm(formula = z ~ L(z) + x, data = ardl)
##
## Coefficients:
## (Intercept)      L(z)          x
##          20.0         0.8         0.7
```

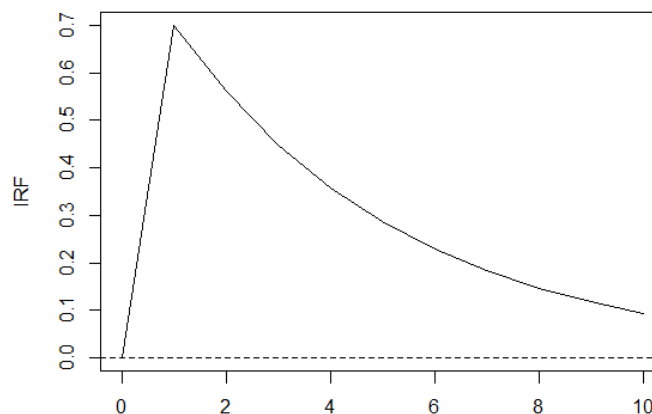
Slika 1.182. Procjena ARDL(1,0) modela

Uočava se da je model sljedeći: $\hat{z}_t = 20 + 0,8z_{t-1} + 0,7x_t$. Učinak u kratkome roku iznosi 0,7, dok se učinak dugog roka iščitava iz izračuna jednadžbe dugog roka:

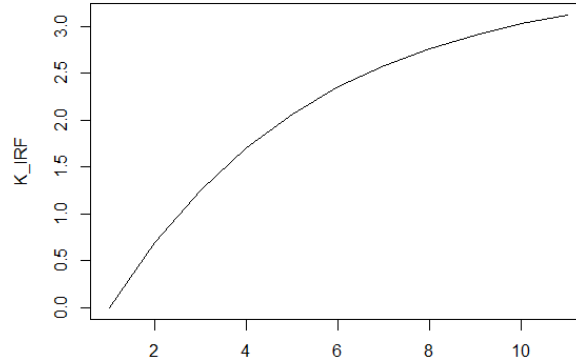
$$E(z_t) = 20 + 0,8E(z_{t-1}) + 0,7E(x_t) \Rightarrow \mu_z - 0,8\mu_z = 20 + 0,7\mu_x,$$

$$(1 - 0,8)\mu_z = 20 + 0,7\mu_x \Rightarrow \mu_z = \frac{20}{1-0,8} + \frac{0,7}{1-0,8}\mu_x \Rightarrow \mu_z = 100 + 3,5\mu_x.$$

U dugom roku je učinak varijable x na varijablu z jednak 3,5. Ako se razmotri funkcija impulsnog odziva z na x , slika 1.183. predočava da je učinak pozitivan, najjači je u drugom razdoblju, nakon čega se umjereno smanjuje prema vrijednosti nula. Ako razmotrimo kumulativni učinak, na slici 1.184., uočavamo da se nakon nekoliko razdoblja učinak približava vrijednosti 3, te bi se uskoro približila vrijednosti 3,5 još nakon nekoliko razdoblja.



Slika 1.183. Funkcija impulsnog odziva ARDL(1,0) modela



Slika 1.184. Kumulativna funkcija impulsnog odziva ARDL(1,0) modela

ARDL(0,1) model oblika je:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t, \quad (1.205)$$

te kako su oba procesa stacionarna, očekivana vrijednost u (1.196) iznosi:

$$E(y_t) = \alpha + \beta_0 E(x_t) + \beta_1 E(x_{t-1}) + 0, \quad (1.206)$$

$$\mu_y = \alpha + \beta_0 \mu_x + \beta_1 \mu_x \Rightarrow \mu_y = \alpha + (\beta_0 + \beta_1) \mu_x, \quad (1.207)$$

pa je očekivani učinak varijable x na varijablu y u dugom roku jednak $\beta_0 + \beta_1$.

Ako se razmotri trenutni učinak (učinak u kratkome roku) varijable x na y , u svrhu konstrukcije funkcije impulsnog odziva (vidjeti odjeljak 1.9.7.), razmatra se parcijalna derivacija od (1.205) za vremensko razdoblje t :

$$\frac{\partial y_t}{\partial x_t} = \beta_0, \quad (1.208)$$

za vremensko razdoblje $t+1$,

$$y_{t+1} = \alpha + \beta_0 x_{t+1} + \beta_1 x_t + \varepsilon_t,$$

$$\frac{\partial y_{t+1}}{\partial x_t} = \beta_1, \quad (1.209)$$

za vremensko razdoblje $t+2$,

$$y_{t+2} = \alpha + \beta_0 x_{t+2} + \beta_1 x_{t+1} + \varepsilon_t,$$

$$\frac{\partial y_{t+2}}{\partial x_t} = 0, \quad (1.210)$$

što se može poopćiti na sljedeći niz:

Razdoblje	t	$t+1$	$t+2$	$t+3$...	τ
Učinak	β_0	β_1	0	0	...	0
Kumulativni učinak	β_0	$\beta_0 + \beta_1$	$\beta_0 + \beta_1$	$\beta_0 + \beta_1$...	$\beta_0 + \beta_1$

Ako se razmatra dugi rok, kumulativni učinak $\beta_0 + \beta_1$ može se računati kao granična vrijednost ako $\tau \rightarrow \infty$:

$$\lim_{\tau \rightarrow \infty} [\beta_0 + \beta_1], \quad (1.211)$$

$$\lim_{\tau \rightarrow \infty} [\beta_0 + \beta_1] = \beta_0 + \beta_1, \quad (1.212)$$

i upravo se dobije učinak u dugome roku kojeg smo već definirali u (1.207).

Primjer 1.41.

Učitajmo datoteku „ardl.txt“ u Rstudio, te definirajmo da se radi o vremenskim nizovima s početnim datumom siječanj 2004. godine, a frekvencija je mjesečna. Datoteka sadrži podatke o varijablama x , y , z i w . Procijenimo ARDL(0,1) model gdje je zavisna varijabla w , a nezavisna x , zapišimo sam model, izračunajmo učinak varijable x na w u kratkome roku, učinak u dugome roku, te formirajmo funkciju impulsnog odziva i komentirajmo rezultate.

Uočava se da je model sljedeći: $\widehat{w}_t = 20 + 0,6x_t + 0,8x_{t-1}$. Učinak u kratkome roku iznosi 0.8, dok se učinak dugog roka iščitava iz izračuna jednadžbe dugog roka:

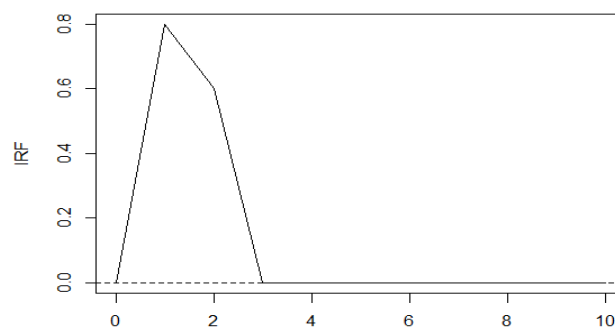
$$E(w_t) = 20 + 0,6E(x_{t-1}) + 0,8E(x_t) \Rightarrow \mu_w = 20 + 0,6\mu_x + 0,8\mu_x \Rightarrow \mu_w = 20 + 1,4\mu_x.$$

U dugom roku je učinak varijable x na varijablu w jednak 1.4. Ako se razmotri funkcija impulsnog odziva w na x , slika 1.186. predočava da je učinak pozitivan, najjači je u drugom razdoblju, nakon čega se u trećem razdoblju potpuno nestaje. Ako razmotrimo kumulativni učinak, na slici 1.187., uočavamo da se u trećem razdoblju ostvaruje učinak dugog roka, u iznosu 1,4 i tu na dalje stagnira.

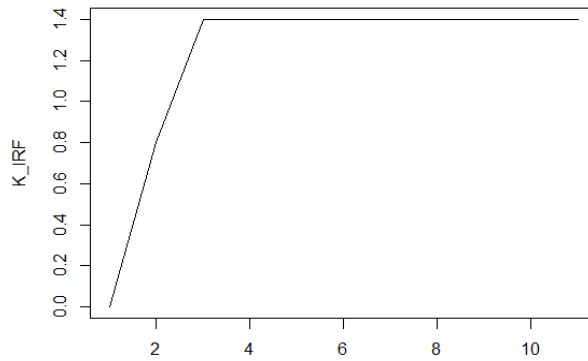
```
dynlm(w~L(x)+x,data=ardl)

## Time series regression with "ts" data:
## Start = 2004(2), End = 2021(6)
##
## Call:
## dynlm(formula = w ~ L(x) + x, data = ard1)
##
## Coefficients:
## (Intercept)          L(x)              x
##          20.0            0.6            0.8
```

Slika 1.185. Procjena ARDL(0,1) modela



Slika 1.186. Funkcija impulsnog odziva ARDL(0,1) modela

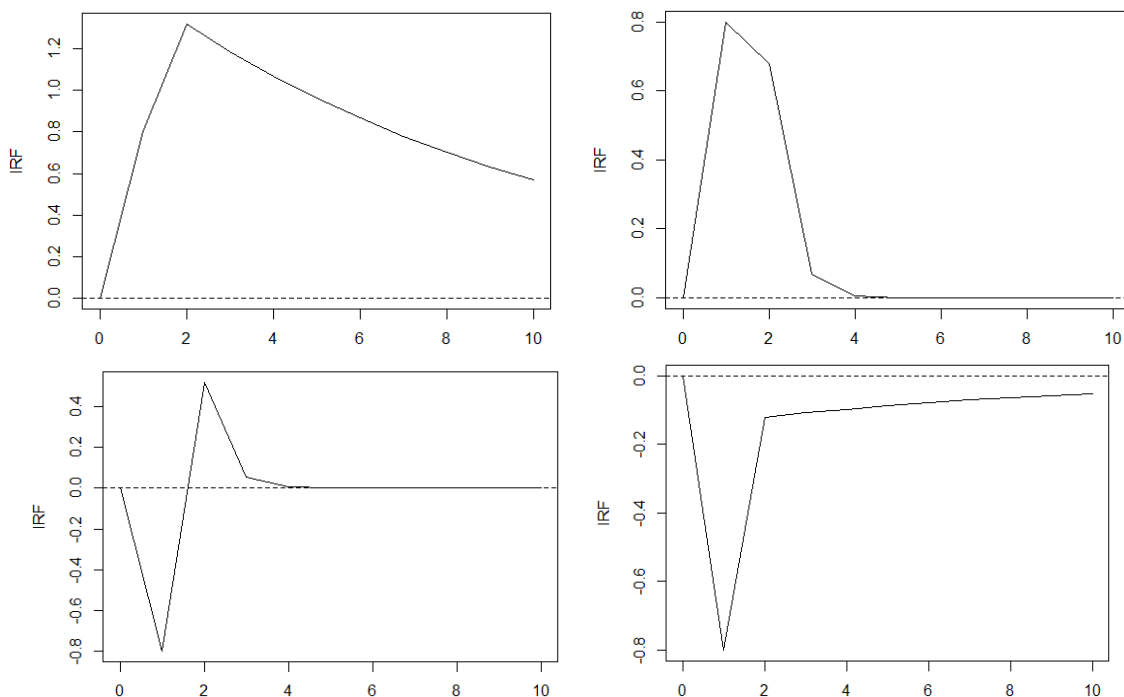


Slika 1.187. Kumulativna funkcija impulsnog odziva ARDL(0,1) modela

Primjer 1.42.

Usporedimo sljedeće funkcije impulsnog odziva dane na panelima na slici 1.188.

Ako se razmotri gornji lijevi panel na slici 1.187., uočava se da šok u varijabli x pozitivno utječe na zavisnu varijablu, najjači učinak je u razdoblju 2, te veoma sporo nestaje kroz naredna razdoblja. Uočimo da u razdoblju 10 je učinak na razini oko 0,6. S druge strane, gornji desni panel upućuje također na pozitivan učinak šoka u varijabli x na zavisnu varijablu, najjači je učinak u prvom razdoblju, no već u četvrtom razdoblju taj učinak nestaje, dakle veoma je kratkotrajan, za razliku od prethodnog procesa. Nadalje, donji lijevi panel upućuje da je inicijalna reakcija zavisne varijable na šok u varijabli x negativna u razdoblju 1, nakon toga je pozitivna u razdoblju 2 (dakle, alternira), i nakon toga brzo nestaje učinak, nestaje u razdoblju 4. Konačno, donji desni panel upućuje na negativan učinak šoka u varijabli x na zavisnu varijablu, najjači učinak je u razdoblju 1, nakon čega brzo iščezava taj učinak od drugog razdoblja na dalje.



Slika 1.188. Primjeri funkcija impulsnog odziva

Primjer 1.43.

Učitajmo datoteku „ardl2.txt“ u RStudio. Datoteka sadrži mjesečne podatke¹⁰ o kretanju broja dolazaka turista u Španjolsku i Italiju, i harmoniziranim indeksima potrošačkih cijena u obje zemlje (HICP, 2015=100). Definirajmo da se radi o vremenskim nizovima s početnim datumom siječanj 1998. godine, desezonirajmo oba niza i izračunajmo stope rasta pomoću formule (1.7) (mjesec ove godine u odnosu na isti mjesec prethodne), i procijenimo ARDL(1,1), ARDL(1,0) i ARDL(0,1) modele u kojemu stopa rasta dolazaka turista u Španjolsku ovisi o stopi rasta cijena u Španjolskoj. Interpretirajmo učinak stope rasta cijena na stopu dolazaka u kratkome i dugome roku, grafički predočimo funkciju impulsnog odziva i komentirajmo. Nakon toga, procijenimo isti model za dva poduzorka: jedan je do prosinca 2007. godine, a drugi od siječnja 2008. godine do kraja razmatranog razdoblja. Usporedimo i komentirajmo rezultate.

```
ardl2<-read.table("ardl2.txt",sep="\t",header=T)
dolasci<-ts(ardl2$S_dol,start=c(1998,1),frequency = 12)
dolasci<-dolasci/decompose(dolasci,type="multiplicative")$seasonal
cijene<-ts(ardl2$S_hicp,start=c(1998,1),frequency = 12)
cijene<-cijene/decompose(cijene,type="multiplicative")$seasonal

stopa.cijene<-diff(log(cijene),12)
stopa.dolasci<-diff(log(dolasci),12)

library(dynlm)
m1<-dynlm(stopa.dolasci~L(stopa.dolasci)+stopa.cijene+L(stopa.cijene))
m2<-dynlm(stopa.dolasci~L(stopa.dolasci)+stopa.cijene)
m3<-dynlm(stopa.dolasci~stopa.cijene+L(stopa.cijene))
library(stargazer)
stargazer(list(m1,m2,m3),type="text")
```

Slika 1.189. Učitavanje podataka i definiranje modela

Slika 1.189. predočava niz naredbi za učitavanje i definiranje varijabli za modele, kao i samu procjenu modela u primjeru. Prvih šest redaka se odnose na već poznato učitavanje podataka i definiranje vremenskog niza koji se potom desezonira. Nakon toga su definirane godišnje stope rasta. Posljednji dio naredbi odnosi se na procjenu tri dinamička modela (m1, m2 i m3), u kojima se definira neka od varijanti ARDL modela. Paket i istoimena naredba stargazer koriste se za spajanje ispisa nekoliko modela u jednu tablicu:

```
stargazer(list(m1,m2,m3),type="text")
```

gdje je prvi argument lista modela koji se žele predočiti u jednoj tablici, dok se drugi argument type odnosi na način kako će se rezultati ispisati (u obliku teksta).

Temeljem ispisa na slici 1.190. možemo zapisati sljedeće modele:

$$\text{ARDL}(1,1): \hat{y}_t = 0,03 + 0,239y_{t-1} + 4,019x_t - 4,02x_{t-1},$$

$$\text{ARDL}(1,0): \hat{y}_t = 0,028 + 0,231y_{t-1} + 0,147x_t,$$

$$\text{ARDL}(0,1): \hat{y}_t = 0,039 + 3,819x_t - 3,792x_{t-1}$$

Sada su redom učinci u kratkome roku: 4,019, 0,147 te 3,819. Dakle, prema sva tri modela kratkoročno povećanje stope rasta cijena u Španjolskoj povećava stopu rasta dolazaka turista u

¹⁰ Izvor: Eurostat (2020).

tu zemlju. Najmanji je učinak u ARDL(1,0) modelu. Pozitivan učinak mogao bi se interpretirati time da se turistička putovanja planiraju unaprijed, stoga povećanje stope rasta cijena u nekome mjesecu ne može u istome mjesecu negativno utjecati na stopu rasta dolazaka turista. Jednadžbe dugog roka i učinci u dugom roku su sljedeći:

ARDL(1,1): $(1 - 0,239)\mu_y = 0,03 + (4,019 - 4,020)\mu_x \Rightarrow \mu_y = 0,039 - 0,0013\mu_x$, učinak u dugome roku: $-0,0013$;

ARDL(1,0): $(1 - 0,231)\mu_y = 0,028 + 0,147\mu_x \Rightarrow \mu_y = 0,036 + 0,191\mu_x$, učinak u dugome roku: $0,191$; te

ARDL(0,1): $\mu_y = 0,039 + (3,819 - 3,792)\mu_x \Rightarrow \mu_y = 0,039 + 0,027\mu_x$, učinak u dugome roku: $0,027$.

Uočava se da je u dugome roku jedini negativni učinak povećanja stope rasta cijena u Španjolskoj na stopu rasta dolazaka turista je u ARDL(1,1) modelu.

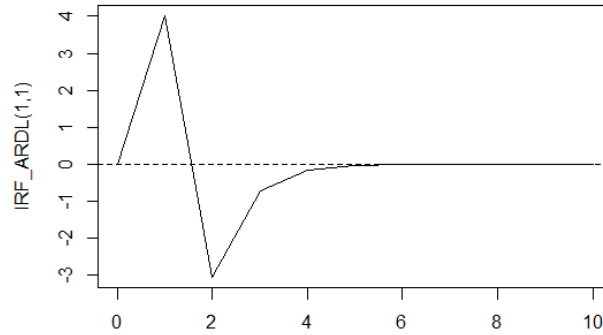
Dependent variable:			

	(1)	(2)	(3)

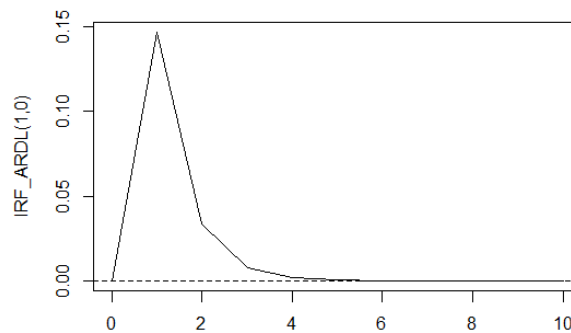
## L(stopa.dolasci)	0.239***	0.231***	
##	(0.060)	(0.061)	
## stopa.cijene	4.019***	0.147	3.819***
##	(1.312)	(0.355)	(1.350)
## L(stopa.cijene)	-4.020***		-3.792***
##	(1.313)		(1.350)
## Constant	0.030***	0.028***	0.039***
##	(0.009)	(0.009)	(0.009)
##			
##			
## Observations	252	252	252
## R2	0.090	0.055	0.031
## Adjusted R2	0.079	0.048	0.024
## Residual Std. Error	0.084 (df = 248)	0.086 (df = 249)	0.087 (df = 249)
## F Statistic	8.140*** (df = 3; 248)	7.279*** (df = 2; 249)	4.048** (df = 2; 249)
##			
## Note:			*p<0.1; **p<0.05; ***p<0.01

Slika 1.190. Ispis sva tri ARDL modela

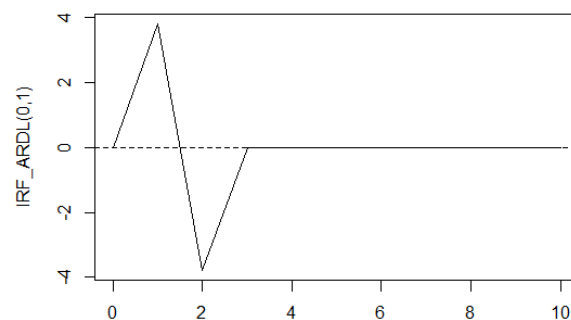
Ako se predoče funkcije impulsnog odziva za sva tri modela, redom na slikama 1.191., 1.192. i 1.193., uočava se da za ARDL(1,1) model povećanje stope rasta cijena u Španjolskoj najprije pozitivno utječe na stopu rasta dolazaka turista u Španjolsku, nakon toga je u drugom razdoblju učinak negativan i brzo nakon toga učinak nestaje i stopa rasta dolazaka turista se vraća na prosječnu razinu. Ako se razmotri funkcija impulsnog odziva za ARDL(1,0) model, učinak povećanja stope rasta cijena u Španjolskoj ima pozitivan učinak na stopu rasta dolazaka turista u Španjolsku u prva tri razdoblja, nakon čega brzo taj učinak nestaje. Za ARDL(0,1) model je slična interpretacija kao za ARDL(1,1) model, pri čemu učinak već nestaje u trećem razdoblju.



Slika 1.191. Funkcija impulsnog odziva za ARDL(1,1) model



Slika 1.192. Funkcija impulsnog odziva za ARDL(1,0) model



Slika 1.193. Funkcija impulsnog odziva za ARDL(0,1) model

Nadalje, razmotrimo još procjenu ARDL(1,1) modela u razdoblju od siječnja 1998. do prosinca 2007. godine, i posebno za drugo podrazdoblje od siječnja 2008. godine do kraja razmatranog razdoblja. Usporedimo rezultate procjena i interpretacije, te funkcije impulsnog odziva za oba podrazdoblja. Slika 1.194. predočava naredbe za procjenu ARDL modela za dva podrazdoblja, kao i koeficijente samog modela, dok su funkcije impulsnog odziva predočene na slici 1.195., a kumulativne funkcije impulsnog odziva na slici 1.196. Modeli za svako podrazdoblje su sljedeći:

Prvo podrazdoblje: $\hat{y}_t = 0,15 + 0,14y_{t-1} - 2,61x_t + 0,04x_{t-1}$, učinak u kratkom roku iznosi $-2,61$; model dugog roka je: $(1 - 0,14)\mu_y = 0,15 + (-2,61 + 0,04)\mu_x \Rightarrow \mu_y = 0,17 - 2,99\mu_x$, te je učinak u dugom roku $-2,99$. Dakle, učinak povećanja stope rasta cijena u Španjolskoj u kratkome roku je nešto manji od učinka u dugome roku na stopu rasta dolazaka turista u prvom podrazdoblju. Učinak je i u kratkom i dugom roku negativan.

Drugo podrazdoblje: $\hat{y}_t = 0,04 - 0,03y_{t-1} + 3,70x_t - 5,35x_{t-1}$, učinak u kratkom roku iznosi 3,70; model dugog roka je: $(1 + 0,03)\mu_y = 0,04 + (3,70 - 5,35)\mu_x \Rightarrow \mu_y = 0,039 - 1,60\mu_x$, te je učinak u dugom roku $-1,60$. Dakle, učinak povećanja stope rasta cijena u Španjolskoj u kratkome roku je pozitivan, dok u dugome roku postaje negativan.

```
m4<-dynam(stopa.dolasci~L(stopa.dolasci)+stopa.cijene+L(stopa.cijene),end=c(2007,12))
m5<-dynam(stopa.dolasci~L(stopa.dolasci)+stopa.cijene+L(stopa.cijene),start=c(2008,1))
m4$coefficients

##      (Intercept) L(stopa.dolasci)      stopa.cijene  L(stopa.cijene)
##      0.14652293      0.14100783      -2.60851966      0.03838064

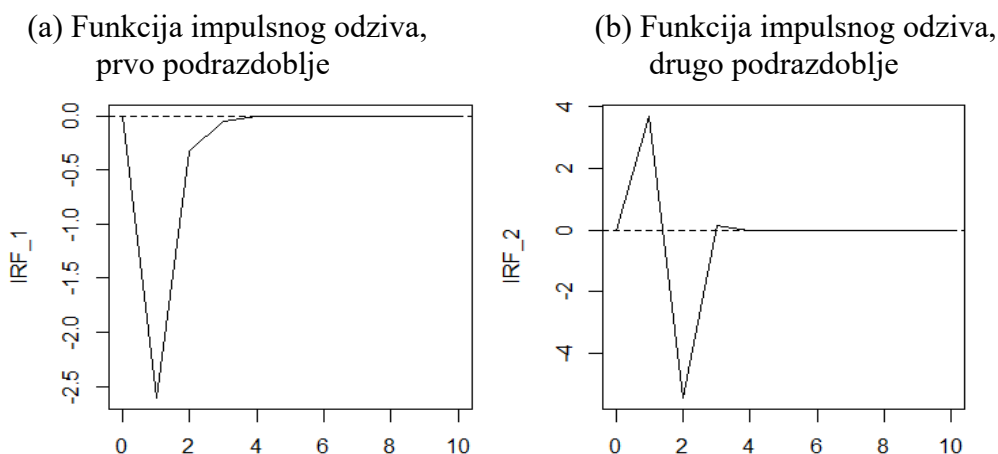
m5$coefficients

##      (Intercept) L(stopa.dolasci)      stopa.cijene  L(stopa.cijene)
##      0.03514457      -0.02923156      3.69706994      -5.35397142
```

Slika 1.193. Naredbe za procjenu ARDL modela u dva podrazdoblja

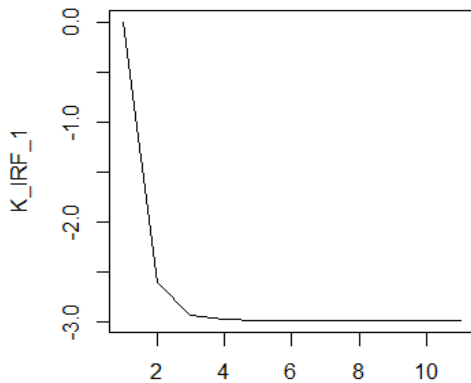
Slika 1.194. predočava da u prvom podrazdoblju učinak povećanja stope rasta cijena u Španjolskoj ima negativan učinak na stopu rasta dolazaka turista u Španjolsku, pri čemu je učinak najjači u prvom mjesecu, i već nestaje oko četvrtog mjeseca od šoka. S druge strane, u drugom podrazdoblju je najprije taj učinak pozitivan u prvom mjesecu, potom je nešto jači negativan u drugom mjesecu i nakon toga brzo nestaje. Ako se razmotre kumulativni učinci na slici 1.195., uočava se da se učinak dugog roka ostvaruje veoma brzo u oba podrazdoblja, $-2,99$ za panel (a), te vrijednost $-1,60$ za panel (b). U prvom podrazdoblju taj je učinak jači u dugome roku, u odnosu na drugo podrazdoblje.

Napomena: U primjeru su razmatrana dva podrazdoblja kao primjer analize u kojoj se ovakva analiza može raditi tako da se određeno razdoblje može podijeliti na podrazdoblja u kojima se pretpostavlja drugačije ponašanje relevantnih varijabli, kao i njihovih međudnosa. U ovome primjeru se radi o relativno kratkim vremenskim nizovima, pa u konačnici analiza vezana za dugi rok u praksi ne bi imala smisla. Primjer je predočen u edukativne svrhe. Stoga u praksi, ako se raspolaže s dovoljno dugim vremenskim nizovima, ovakva analiza bi imala veći značaj.

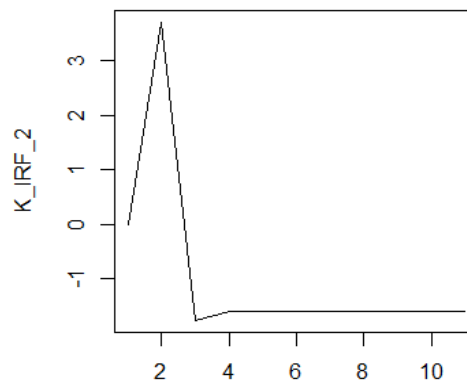


Slika 1.194. Funkcije impulsnog odziva

(a) Kumulativna funkcija impulsnog odziva, prvo podrazdoblje



(b) Kumulativna funkcija impulsnog odziva, drugo podrazdoblje



Slika 1.195. Kumulativne funkcije impulsnog odziva

Primjer 1.44.

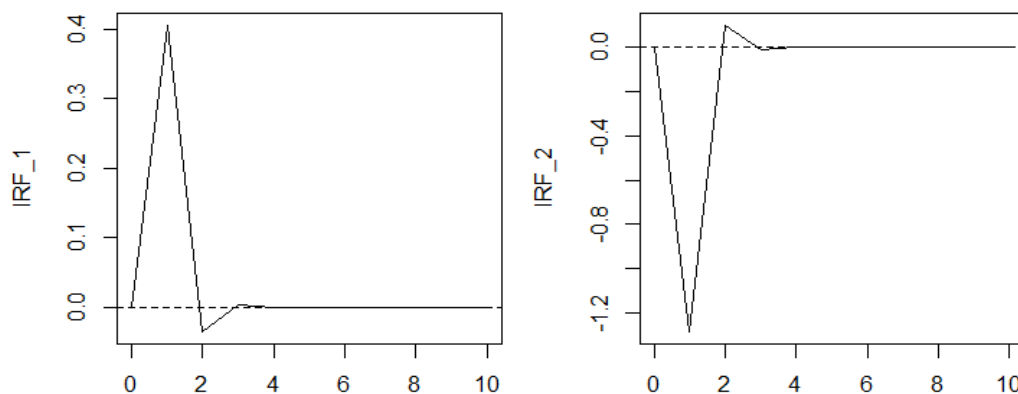
Temeljem prethodnog primjera (1.43) provedimo postupak za dva podniza za podatke o Italiji.

Procijenjen je ARDL(1,0) model za podatke o Italiji (čitatelju preostaje za vježbu da provede sve što je napravljeno u prethodnom primjeru za Španjolsku), za razdoblje do prosinca 2007. i drugo podrazdoblje od siječnja 2008. godine.

```
m6$coefficients
##      (Intercept) L(stopa.dolasci)   stopa.cijene
##      0.02214917   -0.08708713      0.40630577

m7$coefficients
##      (Intercept) L(stopa.dolasci)   stopa.cijene
##      0.03802875   -0.07903659     -1.28841439
```

Slika 1.196. Ispis modela ARDL(1,0) za oba podrazdoblja



Slika 1.197. Funkcije impulsnog odziva

Model u prvom podrazdoblju je $\hat{y}_t = 0,022 - 0,087y_{t-1} + 0,406x_t$, te je učinak u kratkome roku jednak 0,406. Jednadžba dugog roka za ovo podrazdoblje je $\mu_y = 0,02 + 0,374\mu_x$ i vidimo da je učinak u dugome roku nešto manji te iznosi 0,374. U drugom podrazdoblju, model je oblika

$\hat{y}_t = 0,038 - 0,079y_{t-1} - 1,288x_t$, učinak u kratkom roku iznosi $-1,288$, te je model dugog roka sljedeći: $\mu_y = 0,035 - 1,194\mu_x$, te je i ovdje učinak u dugome roku nešto manji i iznosi $-1,194$. Na lijevome panelu slike 1.197. se uočava da u prvome podrazdoblju povećanje stope cijena u Italiji u prvom mjesecu ima pozitivan učinak na stopu rasta dolazaka turista, nakon čega u drugom mjesecu taj učinak ima negativan predznak i već u trećem nestaje. S druge strane, desni panel za drugo podrazoblje prikazuje da je učinak odmah u prvom mjesecu negativan, veoma je blag i pozitivan u drugom mjesecu i ponovno nestaje već u trećem mjesecu.

1.12. Analiza uzročnosti

Često se u radu s ekonomskim vremenskim nizovima razmatra uzročnost između varijabli, bez obzira radi li se o modelu samo s jednom zavisnom varijablom (npr. ARDL model) ili pak s više njih (vektorski modeli, poglavlje 2). Međutim, pojam uzročnosti se može na više načina protumačiti. Primjerice, ako se u svakodnevnom životu razmatraju dva događaja, A i B. Ako je događaj B rezultat događaja A, kažemo da A uzrokuje B. S druge strane, ako se razmatra uzročnost u okviru ekonometrijske analize, tada se uzročnost ili kauzalnost (engl. *causality, causation*) može definirati preko **Grangerove (1969) uzročnosti**. Radi se o jednome od nekoliko mogućih načina¹¹ definiranja i testiranja uzročnosti. Ako varijabla A uzrokuje varijablu B, tada uključivanje prethodnih vrijednosti varijable A u opisivanje dinamike varijable B doprinosi boljem opisivanju te dinamike. Drugim riječima, ako se dio kretanja varijable B u razdoblju t može bolje objasniti uključivanjem prethodnih vrijednosti varijable A, tada kažemo da varijabla A uzrokuje varijablu B u Grangerovom smislu.

Nadalje, u okviru Grangerove uzročnosti se kaže da ako događaj B nastupa nakon događaja A, onda B ne može nikako uzrokovati A. Međutim, ako događaj A nastupa prije događaja B, to nužno ne znači da A uzrokuje B. S time se povezuju pojmovi **endogene i egzogene varijable**. Egzozogena varijabla je ona koja je samo uzrok, a ne i posljedica u modelu, dok endogene varijable mogu biti i uzrok i posljedica. Dakle, ovdje se može samo govoriti o tome da je varijabla uzrokovana drugom varijablom u modelu, tada takva varijabla ne može biti egzozogena. S druge strane, ako se utvrdi da prva varijabla ne uzrokuje drugu, za drugu se ne može reći da je egzozogena.

U okviru testiranja Grangerove uzročnosti za slučaj dvije varijable, formira se poseban slučaj ARDL(p, q) modela, gdje se ne uključuje varijabla x u razdoblju t , već samo prethodne vrijednosti:

$$y_t = \alpha + \sum_{j=1}^p \beta_j y_{t-j} + \sum_{j=1}^q \gamma_j x_{t-j} + \varepsilon_t, \quad (1.213)$$

jer se ispituje prethodno rečeno da ako varijabla x uzrokuje varijablu y u Grangerovom smislu, tada prethodne vrijednosti varijable x bolje objašnjavaju kretanje sadašnje dinamike varijable y , u odnosu na samo uključivanje prethodne dinamike varijable y . Stoga se testiraju sljedeće hipoteze:

$$\begin{aligned} H_0 : x \text{ ne uzrokuje } y \text{ u Grangerovom smislu} \\ H_1 : x \text{ uzrokuje } y \text{ u Grangerovom smislu} \end{aligned}, \quad (1.214)$$

¹¹ Postoje i drugi načini razmatranja uzročnosti, poput DiD (engl. difference in differences), *matching*, itd. Vidjeti više u Abadie (2005), Austin (2011), Pearl (2000).

$$\begin{aligned} \text{odnosno} \quad & H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_q = 0 \\ & H_1 : \exists \gamma_j \neq 0, j \in \{1, 2, \dots, q\} \end{aligned} \quad (1.215)$$

Kako se u nultoj hipotezi provodi test na dio varijabli u modelu (1.213), radi se o parcijalnom F -testu (vidjeti detalje u odjeljku 2.2.7.4 u prvome dijelu). Ako zaključimo da odbacujemo nultu hipotezu da varijabla x ne uzrokuje varijablu y , kažemo da varijabla x uzrokuje varijablu y u Grangerovom smislu. To samo znači da prethodne vrijednosti varijable x poboljšavaju prognoziranje budućih vrijednosti varijable y , a ne znači nužno da postoji kauzalna (uzročna) veza između tih dviju varijabli.

Napomena: u modelu (1.213) se pretpostavlja da su varijable stacionarne. Ako su nestacionarne, potrebno ih je prevesti u stacionarni oblik prije provođenja testa. Nadalje, ako su varijable nestacionarne i ujedno kointegrirane (odjeljak 1.14), u tom slučaju treba razmatrati model korekcije pogreške (odjeljak 1.13).

Primjer 1.45.

Temeljem podataka iz primjera 1.43. provedimo testiranje Grangerove uzročnosti između stopa rasta cijena i stope rasta dolazaka turista u Španjolsku, a potom i Italiju. Za pomake p i q u modelu (1.213) odaberimo redom 1, 2 i 3.

Najprije je provedeno testiranje Grangerove uzročnosti za podatke za Španjolsku gdje je na slici 1.198. predočeno testiranje u kojem je uzrok varijabla stopa rasta cijena. Naredba `grangertest(...)`

$$\text{grangertest}(\text{diff}(\text{cijene}), \text{diff}(\text{dolasci}), 1)$$

se provodi na način da se kao prvi argument unutar naredbe piše varijabla y u modelu (1.213), kao drugi argument varijabla x (za koju se testira uzrokuje li y) i potom broj pomaka (=1, 2 ili 3 u primjeru). Uočava se da za pomak 1 je F -test veličina mala (1,304) i pripadajuća p -vrijednost je velika, veća je od uobičajenih razina značajnosti i ne može se odbaciti nulta hipoteza da stopa rasta cijena ne uzrokuje stopu rasta dolazaka turista u Španjolskoj pri pomaku 1. Pomoćne jednadžbe koje se uspoređuju su sljedeće:

$$y_t = \alpha + \beta_1 y_{t-1} + \gamma_1 x_{t-1} + \varepsilon_t \quad \text{i} \quad y_t = \alpha + \beta_1 y_{t-1} + \varepsilon_t,$$

odnosno u modelu $y_t = \alpha + \beta_1 y_{t-1} + \gamma_1 x_{t-1} + \varepsilon_t$ se u nultoj hipotezi testira $\gamma_1 = 0$.

U ispisu je prvi model s uključenom stopom rasta cijena u „Model 1“, dok je u „Model 2“ uključen samo pomak zavisne varijable, tj. stope rasta dolazaka turista.

Nadalje, za pomake 2 i 3, odgovarajući modeli u kojima se testira značajnost prethodnih pomaka stope rasta cijena su redom:

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \gamma_1 x_{t-1} + \gamma_2 x_{t-2} + \varepsilon_t \quad \text{za koji se testira} \quad \gamma_1 = \gamma_2 = 0 \quad \text{i}$$

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \gamma_1 x_{t-1} + \gamma_2 x_{t-2} + \gamma_3 x_{t-3} + \varepsilon_t \quad \text{za koji se testira} \quad \gamma_1 = \gamma_2 = \gamma_3 = 0.$$

Kako je p -vrijednost u slučaju $p = q = 2$ jednaka $0,56 > 0,05$, ne može se odbaciti hipoteza da stopa rasta cijena ne uzrokuje stopu rasta dolazaka turista u Španjolskoj pri pomaku 2. Za treći pomak je p -vrijednost jednaka $0,093 > 0,05$, te se i ovdje ne odbacuje hipoteza da stopa rasta cijena ne uzrokuje stopu rasta dolazaka turista u Španjolskoj pri pomaku 3.

```
library(lmtest)
grangertest(diff(cijene),diff(dolasci),1)

## Granger causality test
##
## Model 1: diff(dolasci) ~ Lags(diff(dolasci), 1:1) + Lags(diff(cijene), 1:1)
## Model 2: diff(dolasci) ~ Lags(diff(dolasci), 1:1)
##   Res.Df Df      F Pr(>F)
## 1      260
## 2      261 -1 1.3043 0.2545

grangertest(diff(cijene),diff(dolasci),2)

## Granger causality test
##
## Model 1: diff(dolasci) ~ Lags(diff(dolasci), 1:2) + Lags(diff(cijene), 1:2)
## Model 2: diff(dolasci) ~ Lags(diff(dolasci), 1:2)
##   Res.Df Df      F Pr(>F)
## 1      257
## 2      259 -2 0.5857 0.5575

grangertest(diff(cijene),diff(dolasci),3)

## Granger causality test
##
## Model 1: diff(dolasci) ~ Lags(diff(dolasci), 1:3) + Lags(diff(cijene), 1:3)
## Model 2: diff(dolasci) ~ Lags(diff(dolasci), 1:3)
##   Res.Df Df      F Pr(>F)
## 1      254
## 2      257 -3 2.1637 0.09278 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Slika 1.198. Testiranje Grangerove uzročnosti, uzrok = stopa rasta cijena, Španjolska

Ako se sad razmotri Grangerova uzročnost stope rasta dolazaka turista na stopu rasta cijena (slika 1.199.), uočava se da za sve pomake je pripadajuća p -vrijednost parcijalnog F -testa manja od uobičajenih razina značajnosti, stoga se odbacuje hipoteza da stopa rasta dolazaka turista ne uzrokuje u Grangerovom smislu stopu rasta cijena. Kada se pri razini značajnosti od 1% usporede ishodi na slici 1.198. i 1.199., zaključuje se da postoji jednosmjerna uzročnost od stope rasta cijena prema stopi rasta dolazaka turista. Može se reći da je stopa rasta dolazaka turista endogena varijabla.

Nadalje, ako se provedu isti testovi za slučaj Italije (slike 1.200. i 1.201.) te se razmotri razina značajnosti od 5% zaključci su sljedeći: za pomak 1 ne odbacuje hipoteza da stopa rasta cijena ne uzrokuje stopu rasta dolazaka turista i obratno. Pri pomaku 2 se ne odbacuje nulta hipoteza za slučaj cijena, ali se odbacuje za slučaj dolazaka. Konačno, pri pomaku 3 se odbacuje nulta hipoteza za obje varijable.

```
## Granger causality test
##
## Model 1: diff(cijene) ~ Lags(diff(cijene), 1:1) + Lags(diff(dolasci), 1:1)
## Model 2: diff(cijene) ~ Lags(diff(cijene), 1:1)
##   Res.Df Df      F Pr(>F)
## 1      260
```

```
## 2    261 -1 7.2763 0.007444 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(diff(dolasci),diff(cijene),2)

## Granger causality test
##
## Model 1: diff(cijene) ~ Lags(diff(cijene), 1:2) + Lags(diff(dolasci), 1:2)
## Model 2: diff(cijene) ~ Lags(diff(cijene), 1:2)
##   Res.Df Df       F    Pr(>F)
## 1      257
## 2      259 -2 10.904 2.849e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(diff(dolasci),diff(cijene),3)

## Granger causality test
##
## Model 1: diff(cijene) ~ Lags(diff(cijene), 1:3) + Lags(diff(dolasci), 1:3)
## Model 2: diff(cijene) ~ Lags(diff(cijene), 1:3)
##   Res.Df Df       F    Pr(>F)
## 1      254
## 2      257 -3 7.2523 0.0001096 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Slika 1.199. Testiranje Grangerove uzročnosti, uzrok = stopa rasta dolazaka turista, Španjolska

```
## Granger causality test
##
## Model 1: diff(dolasci) ~ Lags(diff(dolasci), 1:1) + Lags(diff(cijene), 1:1)
## Model 2: diff(dolasci) ~ Lags(diff(dolasci), 1:1)
##   Res.Df Df       F Pr(>F)
## 1      260
## 2      261 -1 1.3043 0.2545

grangertest(diff(cijene),diff(dolasci),2)

## Granger causality test
##
## Model 1: diff(dolasci) ~ Lags(diff(dolasci), 1:2) + Lags(diff(cijene), 1:2)
## Model 2: diff(dolasci) ~ Lags(diff(dolasci), 1:2)
##   Res.Df Df       F Pr(>F)
## 1      257
## 2      259 -2 0.5857 0.5575

grangertest(diff(cijene),diff(dolasci),3)

## Granger causality test
##
## Model 1: diff(dolasci) ~ Lags(diff(dolasci), 1:3) + Lags(diff(cijene), 1:3)
## Model 2: diff(dolasci) ~ Lags(diff(dolasci), 1:3)
##   Res.Df Df       F Pr(>F)
## 1      254
## 2      257 -3 2.1637 0.09278 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Slika 1.200. Testiranje Grangerove uzročnosti, uzrok = stopa rasta cijena, Italija

```

## Granger causality test
##
## Model 1: diff(cijene) ~ Lags(diff(cijene), 1:1) + Lags(diff(dolasci), 1:1)
## Model 2: diff(cijene) ~ Lags(diff(cijene), 1:1)
## Res.Df Df      F    Pr(>F)
## 1      260
## 2      261 -1 7.2763 0.007444 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(diff(dolasci),diff(cijene),2)

## Granger causality test
##
## Model 1: diff(cijene) ~ Lags(diff(cijene), 1:2) + Lags(diff(dolasci), 1:2)
## Model 2: diff(cijene) ~ Lags(diff(cijene), 1:2)
## Res.Df Df      F    Pr(>F)
## 1      257
## 2      259 -2 10.904 2.849e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(diff(dolasci),diff(cijene),3)

## Granger causality test
##
## Model 1: diff(cijene) ~ Lags(diff(cijene), 1:3) + Lags(diff(dolasci), 1:3)
## Model 2: diff(cijene) ~ Lags(diff(cijene), 1:3)
## Res.Df Df      F    Pr(>F)
## 1      254
## 2      257 -3 7.2523 0.0001096 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Slika 1.201. Testiranje Grangerove uzročnosti, uzrok = stopa rasta dolazaka turista, Italija

1.13. Kointegracija

Prije samog definiranja kointegriranosti dviju varijabli, razmotrimo sliku 1.202., gdje desni panel predočava dva niza koji imaju sličnu vremensku dinamiku kroz vrijeme, dok na lijevom panelu uočavamo kako se razlikuju dugoročne dinamike obiju pojava.

Uočimo da su na oba panela nizovi nestacionarni, no na lijevome panelu kao da sličan trend upravlja dinamikom oba niza, dok na desnome panelu ne postoji usklađenost kretanja kroz vrijeme. Fokusiranjem na lijevi panel, razlika između nizova koji imaju sličnu vremensku dinamiku tijekom vremena je podjednaka.

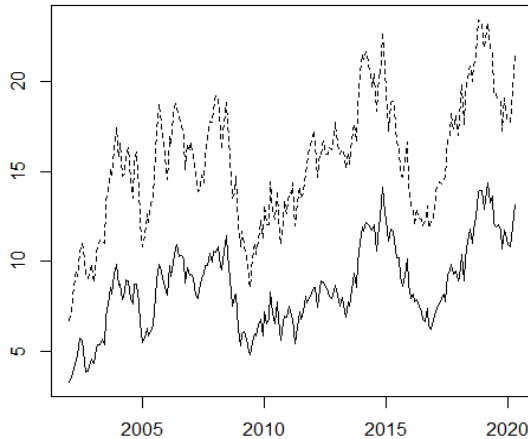
Dakle, ako se razmatraju x_t i y_t kao dva **nestacionarna niza** i neka je y_t linearna funkcija od x_t :

$$y_t = \alpha + \beta x_t + \varepsilon_t, \quad (1.216)$$

tada su odstupanja y_t od linearne funkcije $\alpha + \beta x_t$ jednaka $\varepsilon_t = y_t - \alpha - \beta x_t$, i upravo su ta **odstupanja podjednaka** ako je y_t linearna funkcija od x_t . To znači da procjenom modela (1.216) dobivamo rezidualne koji će biti stacionarni. U tom slučaju govorimo da su varijable x_t i y_t **kointegrirane**, u zapisu: $CI(1,1)$, čitamo kointegrirane prvog reda. U tom slučaju je model (1.216) dugog roka ili dugoročna jednadžba, dok se parametar β zove **kointegracijski parametar** (učinak u dugome roku). Prva jedinica u $CI(1,1)$ odnosi se na red integracije pojedine varijable čija se kointegriranost razmatra, dok se druga jedinica odnosi na vrijednost

koliko ćemo umanjiti red integriranosti odstupanja ε_t . Dakle, obje varijable x i y su integrirane prvog reda, te ako su kointegrirane, odstupanje će biti integrirano reda 1–1. Uočimo da obje varijable čija se kointegriranost ispituje moraju biti **istog reda integracije**.

(a) Slična dinamika nizova kroz vrijeme



(b) Razlika u dinamici između nizova



Slika 1.202. Vremenski nizovi

Engle i Granger (1987) prvi definiraju kointegriranost varijabli na sljedeći način. Razmatrali su skup ekonomskih varijabli i ako u dugome roku postoji dugoročna ravnoteža između njih, tada vrijedi:

$$\beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} = \beta x_t = 0, \quad (1.217)$$

gdje je β vektor parametara, a x_t vektor razmatranih varijabli. Odstupanje od dugog roka (engl. *equilibrium error*) definirali su kao:

$$\varepsilon_t = \beta x_t, \quad (1.218)$$

i zaključili da ako postoji ravnoteža dugog roka, odstupanje dugog roka mora biti stacionarno. Stoga su komponente vektora x_t definirali kao kointegriranimi reda d, b , u oznaci $x_t \sim CI(d, b)$ ako su sve komponente vektora integrirane reda d , te postoji vektor β takav da je linearna kombinacija $\beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk}$ integrirana reda $d-b$, $b > 0$. Vektor β se naziva kointegracijski vektor. Najčešće se u radu s ekonomskim varijablama radi o integriranosti pojedinih varijabli prvog reda, a rezultirajuće odstupanje dugog roka je stacionarno.

Ako se razmatraju dvije varijable kao na prethodnoj slici, potrebno je u prvome koraku testirati pomoću testa jediničnog korijena integriranost svake varijable. Ako se utvrdi da su obje integrirane prvog reda, slijedi procjena modela dugog roka u (1.216), te procjena reziduala iz tog modela. Potom se u zadnjem koraku provodi test jediničnog korijena o stacionarnosti reziduala. Ako su obje varijable integrirane prvog reda, dok su reziduali stacionarni, zaključuje se da su varijable kointegrirane.

Nadalje, Stock i Watson (1988) razmatraju kointegiranost sljedeća dva stohastička procesa: $y_t = y_{t-1} + e_t$ i $x_t = x_{t-1} + u_t$ (dva procesa slučajnog pomaka). Ako su ta dva procesa

kointegrirani reda (1,1), tada mora postojati linearna kombinacija $\beta_1 y_t + \beta_2 x_t$ koja je stacionarna, pri čemu su $\beta_1 \neq 0, \beta_2 \neq 0$. Dakle, ako se razmotri linearna kombinacija:

$$\begin{aligned} \beta_1 y_t + \beta_2 x_t &= \beta_1 (y_{t-1} + e_t) + \beta_2 (x_{t-1} + u_t) \\ &= \beta_1 y_{t-1} + \beta_2 x_{t-1} + \beta_1 e_t + \beta_2 u_t \end{aligned} \quad (1.219)$$

uočava se da $\beta_1 y_{t-1} + \beta_2 x_{t-1}$ mora biti jednako 0 kako bi linearna kombinacija bila stacionaran proces ($\beta_1 e_t + \beta_2 u_t$ je stacionaran proces jer je linearna kombinacija dva stacionarna, čista slučajna procesa). Stoga vrijedi:

$$\beta_1 y_{t-1} + \beta_2 x_{t-1} = 0 \Rightarrow y_{t-1} = -\frac{\beta_2}{\beta_1} x_{t-1}, \quad (1.220)$$

gdje se uočava da su stohastički trendovi varijabli y i x jednaki do na faktor $-\frac{\beta_2}{\beta_1}$.

Napomena: Hill i dr. (2018) napominju da se kod testiranja stacionarnosti reziduala iz modela (1.216) kritične granice razlikuju od tipičnih granica ADF testa, jer se ne razmatra slučajna varijabla, već njena procjena. Stoga se razmatraju granice prikazane u tablici 1.2. (Hamilton, 1994). U RStudiu se u okviru ispisa ADF testa ne navode spomenute granice, stoga treba paziti prilikom donošenja odluka o ishodu testa vezano uz rezidualne iz kointegracijske jednadžbe.

Tablica 1.2. Kritične granice testova kointegracije za rezidualne

Model	1%	5%	10%
$y_t = \beta x_t + \varepsilon_t$	-3,39	-2,76	-2,45
$y_t = \alpha + \beta x_t + \varepsilon_t$	-3,96	-3,37	-3,07
$y_t = \alpha + \beta x_t + \gamma t + \varepsilon_t$	-3,98	-3,42	-3,13

Za ostale kritične granice temeljem više od 2 varijable u modelu, vidjeti Hamilton (1994).

Primjer 1.46.

Razmatraju se tri niza: x, y_1 i y_2 . Provedimo testiranje kointegracije između x i y_1 te potom x i y_2 i komentirajmo rezultate temeljem ispisa na sljedećim slikama.

Najprije se provodi test za x i y_1 : potrebno je za svaku varijablu provesti ADF test u razinama, što je predočeno na slici 1.203. Uočava se da su pri razini značajnosti od 5% obje test veličine veće od kritične granice (-2,88), stoga se ne odbacuje nulta hipoteza da je niz x nestacionaran i da je niz y_1 nestacionaran. Potom su oba niza diferencirana i proveden je ADF test za obje varijable u prvim diferencijama na slici 1.204. Sada se uočava da pri razini značajnosti od 5% su obje test veličine manje od kritične granice (-1,95), stoga se odbacuje nulta hipoteza da je niz x nestacionaran i da je niz y_1 nestacionaran. Objе varijable su integrirane prvog reda. Potom je procijenjen model (1.216) i pripadajući reziduali. ADF test nad rezidualima je predočen na slici 1.205. Uočava se da pri razini značajnosti od 5% je test veličina manja od kritične granice (koja iznosi -3,37, prema tablici 1.2.), što znači da se odbacuje nulta hipoteza da je niz reziduala nestacionaran. Zaključujemo kako su varijable x i y_1 kointegrirane, jer su obje integrirane prvog reda, dok su reziduali integrirani jednog reda manje, tj. nultog reda, x i $y_1 \sim CI(1,1)$.

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.51942    0.19782   2.626 0.00927 **
## z.lag.1      -0.05494    0.02218  -2.477 0.01401 *
## z.diff.lag   -0.02076    0.06822  -0.304 0.76114
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -2.4772 3.4746
##
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
```

```
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.05113    0.36462   2.883 0.00434 **
## z.lag.1      -0.06327    0.02293  -2.759 0.00630 **
## z.diff.lag   -0.01944    0.06805  -0.286 0.77542
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -2.7588 4.1915
##
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

Slika 1.203. ADF testovi za x i y_1 , u razinama

```
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -1.12504    0.09942 -11.316 <2e-16 ***
## z.diff.lag    0.07860    0.06868   1.144  0.254
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -11.3164
##
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

```
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -1.08045    0.09976 -10.830 <2e-16 ***
## z.diff.lag    0.03470    0.06885   0.504  0.615
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -10.83
##
## Critical values for test statistics:
```

```
##      1pct 5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.204. ADF testovi za x i y_1 , u prvim diferencijama

```
rez1<-resid(lm(y1~x))
summary(ur.df(rez1,type="none",selectlags = "BIC"))

##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -0.10218    0.03054  -3.346 0.000967 ***
## z.diff.lag -0.01499    0.06801  -0.220 0.825742
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -3.3458
##
## Critical values for test statistics:
##      1pct 5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.205. ADF test za rezidualne iz modela (1.216), x i y_1

Za varijablu x smo prethodno zaključili da je integrirana prvog reda. Sada se još provode ADF testovi za varijablu y_2 na slici 1.206. Uočava se da se u razinama ne odbacuje nulta hipoteza o nestacionarnosti niza y_2 (gornji panel), dok se u prvim diferencijama odbacuje nulta hipoteza. Dakle, i varijabla y_2 je integrirana prvog reda. Procijenjena je kointegracijska jednadžba između x i y_2 , te je proveden ADF test za rezidualne iz novog modela, vidjeti sliku 1.207. Sada se ne odbacuje nulta hipoteza o nestacionarnosti reziduala u ovome slučaju, što vodi do zaključka da varijable x i y_2 nisu kointegrirane jer ne zadovoljavaju definiciju kointegriranosti.

```
summary(ur.df(y2,type="drift",selectlags="BIC"))

##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.430033    0.201066   2.139  0.0336 *
## z.lag.1      0.003331    0.003080   1.081  0.2807
## z.diff.lag  -0.088506    0.068653  -1.289  0.1987
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: 1.0815 10.8131
##
## Critical values for test statistics:
##      1pct 5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

summary(ur.df(diff(y2),type="none",selectlags="BIC"))

##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
```

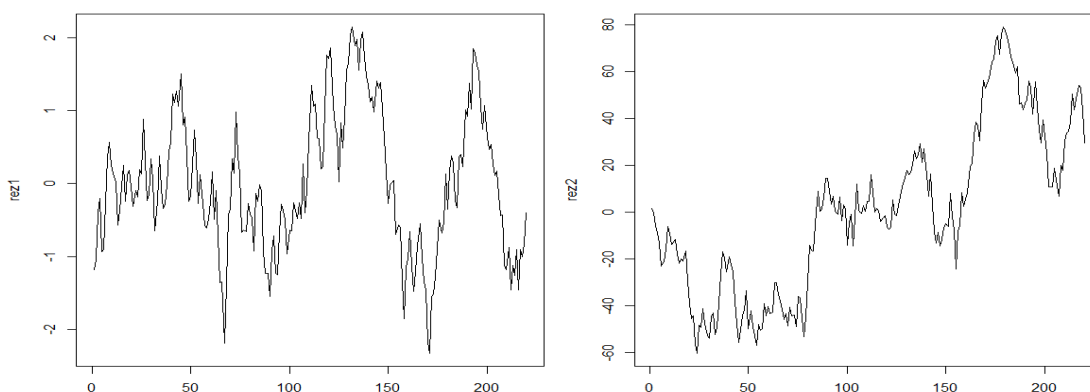


```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1   -0.90210    0.09577  -9.419  <2e-16 ***
## z.diff.lag -0.08791    0.06811  -1.291    0.198
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -9.419
##
## Critical values for test statistics:
##           1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.206. ADF test y_2 , u razinama (gornji panel) i prvim diferencijama (donji panel)

```
rez2<-resid(lm(y2~x))
summary(ur.df(rez2,type="none",selectlags = "BIC"))

##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.0290  -5.3107   0.0685   4.8270  16.1864
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1   -0.02032    0.01443  -1.408    0.160
## z.diff.lag -0.01703    0.06874  -0.248    0.805
##
## Value of test-statistic is: -1.4082
##
## Critical values for test statistics:
##           1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.207. ADF test za reziduale iz modela (1.227), x i y_2 Slika 1.208. Reziduali iz modela (1.227), lijevi panel: x i y_1 , desni panel x i y_2

Konačno, ako se i razmotre grafički prikazi oba reziduala iz kointegracijskih jednadžbi, uočava se da reziduali za x i y_1 (lijevi panel slike 1.208.) izgledaju kao stacionaran niz, dok reziduali za x i y_2 i dalje sadržavaju stohastički trend (desni panel). Čitatelju ostaje za vježbu da za sve provedene ADF testove zapiše odgovarajuće ADF jednadžbe, hipoteze testova i donese zaključke pri uobičajenim razinama značajnosti.

Primjer 1.47.

Temeljem podataka iz primjera 1.36. (paritet kupovne moći, PPP) gdje se upućuje na vezu između domaće razine cijena i tečaja. Jednostavan oblik testiranja PPP-a odnosi se na razmatranje razlike između domaće i strane stope inflacije. Neka su p_t i p_{ft} logaritmirane vrijednosti domaće razine cijena i strane razine cijena (engl. *foreign, f*), dok je e_t logaritam tečaja domaće valute i strane valute, PPP implicira da vrijedi:

$$e_t = p_t - p_{ft} + d_t,$$

gdje d_t predstavlja odstupanje od PPP-a. Prethodno se razmotrio realni tečaj $r_t = e_t + p_{ft} - p_t$ i ispitala se njegova (ne)stacionarnost. Ekonometrijsko testiranje PPP-a se sada može provesti na sljedeći način, pomoću testiranja kointegracije (vidjeti Enders, 2014). Sada ako vrijedi PPP, niz $e_t + p_{ft}$ treba biti kointegriran s p_t . Dakle, u prvome koraku ispituje se integraciju varijabli $e_t + p_{ft}$ i p_t , te potom procijenimo kointegracijsku jednadžbu između njih i za rezidualne tog modela provedemo testiranje nestacionarnosti. Reziduali trebaju biti stacionarni, te u kointegracijskoj jednadžbi parametar uz p_t treba biti jednak 1.

U tablici su predočeni rezultati svih ADF testova. Ono što se odmah uočava da se za varijable $e_t + p_{ft}$ i p_t već u razinama odbacuje nulta hipoteza ADF testa, stoga su te varijable stacionarne. Stoga se ne može razmatrati kointegracija između njih u dugome roku. Drugim riječima, PPP ne vrijedi za slučaj razmatranih podataka.

Varijable:	Test veličine	Kritične granice		
		1%	5%	10%
$e_t + p_{ft}$	-4,038	-3,46	-2,88	-2,57
p_t	-4.649			
$\Delta(e_t + p_{ft})$	-10,501	-2,58	-1,95	-1,62
Δp_t	-10.241			
$\hat{\varepsilon}_t$	-9.388	-4,35	-3,78	-3,47

1.14. Model korekcije pogreške

1.14.1. Definiranje modela korekcije pogreške

Engle i Granger (1987, 1991) popularizirali su model korekcije pogreške, za istovremeno razmatranje kratkoročne i dugoročne dinamike između dviju varijabli.

Ako se ponovno razmotri ARDL(1,1) model:

$$y_t = \alpha + \phi_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t, \quad (1.221)$$

te se od lijeve i desne strane jednakosti oduzmu y_{t-1} i $\beta_0 x_{t-1}$:

$$y_t - y_{t-1} - \beta_0 x_{t-1} = \alpha + \phi_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t - y_{t-1} - \beta_0 x_{t-1}, \quad (1.222)$$

$$\Delta y_t - \beta_0 x_{t-1} = \alpha + \phi_1 y_{t-1} - y_{t-1} + \beta_0 \Delta x_t + \beta_1 x_{t-1} + \varepsilon_t, \quad (1.223)$$

$$\Delta y_t = \alpha + \phi_1 y_{t-1} - y_{t-1} + \beta_0 \Delta x_t + \beta_0 x_{t-1} + \beta_1 x_{t-1} + \varepsilon_t, \quad (1.224)$$

$$\Delta y_t = \alpha + \underbrace{(\phi_1 - 1)}_{\lambda} y_{t-1} + \beta_0 \Delta x_t + \underbrace{(\beta_0 + \beta_1)}_{\pi} x_{t-1} + \varepsilon_t, \quad (1.225)$$

$$\Delta y_t = \alpha + \lambda y_{t-1} + \beta_0 \Delta x_t + \pi x_{t-1} + \varepsilon_t, \quad (1.226)$$

te se sada model (1.226) može zapisati u obliku:

$$\Delta y_t = \alpha + \beta_0 \Delta x_t + \lambda \left(y_{t-1} + \frac{\pi}{\lambda} x_{t-1} \right) + \varepsilon_t, \quad (1.227)$$

$$\Delta y_t = \alpha + \beta_0 \Delta x_t + (\phi_1 - 1) \left(y_{t-1} + \frac{\beta_0 + \beta_1}{\phi_1 - 1} x_{t-1} \right) + \varepsilon_t, \quad (1.228)$$

$$\Delta y_t = \alpha + \beta_0 \Delta x_t + (\phi_1 - 1) \left(y_{t-1} + \frac{\beta_0 + \beta_1}{\phi_1 - 1} x_{t-1} \right) + \varepsilon_t, \quad (1.229)$$

$$\Delta y_t = \beta_0 \Delta x_t - (1 - \phi_1) \left(\underbrace{y_{t-1} - \frac{\alpha}{1 - \phi_1} - \frac{\beta_0 + \beta_1}{1 - \phi_1} x_{t-1}}_{=\varepsilon_{t-1}} \right) + \varepsilon_t, \quad (1.230)$$

$$\Delta y_t = \underbrace{\beta_0}_{\text{kratkoročna dinamika}} \Delta x_t + \underbrace{\gamma \varepsilon_{t-1}}_{\text{dugoročna dinamika}} + \varepsilon_t, \quad (1.231)$$

te se model (1.230), odnosno (1.231) naziva **model korekcije pogreške** (engl. *error correction model*, ECM). Model korekcije pogreške razmatra kratkoročnu i dugoročnu dinamiku između varijabli x i y , gdje se učinak varijable x na y u **kratkome roku očituje u parametru β_0** , dok se **dugoročna dinamika** razmatra u članu ε_{t-1} u (1.231), koji je definiran u (1.230), te je učinak varijable x na y u dugom roku jednak $\frac{\beta_0 + \beta_1}{1 - \phi_1}$ (što je već izvedeno u (1.190)). Nadalje,

vrijednost γ se interpretira kao **odstupanje od ravnotežnog stanja**, jer se u zagradi u (1.230) razmatra odstupanje y_{t-1} od ravnotežnog stanja, tj. dugog roka; $\gamma = -(1 - \phi_1) = \phi_1 - 1$, što se može interpretirati i kao **brzina prilagodbe varijable y** .

Naziv modela, model korekcije pogreške stoga proizlazi iz analize kratkog i dugog roka istovremeno u (1.230), gdje se pretpostavlja da postoji dugoročna ravnoteža između varijabli u dugome roku i ako postoje kratkoročne neravnoteže, odstupanje od ravnotežnog stanja (prema dugome roku) se treba smanjivati, što će se očitovati u negativnom predznaku parametra γ , i zato se pogreška kratkog roka korigira prema ravnoteži dugog roka.

Model se procjenjuje u dva koraka. Najprije se procijeni model dugog roka (1.190), na način da se procjeni relacija $y_t = a + bx_t + \varepsilon_t$, te se procjene reziduali iz modela dugog roka: $\hat{\varepsilon}_t = y_t - \hat{a} - \hat{b}x_t$. Potom se u drugom koraku procijeni model (1.231), na način da se za procjenu ε_{t-1} koriste vrijednosti reziduala iz modela dugog roka, te se procijeni model: $\Delta y_t = \beta_0 \Delta x_t + \gamma \hat{\varepsilon}_{t-1} + \varepsilon_t$.

Primjer 1.48.

Temeljem podatka u primjeru 1.43, procijenimo model korekcije pogreške kojeg su predložili Engle i Granger između indeksa cijena i broja dolazaka turista u Španjolsku, te potom za podatke za Italiju te interpretirajmo sve rezultate. Najprije desezonirajmo vremenske nizove i potom ih logaritmirajmo prije same analize.

Sve naredbe potrebne za provedbu primjera prikazane su na slici 1.209., dok su vrijednosti parametara modela dugog roka i modela korekcije pogreške predočeni na slici 1.210. Na slici 1.209. prvih pet redaka naredbi se odnose na učitavanje podataka, definiranje vremenskih nizova i njihovo desezoniranje, nakon čega su varijable logaritmirane. Jednadžba dugog roka dana je naredbom `lm(...)`, te su potom procijenjeni reziduali definirani također kao vremenski niz. Model korekcije pogreške je u posljednjem retku definiran pomoću naredbe `dynlm(...)`.

```
ard12<-read.table("ard12.txt",sep="\t",header=T)
dolasci<-ts(ard12$S_dol,start=c(1998,1),frequency = 12)
dolasci<-dolasci/decompose(dolasci,type="multiplicative")$seasonal
cijene<-ts(ard12$S_hicp,start=c(1998,1),frequency = 12)
cijene<-cijene/decompose(cijene,type="multiplicative")$seasonal

dolasci<-log(dolasci)
cijene<-log(cijene)
#dugi rok:
lm(dolasci~cijene)
reziduali<-ts(resid(lm(dolasci~cijene)),start=c(1998,1),frequency = 12)

#korekcija:
dynlm(diff(dolasci)~diff(cijene)+L(reziduali))
```

Slika 1.209. Naredbe za procjenu modela korekcije pogreške

```
##
## Call:
## lm(formula = dolasci ~ cijene)
##
## Coefficients:
## (Intercept)      cijene
##      8.848         1.429
##
## Call:
## dynlm(formula = diff(dolasci) ~ diff(cijene) + L(reziduali))
##
## Coefficients:
## (Intercept)  diff(cijene)  L(reziduali)
## -0.002354    3.473550    -0.280365
```

Slika 1.210. Model dugog roka i model korekcije pogreške, Španjolska

Model dugog roka je $\hat{\mu}_y = 8,85 + 1,43\mu_x$, dok je model korekcije pogreške sljedeći: $\Delta\hat{y}_t = -0,002 + 3,47 \Delta x_t - 0,28\hat{\epsilon}_{t-1}$. U dugome roku povećanje indeksa cijena u Španjolskoj za 1% vodi do povećanja dolazaka turista za približno 1,43% (varijable su logaritmirane .), dok u kratkome roku je taj učinak jednak 3,47. Dakle, učinak je jači u kratkome roku. Nadalje, vrijednost parametra $\hat{\gamma}$ iznosi $-0,28$, te se interpretira na sljedeći način: u svakome mjesecu smanjenje odstupanja od ravnotežnog stanja iznosi 28%.

Za Italiju su procjene modela prikazane na slici 1.211. Model dugog roka je sada $\hat{\mu}_y = 10,817 + 0,998\mu_x$, dok je model korekcije pogreške: $\Delta\hat{y}_t = 4,7 \cdot 10^{-5} + 1,04 \Delta x_t - 0,066\hat{\epsilon}_{t-1}$. Sada

učinak povećanja indeksa cijena u Italiji za 1% u dugome roku vodi do povećanja dolazaka turista za približno 0,998%, a učinak u kratkome roku jednak je 1,04. Vrijednost parametra korekcije iznosi $-0,066$, što znači da u svakome mjesecu smanjenje odstupanja od ravnotežnog stanja iznosi 6,6%. Uočavamo da će se ravnoteža dugog roka ostvariti sporije za slučaj Italije u odnosu na Španjolsku.

```
##
## Call:
## lm(formula = dolasci ~ cijene)
##
## Coefficients:
## (Intercept)      cijene
##    10.8174         0.9977
##
## Call:
## dynlm(formula = diff(dolasci) ~ diff(cijene) + L(reziduali))
##
## Coefficients:
## (Intercept)  diff(cijene)  L(reziduali)
##  4.745e-05    1.036e+00   -6.561e-01
```

Slika 1.211. Model dugog roka i model korekcije pogreške, Italija

1.14.2. Prednosti i nedostaci pristupa modela korekcije pogreške

Prethodno opisani pristup ima određene prednosti i nedostatke. Prednost se očituje u tome što je postupak jednostavno provesti za slučaj dvije varijable. Nedostaci se očituju u tome što za slučaj više od 2 varijable istovremeno može postojati više jednadžbi dugoga roka (ako je broj varijabli koje se razmatraju jednak $m > 2$, tada može postojati $n \leq m-1$ jednadžbi dugog roka). U tom slučaju se ne može provoditi prethodno opisana procedura, već se razmatraju vektorski autoregresijski modeli i vektorski modeli korekcije pogreške (vidjeti poglavlja 2 i 3).

1.14.3. Napomene vezane uz model korekcije pogreške

Prije same procjene modela (1.231) potrebno je najprije testirati kointegriranost varijabli (vidjeti odjeljak 1.13.), odnosno procijeniti jednadžbu dugog roka, a potom ako se zaključi da su varijable kointegrirane, razmatramo model korekcije pogreške.

Uočimo da u modelu korekcije pogreške (1.231) razmatramo diferencije varijabli x i y , te rezidualne. Stoga bi u slučaju da su varijable x i y integrirane prvog reda, njihove prve diferencije, kao i reziduali trebali biti stacionarni pa se može regresijskom analizom razmatrati sam model (1.231) i provoditi t -test, kao i F -test.

Grangerov teorem reprezentacije sažima prethodno obrađivani model na sljedeći način. Razmotrimo skup varijabli koje su sve integrirane prvog reda. Ako je taj skup varijabli kointegriran, tada uvijek postoji model korekcije pogreške koji opisuje odnos između njih. Ako taj skup varijabli ima reprezentaciju u obliku modela korekcije pogreške, tada su kointegrirane.

1.15. Zaključno o (ne)stacionarnosti varijabli i odgovarajućim modelima

U ekonometrijskoj analizi potrebno je voditi računa o (ne)stacionarnosti varijabli koje se analiziraju, čemu se velik dio ovog poglavlja posvetio. Sumarno se može reći da (Enders, 2015):

1. Ako su razmatrani stohastički procesi u analizi stacionarni, klasičan pristup regresijske analize se može primijeniti i dalje analizirati.
2. Ako su razmatrani stohastički procesi različitih redova integracije, nema smisla raditi regresijsku analizu (vidjeti zaključke kod problema prividne regresije, odjeljak 1.10.6).
3. Ako su razmatrani stohastički procesi istog reda integracije, te su reziduali iz regresijskog modela u kojem se procjenjuje veza između stohastičkih procesa također nestacionarni, radi se o problemu prividne regresije. Tada se preporuča procjena regresijskog modela sa svim varijablama u prvim diferencijama.
4. Ako su razmatrani stohastički procesi istog reda integracije, te su reziduali iz regresijskog modela stacionarni, radi se o kointegriranosti varijabli, vidjeti odjeljak 1.13.

Hill i dr. (2018) upotpunjuju prethodno sa sljedećim: ako su varijable trend-stacionarne, tada se može procijeniti ARDL model u razinama, s uključenim trendom.

1.16. Primjer

Učitajmo datoteku „rh_dinamicki.txt“ u RStudio. Datoteka sadrži mjesečne podatke o kretanju ukupne nezaposlenosti u Hrvatskoj, te IIP (indeks industrijske proizvodnje, 2015 = 100). Desezonirajmo obje varijable pomoću multiplikativne metode dekomponiranja vremenskog niza, te potom izračunajmo stope rasta obiju varijabli, pomoću formule

$r_t = \ln\left(\frac{y_t}{y_{t-12}}\right)$. Neka je stopa rasta IIP-a varijabla x , a stopa rasta nezaposlenosti y . Početni datum je siječanj 2000. godine.

- a) Procijenimo ARDL(1,1) model između varijabli x i y te komentirajmo učinak x na y u kratkom i dugom roku, zapišimo jednadžbu dugog roka te prikažimo funkciju impulsnog odziva, kao i kumulativnu funkciju i komentirajmo dobiveno.

Nakon što su učitani podaci u RStudio, varijable desezonirane te definirane stope rasta (s.iip i s.nez), procijenjen je ARDL(1,1) model (slika 1.212.):

$$\hat{y}_t = 0,0003 + 0,955y_{t-1} - 0,114x_t - 0,068x_{t-1}.$$

Uočava se da je učinak stope rasta IIP-a na stopu rasta nezaposlenosti u kratkome roku jednak $-0,114$, dok je jednadžba dugog roka sljedeća:

$$\begin{aligned} \mu_y &= 0,0003 + 0,955\mu_y - 0,114\mu_x - 0,068\mu_x \\ \Rightarrow (1 - 0,955)\mu_y &= 0,0003 + (-0,114 - 0,068)\mu_x \Rightarrow \mu_y = 0,007 - 4,04\mu_x, \end{aligned}$$

pa je učinak u dugome roku veći i iznosi $-4,04$. Dakle, povećanje industrijske aktivnosti vodi i u kratkome i dugome roku do smanjenja nezaposlenosti.

```

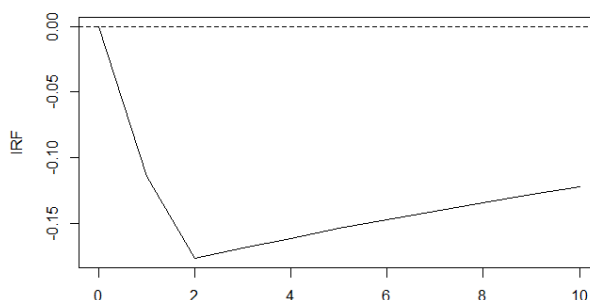
library(dynlm)
dynlm(s.nez~L(s.nez)+s.iip+L(s.iip))

##
## Time series regression with "ts" data:
## Start = 2001(2), End = 2020(9)
##
## Call:
## dynlm(formula = s.nez ~ L(s.nez) + s.iip + L(s.iip))
##
## Coefficients:
## (Intercept)      L(s.nez)          s.iip      L(s.iip)
##  0.0003051    0.9552513   -0.1141376   -0.0682544

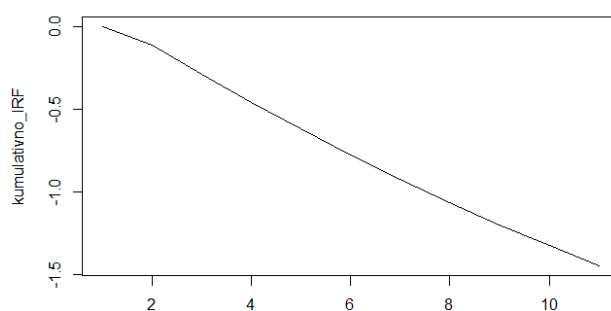
```

Slika 1.212. ARDL(1,1) model

Slike 1.213. i 1.214. predočavaju funkciju impulsnog odziva i kumulativnu funkciju. Uočava se da povećanje stope rasta IIP-a ima negativan učinak stopu rasta nezaposlenosti, pri čemu je učinak najjači u drugome mjesecu od šoka, i nakon toga se relativno sporo stopa rasta nezaposlenosti vraća na razinu prije samog šoka. Kumulativno gledajući, uočava se da se oko desetog mjeseca doseže učinak dugog roka.



Slika 1.213. Funkcija impulsnog odziva



Slika 1.214. Kumulativna funkcija impulsnog odziva

- b) Procijenimo ARDL(1,0) model između varijabli x i y te komentirajmo učinak x na y u kratkom i dugom roku, zapišimo jednadžbu dugog roka te prikazimo funkciju impulsnog odziva kao i kumulativnu funkciju i komentirajmo dobiveno.

Procijenjen je ARDL(1,0) model (slika 1.205.): $\hat{y}_t = 0,0003 + 0,963y_{t-1} - 0,145x_{t-1}$. Uočava se da je učinak stope rasta IIP-a na stopu rasta nezaposlenosti u kratkome roku jednak $-0,145$, dok je jednadžba dugog roka sljedeća:

$$\mu_y = 0,0003 + 0,963\mu_y - 0,145\mu_x$$

$$\Rightarrow (1 - 0,963)\mu_y = 0,0003 + 0,145\mu_x \Rightarrow \mu_y = 0,0088 - 3,919\mu_x,$$

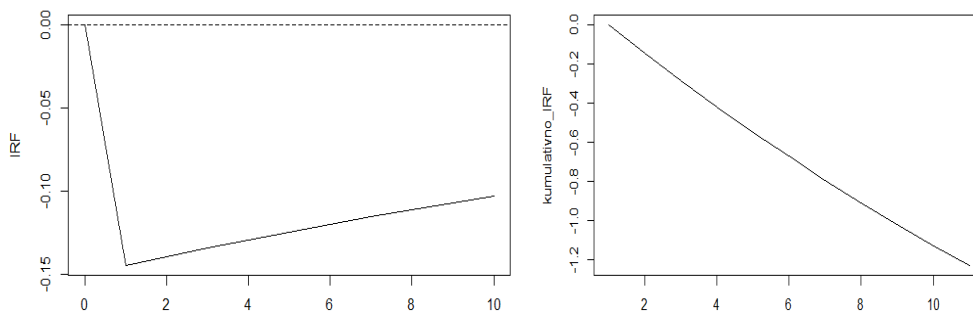
pa je učinak u dugome roku veći i iznosi $-3,919$. Dakle, povećanje industrijske aktivnosti vodi i u kratkome i dugome roku do smanjenja nezaposlenosti i u ovome modelu, pri čemu je učinak u dugome roku jači u odnosu na prethodni model. Slika 1.216. predočava obje funkcije: impulsnog odziva na lijevom panelu i uočava se slično ponašanje kao za prethodni model, no sad je najjači učinak u prvome mjesecu, dok kumulativna funkcija predočava smanjivanje tijekom vremena (desni panel) i uočava se da u desetom mjesecu kumulativni učinak iznosi oko $-1,2$, što znači da se još ne ostvaruje dugoročna ravnoteža. Potrebno je više vremena u ovome modelu da se ostvari ravnoteža dugog roka.

```

dynlm(s.nez~L(s.nez)+s.iip)

##
## Time series regression with "ts" data:
## Start = 2001(2), End = 2020(9)
##
## Call:
## dynlm(formula = s.nez ~ L(s.nez) + s.iip)
##
## Coefficients:
## (Intercept)      L(s.nez)          s.iip
##  0.000308      0.963276     -0.145049
    
```

Slika 1.215. ARDL(1,0) model



Slika 1.216. Funkcija impulsnog odziva (lijevi panel), kumulativna funkcija impulsnog odziva (desni panel)

- c) Procijenimo ARDL(0,1) model između varijabli x i y te komentirajmo učinak x na y u kratkom i dugom roku, zapišimo jednadžbu dugog roka te prikažimo funkciju impulsnog odziva kao i kumulativnu funkciju i komentirajmo dobiveno.

Sada je model (slika 1.217.) sljedeći: $\hat{y}_t = -0,036 - 0,798x_t - 0,927x_{t-1}$, te je učinak stope rasta IIP-a na stopu rasta nezaposlenosti u kratkome roku jednak $-0,798$, dok je jednadžba dugog roka $\mu_y = -0,036 - 0,798\mu_x - 0,927\mu_x \Rightarrow \mu_y = -0,036 - 1,725\mu_x$, i učinak u dugome roku jednak je $-1,725$. Ponovno je učinak u dugome roku veći od učinka u kratkome. Funkcija impulsnog odziva (lijevi panel slike 1.218.) prikazuje da je učinak povećanja stope rasta IIP-a na stopu raste nezaposlenosti negativan, najjači je u drugome mjesecu i već u trećem nestaje, dok je kumulativni učinak (desni panel slike 1.218.) dugog roka ostvaren već u trećem mjesecu.

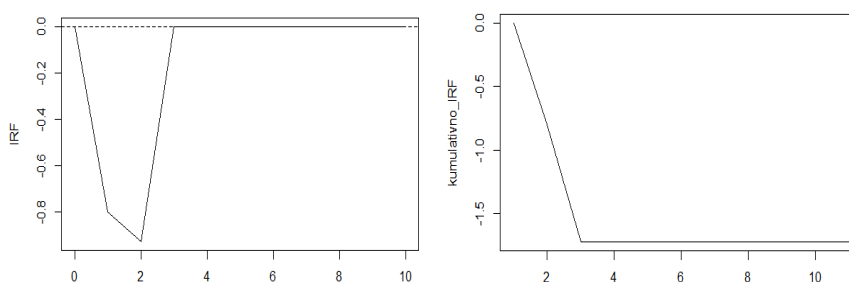

```

dynlm(s.nez~s.iip+L(s.iip))

##
## Time series regression with "ts" data:
## Start = 2001(2), End = 2020(9)
##
## Call:
## dynlm(formula = s.nez ~ s.iip + L(s.iip))
##
## Coefficients:
## (Intercept)      s.iip      L(s.iip)
##   -0.02623    -0.79815    -0.92707

```

Slika 1.217. ARDL(0,1) model



Slika 1.218. Funkcija impulsnog odziva (lijevi panel), kumulativna funkcija impulsnog odziva (desni panel)

- d) Procijenimo ARDL(1,0) za razmatrane varijable u prethodnim postupcima, ali za tri podrazdoblja: prvo od početka uzorka do prosinca 2007., za drugo od siječnja 2008. do prosinca 2015., te treće od siječnja 2016. do kraja razmatranog uzorka. Spojimo pomoću naredbe `stargazer(...)` rezultate, komentirajmo učinak varijable x na varijablu y u kratkom i dugom roku, te usporedimo funkcije impulsnog odziva za sva tri podrazdoblja.

Sada su na slici 1.219. predočeni modeli za sva tri podrazdoblja, dok su detaljni izračuni za svako podrazdoblje prikazani u tablici 1.3. Uočava se da je u trećem podrazdoblju učinak promjene stope rasta IIP-a na stopu rasta nezaposlenosti najjači i u kratkome i u dugome roku. Može se zaključiti da je razdoblje krize utjecalo na promjenu u vezi između dviju razmatranih varijabli, jer su učinci u kratkom i dugom roku prije i nakon krize značajno različitih vrijednosti. Nadalje, slika 1.220. uspoređuje funkcije impulsnog odziva za sva tri podrazdoblja i uočava se da povećanje stope rasta IIP-a u prvom podrazdoblju na stopu rasta nezaposlenosti utječe manje u odnosu na druga dva podrazdoblja, i stopa rasta nezaposlenosti se brže vraća na razinu prije povećanja stope rasta IIP-a. Veći učinak i dulje trajanje samog učinka je prisutno u drugom podrazdoblju, te najveće u posljednjem, za koje se s obzirom na rezultat dugog roka očekuje nadalje još trajanje šoka.

```

m1<-dynlm(s.nez~L(s.nez)+s.iip+L(s.iip),end=c(2017,12))
m2<-dynlm(s.nez~L(s.nez)+s.iip+L(s.iip),start=c(2008,1),end=c(2015,12))
m3<-dynlm(s.nez~L(s.nez)+s.iip+L(s.iip),start=c(2016,1))
library(stargazer)
stargazer(list(m1,m2,m3),type="text")

```

```

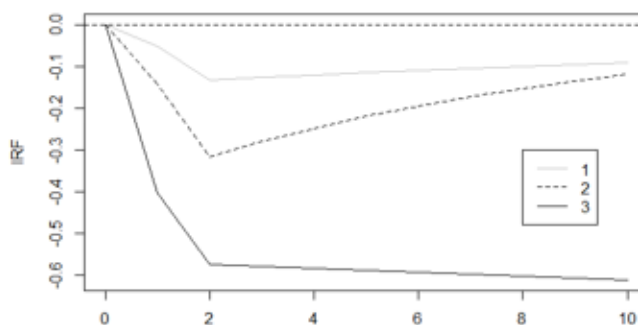
## =====
##                                     Dependent variable:
## -----
##                                     s.nez
##                                     (2)
## (1)                                     (3)
## -----
## L(s.nez)          0.953***          0.886***          1.008***
##                  (0.015)          (0.022)          (0.035)
##
## s.iip            -0.052*           -0.143***          -0.402***
##                  (0.031)          (0.043)          (0.109)
##
## L(s.iip)         -0.083***          -0.189***          -0.168
##                  (0.032)          (0.045)          (0.113)
##
## Constant         -0.001            -0.003            0.016**
##                  (0.001)          (0.002)          (0.008)
## -----
## Observations      203              96                57
## R2                0.979            0.981            0.949
## Adjusted R2       0.979            0.980            0.947
## Residual Std. Error 0.018 (df = 199)    0.017 (df = 92)    0.036 (df = 53)
## F Statistic       3,088.425*** (df = 3; 199) 1,543.069*** (df = 3; 92) 331.628*** (df = 3; 53)
## =====
## Note:                                     *p<0.1; **p<0.05; ***p<0.01

```

Slika 1.219. ARDL(1,1) model za tri podrazdoblja

Tablica 1.3. Procijenjeni modeli sa slike 1.208.

Razdoblje:	ARDL(1,1) model:	Učinak u kratkome roku:
Prvo:	$\hat{y}_t = -0,001 + 0,953y_{t-1} - 0,052x_t - 0,008x_{t-1}$	-0,052
Drugo:	$\hat{y}_t = -0,003 + 0,886y_{t-1} - 0,143x_t - 0,189x_{t-1}$	-0,143
Treće:	$\hat{y}_t = 0,016 + 1,008y_{t-1} - 0,402x_t - 0,168x_{t-1}$	-0,402
Razdoblje:	Model dugog roka:	Učinak u dugome roku:
Prvo:	$\mu_y = -0,0213 - 1,28\mu_x$	-1,28
Drugo:	$\mu_y = -0,025 - 2,77\mu_x$	-2,77
Treće:	$\mu_y = -2 - 71,25\mu_x$	-71,25



Slika 1.220. Funkcije impulsnog odziva za tri podrazdoblja

- e) Provedimo testiranje Grangerove uzročnosti između varijabli x i y . Za pomake p i q u modelu (1.213) odaberimo redom 1, 2 i 3.

Sada se testiraju sljedeće hipoteze, predočene u tablici 1.4., temeljem rezultata na slikama 1.221. i 1.222. Uz razinu značajnosti od 5%, odbacuje se hipoteza da stopa rasta IIP-a ne uzrokuje stopu rasta nezaposlenosti pri prvom i trećem pomaku, dok se ne odbacuje pri drugom pomaku. S druge strane, uz istu razinu značajnosti od 5%, odbacuje se hipoteza da stopa rasta nezaposlenosti ne uzrokuje stopu rasta IIP-a pri sva tri pomaka.

Tablica 1.4. Rezultati testova Grangerove uzročnosti

Pomak /	Uzrok:	Stopa rasta IIP-a	Stopa rasta nezaposlenosti
Prvi	$H_0:$	$\gamma_1 = 0$	
	Test veličina	12,744	17,231
	p -v	0,0004	≈ 0
Drugi	$H_0:$	$\gamma_1 = \gamma_2 = 0$	
	Test veličina	2,861	4,786
	p -v	0,0593	0,009
Treći	$H_0:$	$\gamma_1 = \gamma_2 = \gamma_3 = 0$	
	Test veličina	3,052	3,804
	p -v	0,0294	0,011

```
grangertest(s.iip,s.nez,1)

## Granger causality test
##
## Model 1: s.nez ~ Lags(s.nez, 1:1) + Lags(s.iip, 1:1)
## Model 2: s.nez ~ Lags(s.nez, 1:1)
##   Res.Df Df      F    Pr(>F)
## 1      233
## 2      234 -1 12.744 0.0004337 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(s.iip,s.nez,2)

## Granger causality test
##
## Model 1: s.nez ~ Lags(s.nez, 1:2) + Lags(s.iip, 1:2)
## Model 2: s.nez ~ Lags(s.nez, 1:2)
##   Res.Df Df      F Pr(>F)
## 1      230
## 2      232 -2 2.8606 0.05928 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(s.iip,s.nez,3)

## Granger causality test
##
## Model 1: s.nez ~ Lags(s.nez, 1:3) + Lags(s.iip, 1:3)
## Model 2: s.nez ~ Lags(s.nez, 1:3)
##   Res.Df Df      F Pr(>F)
## 1      227
## 2      230 -3 3.0519 0.02935 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Slika 1.221. Grangerov test uzročnosti, uzrok = stopa rasta IIP-a

```

grangertest(s.nez,s.iip,1)

## Granger causality test
##
## Model 1: s.iip ~ Lags(s.iip, 1:1) + Lags(s.nez, 1:1)
## Model 2: s.iip ~ Lags(s.iip, 1:1)
##   Res.Df Df       F    Pr(>F)
## 1      233
## 2      234 -1 17.241 4.622e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(s.nez,s.iip,2)

## Granger causality test
##
## Model 1: s.iip ~ Lags(s.iip, 1:2) + Lags(s.nez, 1:2)
## Model 2: s.iip ~ Lags(s.iip, 1:2)
##   Res.Df Df       F    Pr(>F)
## 1      230
## 2      232 -2 4.7863 0.009192 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(s.nez,s.iip,3)

## Granger causality test
##
## Model 1: s.iip ~ Lags(s.iip, 1:3) + Lags(s.nez, 1:3)
## Model 2: s.iip ~ Lags(s.iip, 1:3)
##   Res.Df Df       F    Pr(>F)
## 1      227
## 2      230 -3 3.8042 0.01089 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Slika 1.222. Grangerov test uzročnosti, uzrok = stopa rasta nezaposlenosti

- f) Procijenimo model korekcije pogreške kojeg su predložili Engle i Granger između desezoniranih vrijednosti nezaposlenosti u Hrvatskoj i IIP-a za kointegracijsku jednadžbu (dakle, varijable u razinama, ne stope rasta). Pretpostavlja se da je zavisna varijabla nezaposlenost. Interpretirajmo rezultate.

```

lm(nez~iip)

## Call:
## lm(formula = nez ~ iip)
##
## Coefficients:
## (Intercept)          iip
##   581.963         -3.206

rez<-resid(lm(nez~iip))
rez<-ts(rez,start=c(2001,1),frequency = 12)
library(dynlm)
dynlm(diff(nez)~diff(iip)+L(rez))

## Time series regression with "ts" data:
## Start = 2001(2), End = 2020(9)
##
## Call:
## dynlm(formula = diff(nez) ~ diff(iip) + L(rez))
##
## Coefficients:

```

## (Intercept)	diff(iip)	L(rez)
## -0.48103	0.02880	-0.01843

Slika 1.223. Model korekcije pogreške

Engle i Granger su predložili procjenu modela dugog roka za varijable u razinama, stoga je temeljem slike 1.223. model dugog roka sljedeći: $\mu_y = 581963 - 3,206\mu_x$, dok je model korekcije pogreške: $\Delta\hat{y}_t = -0,4813 + 0,0289 \Delta x_t - 0,018\hat{\epsilon}_{t-1}$. Učinak u kratkome roku u ovome modelu je pozitivan i iznosi 0,0289, dok je u dugome roku negativan (-3,206). Parametar odstupanja od ravnotežnog stanja iznosi -0,018 i govori da se svakog mjeseca se 1,8% kratkoročne neravnoteže korigira prema ravnoteži dugog roka (ili, u svakome mjesecu smanjenje odstupanja od ravnotežnog stanja iznosi 1,8%). Zato je potrebno dulje vrijeme za postizanje ravnoteže dugog roka.

- g) Provedimo testiranje kointegracije između razmatranih varijabli u razinama i interpretirajmo rezultate.

Čitatelju ostaje za vježbu pisanje svih ADF jednadžbi sa slika 1.124., 1.125. i 1.126. Interpretacije su sljedeće. Kada se provodi ADF test za varijablu IIP i nezaposlenost u razinama (slika 1.124.), uočava se da su odgovarajuće test veličine jednake -2,47 i -0,93, te pri uobičajenim razinama značajnosti su obje test veličine veće od kritičnih granica pa se ne odbacuje nulta hipoteza da je varijabla IIP nestacionarna i da je varijabla nezaposlenost nestacionarna.

Nadalje, kada se provode ADF testovi za varijable u prvim diferencijama (slika 1.225.), uočava se da su test veličine redom za varijable IIP i nezaposlenost u diferencijama jednake -19,29 i -7,31, što je pri uobičajenim razinama značajnosti manje od kritičnih granica testa. Stoga se za obje varijable odbacuje nulta hipoteza da su nestacionarne u prvim diferencijama. Kako je bilo potrebno jedanput diferencirati svaku varijablu da bi postala stacionarna, zaključujemo kako su varijable, IIP i nezaposlenost, integrirane prvog reda.

```
summary(ur.df(iip,type="drift",selectlags = "BIC"))
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.84518    2.74211   2.496  0.0132 *
## z.lag.1      -0.06232    0.02523  -2.470  0.0142 *
## z.diff.lag   -0.46201    0.05667  -8.153 1.87e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -2.4701 3.1271
##
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
```

```
summary(ur.df(nez,type="drift",selectlags = "BIC"))
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Coefficients:
```

```
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.747693  1.033493  0.723  0.470
## z.lag.1     -0.003937  0.004245 -0.927  0.355
## z.diff.lag  0.636595  0.049814 12.779 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -0.9274 0.6912
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
```

Slika 1.224. ADF testovi varijabli u razinama

```
summary(ur.df(diff(iip),type="none",selectlags = "BIC"))
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -2.00417   0.10388 -19.293 < 2e-16 ***
## z.diff.lag  0.34574   0.06064  5.701 3.42e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -19.293
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

```
summary(ur.df(diff(nez),type="none",selectlags = "BIC"))
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -0.39687   0.05428 -7.311 3.78e-12 ***
## z.diff.lag  0.08285   0.06356  1.304  0.194
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -7.3111
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.225. ADF testovi varijabli u prvim diferencijama

Nadalje, procijenjen je model dugog roka u kojemu je varijabla nezaposlenost zavisna, a varijabla IIP nezavisna i spremljeni su reziduali iz tog modela, nad kojima je proveden ADF test nestacionarnosti (slika 1.226.). Uočava se da je test veličina ovog testa jednaka $-1,399$, što je veće od uobičajenih kritičnih granica iz tablice 1.2 pa se ne može odbaciti hipoteza da je niz reziduala nestacionaran. Kako je ovaj niz nestacionaran, zaključuje se kako varijable IIP i nezaposlenost nisu kointegrirane u dugome roku, jer njihova linearna kombinacija (reziduali) nije stacionarna, tj. integrirana jednog reda manje.

```
rez2<-resid(lm(nez~iip))
summary(ur.df(rez2,type="none",selectlags = "BIC"))

##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -0.02860    0.02044  -1.399   0.163
## z.diff.lag -0.35752    0.06030  -5.929 1.03e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Value of test-statistic is: -1.3992
##
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Slika 1.226. ADF test nad rezidualima kointegracijske jednadžbe

1.17. Informacijski kriteriji

S obzirom da je u praksi često teško odrediti o kakvom $ARMA(p,q)$ procesu se radi temeljem korelograma ACF i PACF funkcija, postavlja se pitanje postoji li kakav formalan način usporedbe procijenjenih modela kojima opisujemo ponašanje vremenskih nizova. Jedan način je usporediti optimalnu vrijednost funkcije vjerodostojnosti za svaki procijenjen model (vidjeti prvi dio u ovoj seriji publikacija), ali kako se radi o modelima koji se razlikuju i po broju procijenjenih parametara, nastoje se koristiti određene mjere koje uvažavaju i tu činjenicu. Zato se često u praksi koriste informacijski kriteriji. **Informacijski kriteriji** su mjere temeljene na kompromisu između modelovog dobrog opisivanja dinamike vremenskih nizova (engl. *goodnes of fit*) i očuvanja dovoljnog broja stupnjeva slobode u modelu (parsimonija).

Tri su osnovna informacijska kriterija koji se najčešće koriste u praksi: *AIC* (*Akaike information criterion*), *SIC* (*Schwartz criterion*) te *HQC* (*Hannan-Quinn information criterion*):

$$AIC = -2 \log L / T + 2k/T, \quad (1.232)$$

$$SIC = -2 \log L / T + (k \cdot \log T) / T \quad (1.233)$$

i

$$HQC = -2 \log L / T + 2k \log(\log T) / T, \quad (1.232)$$

gdje je $\log L$ optimalna vrijednost funkcije cilja (funkcija vjerodostojnosti), T je broj opažanja temeljem kojih se procjenjuje model, a k je broj parametara u modelu koji se procjenjuju. Kako se u slučaju optimizacije funkcije vjerodostojnosti radi o njenom maksimizaciji, veća vrijednost $\log L$ je bolja, no zbog množenja s vrijednošću -2 , bolji model koji ima veću optimalnu vrijednost funkcije cilja imat će što manju vrijednost $-2 \log L$. Kako bi se penalizirao prevelik broj nepotrebno procijenjenih parametara, veća vrijednost k u modelu znači veću vrijednost informacijskog kriterija, jer je predznak u drugom dijelu izračuna svakog kriterija, vezan uz k , pozitivnog predznaka. Dakle, "nagrađuje" se optimalna vrijednost funkcije cilja kao prvi pribrojnik, dok se penalizira složenost modela kao drugi pribrojnik. **Zato je poželjno da je vrijednost informacijskih kriterija što manja.** Detalji o ovim mjerama mogu se vidjeti u Tsay (2010).

Primjer 1.49.

Učitajmo datoteku „**gen_nizovi.txt**“ u RStudio. Datoteka sadrži 200 opservacija za 7 različitih procesa (Y1 do Y7 nazivi). Definirajmo za Y1 niz da se radi o vremenskom nizu, s početnom vrijednošću u rujnu 2003. godine, podaci su mjesečni te potom procijenimo nekoliko različitih ARMA modela: AR(1), AR(2), MA(1), MA(2), ARMA(1,1) i usporedimo sve modele temeljem AIC, SIC i HQC informacijskih kriterija.

Na slici 1.227. prikazan je postupak učitavanja podataka, procjene svih modela, te naredbe za izračun sva tri informacijska kriterija za svaki model, i konačno spajanje u jednu tablicu za lakšu usporedbu. Potrebno je procijeniti odgovarajuće modele naredbom `arima(...)` i spremiti ih (nazivi `ar_1` do `arma`). Za spremljeni model se pomoću naredbi AIC, BIC i HQIC računaju sva tri informacijska kriterija na način da je argument svake naredbe naziv modela za koji se računa. Primjerice `BIC(ma_1)` je naredba za izračun BIC informacijskog kriterija za model MA(1). Nakon toga su sve vrijednosti za pojedine modele spremljene u tri kategorije: `aic`, `bic` i `hqc`, temeljem kojih je definirana tablica, dodani su nazivi retcima naredbom `row.names(tablica)`, pri čemu nazive pojedinog retka obavezno pišemo pod navodnike. Naredba tablica na samom dnu slike 1.227. omogućava lakšu usporedbu svih modela. Uočava se da su MA modeli veoma loši jer su njihovi informacijski kriteriji puno veći od preostalih modela. Između preostala tri (AR(1), AR(2) i ARMA(1,1)), najmanje vrijednosti svih kriterija je upravo za AR(1) pa bi daljnja analiza bila temeljena upravo na tom modelu.

```
nizovi<-read.table("gen_nizovi.txt",header=T,sep="\t")
y<-ts(nizovi$Y1,start=c(2003,9),frequency=12)

library("ICglm")

ar_1<-arima(y,order=c(1,0,0))
ar_2<-arima(y,order=c(2,0,0))
ma_1<-arima(y,order=c(0,0,1))
ma_2<-arima(y,order=c(0,0,2))
arma<-arima(y,order=c(1,0,1))

aic_ar_1<-AIC(ar_1);bic_ar_1<-BIC(ar_1);hqc_ar_1<-HQIC(ar_1)
aic_ar_2<-AIC(ar_2);bic_ar_2<-BIC(ar_2);hqc_ar_2<-HQIC(ar_2)
aic_ma_1<-AIC(ma_1);bic_ma_1<-BIC(ma_1);hqc_ma_1<-HQIC(ma_1)
aic_ma_2<-AIC(ma_2);bic_ma_2<-BIC(ma_2);hqc_ma_2<-HQIC(ma_2)
aic_arma<-AIC(arma);bic_arma<-BIC(arma);hqc_arma<-HQIC(arma)

aic<-c(aic_ar_1,aic_ar_2,aic_ma_1,aic_ma_2,aic_arma)
bic<-c(bic_ar_1,bic_ar_2,bic_ma_1,bic_ma_2,bic_arma)
hqc<-c(hqc_ar_1,hqc_ar_2,hqc_ma_1,hqc_ma_2,hqc_arma)

tablica<-cbind(aic,bic,hqc)
row.names(tablica)<-c("ar_1","ar_2","ma_1","ma_2","arma")
tablica
```

Slika 1.227. Učitavanje podataka, procjena modela i izračun informacijskih kriterija

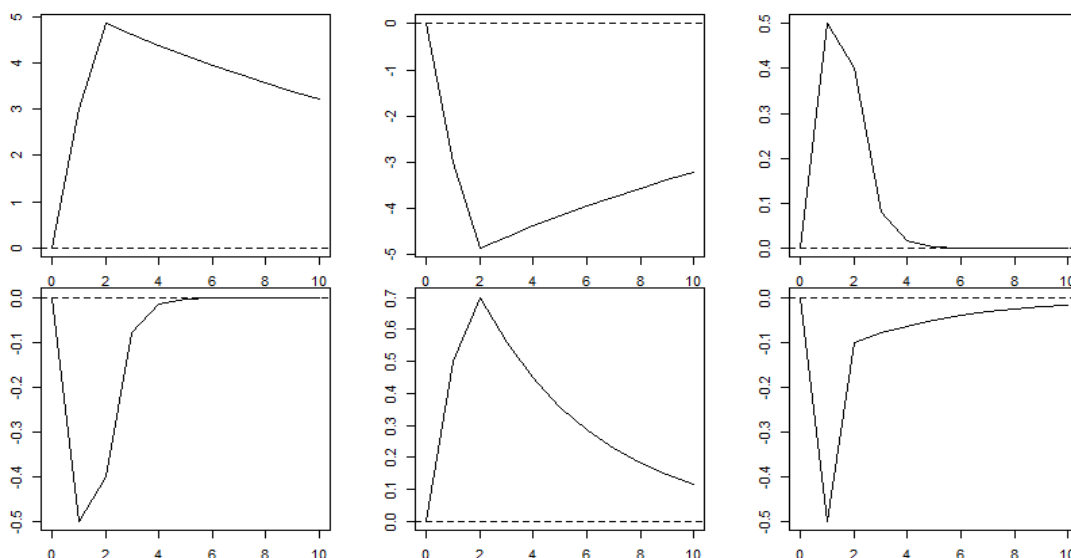
```
tablica
##          aic          bic          hqc
## ar_1  69.18479  79.07975  73.18913
## ar_2  71.06863  84.26190  76.40774
## ma_1 287.78625 297.68120 291.79058
## ma_2 207.58605 220.77932 212.92517
## arma  71.08860  84.28187  76.42772
```

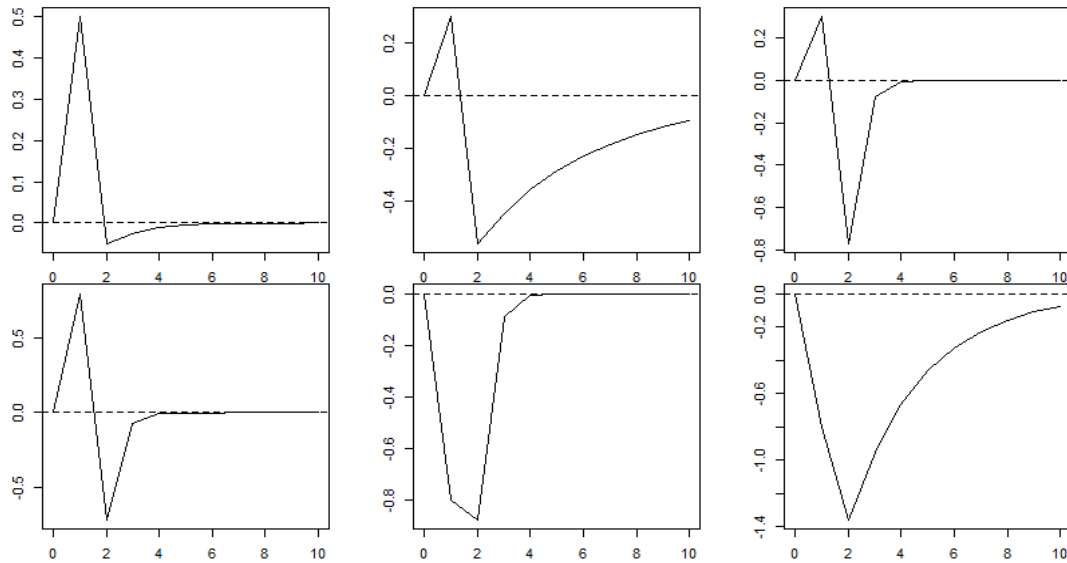
Slika 1.228. Prikaz vrijednosti informacijskih kriterija za sve modele

1.18. Pitanja za ponavljanje

- Što je to autoregresijski model distribuiranih pomaka?
- Zapišite model s beskonačnim brojem pomaka. Koja mu je očekivana vrijednost?
- Koliki je učinak nezavisne na zavisnu varijablu u dugome roku u modelu s beskonačnim brojem pomaka?
- Koliko iznosi učinak nezavisne na zavisnu varijablu u kratkome roku u prethodno navedenom modelu?
- Zapišite ARDL model s konačnim brojem pomaka. Koliki je učinak nezavisne na zavisnu varijablu u kratkome, a koliki u dugome roku u tom modelu?
- Zapišite ARDL(1,1) model te odredite učinak nezavisne na zavisnu varijablu u kratkome i dugome roku i zapišite jednadžbu dugog roka.
- Zapišite ARDL(1,0) model te odredite učinak nezavisne na zavisnu varijablu u kratkome i dugome roku i zapišite jednadžbu dugog roka.
- Zapišite ARDL(0,1) model te odredite učinak nezavisne na zavisnu varijablu u kratkome i dugome roku i zapišite jednadžbu dugog roka.
- Za sljedeće ARDL modele zapišite učinak u kratkome i dugome roku:

a) $y_t = 2 + 0,8y_{t-1} + 0,8x_t - 0,2x_{t-1} + \varepsilon_t$	f) $y_t = 2 - 0,8x_t - 0,2x_{t-1} + \varepsilon_t$
b) $y_t = 2 - 0,8y_{t-1} - 0,8x_t + 0,2x_{t-1} + \varepsilon_t$	g) $y_t = 2 - 0,8x_t + 0,2x_{t-1} + \varepsilon_t$
c) $y_t = 2 + 0,8y_{t-1} - 0,8x_t + 0,2x_{t-1} + \varepsilon_t$	h) $y_t = 2 + 0,8x_t + 0,2x_{t-1} + \varepsilon_t$
d) $y_t = 2 - 0,8y_{t-1} + 0,8x_t - 0,2x_{t-1} + \varepsilon_t$	i) $y_t = 2 + 0,8y_{t-1} + 0,8x_t + \varepsilon_t$
e) $y_t = 2 + 0,8x_t - 0,2x_{t-1} + \varepsilon_t$	j) $y_t = 2 + 0,8y_{t-1} - 0,8x_t + \varepsilon_t$
- Zapišite osnovnu jednadžbu modela korekcije pogreške kojeg su definirali Engle i Granger (1987, 1991). Interpretirajte parametre tog modela. Opišite postupak procjene modela u dva koraka.
- Definirajte što znači da su dvije varijable kointegrirane. Opišite postupak testiranja kointegracije.
- Za sljedeće funkcije impulsnog odziva odabranih ARDL modela komentirajte učinak nezavisne na zavisnu varijablu:





13. Učitajte u RStudio datoteku „**it_dinamicki.txt**“ u RStudio. Datoteka sadrži mjesečne podatke o kretanju ukupne nezaposlenosti u Italiji, te IIP (indeks industrijske proizvodnje, 2015 = 100). Desezonirajte obje varijable pomoću multiplikativne metode dekomponiranja vremenskog niza, te potom izračunajte stope rasta obiju varijabli, pomoću formule $r_t = \ln\left(\frac{y_t}{y_{t-12}}\right)$. Neka je stopa rasta IIP-a varijabla y , a stopa rasta nezaposlenosti x . Početni datum je siječanj 1990. godine.

- Procijenite ARDL(1,1) model između varijabli x i y te komentirajte učinak x na y u kratkom i dugom roku, zapišite jednadžbu dugog roka te prikažite funkciju impulsnog odziva, kao i kumulativnu funkciju i komentirajte dobiveno.
- Procijenite ARDL(1,0) model između varijabli x i y te komentirajte učinak x na y u kratkom i dugom roku, zapišite jednadžbu dugog roka te prikažimo funkciju impulsnog odziva kao i kumulativnu funkciju i komentirajte dobiveno.
- Procijenite ARDL(0,1) model između varijabli x i y te komentirajte učinak x na y u kratkom i dugom roku, zapišite jednadžbu dugog roka te prikažite funkciju impulsnog odziva kao i kumulativnu funkciju i komentirajte dobiveno.
- Procijenite ARDL(1,1) za razmatrane varijable u prethodnim postupcima, ali za dva podrazdoblja: prvo od početka uzorka do prosinca 2007., za drugo od siječnja 2008. do kraja razmatranog uzorka. Spojite pomoću naredbe `stargazer(...)` rezultate, komentirajmo učinak varijable x na varijablu y u kratkom i dugom roku, te usporedite funkcije impulsnog odziva za oba podrazdoblja.
- Provedite testiranje Grangerove uzročnosti između varijabli x i y . Za pomake p i q u modelu (1.213) odaberite redom 1, 2 i 3.
- Procijenite model korekcije pogreške kojeg su predložili Engle i Granger između desezoniranih vrijednosti nezaposlenosti u Italiji i IIP-a (dakle, varijable u razinama, ne stope rasta). Pretpostavlja se da je zavisna varijabla IIP. Interpretirajte rezultate.
- Provedite testiranje kointegracije između razmatranih varijabli u razinama i interpretirajte rezultate.

14. Što su informacijski kriteriji i čemu služe?

Rješenja**Zadatak 9)**

$$a) \quad y_t = 2 + 0,8y_{t-1} + 0,8x_t - 0,2x_{t-1} + \varepsilon_t$$

Učinak u kratkom roku: 0,8

Učinak u dugom roku: $y - 0,8y = 2 + 0,8x - 0,2x$, $y = 10 + 0,6x$, iznosi 3

$$b) \quad y_t = 2 - 0,8y_{t-1} - 0,8x_t + 0,2x_{t-1} + \varepsilon_t$$

Učinak u kratkom roku: -0,8

Učinak u dugom roku: $y + 0,8y = 2 - 0,8x + 0,2x$, $y = 1,11 - 0,33x$, iznosi -0,33

$$c) \quad y_t = 2 + 0,8y_{t-1} - 0,8x_t + 0,2x_{t-1} + \varepsilon_t$$

Učinak u kratkom roku: -0,8

Učinak u dugom roku: $y - 0,8y = 2 - 0,8x + 0,2x$, $y = 10 - 3x$, iznosi -3

$$d) \quad y_t = 2 - 0,8y_{t-1} + 0,8x_t - 0,2x_{t-1} + \varepsilon_t$$

Učinak u kratkom roku: 0,8

Učinak u dugom roku: $y + 0,8y = 2 + 0,8x - 0,2x$, $y = 1,11 - 0,33x$, iznosi -0,33

$$e) \quad y_t = 2 + 0,8x_t - 0,2x_{t-1} + \varepsilon_t$$

Učinak u kratkom roku: 0,8

Učinak u dugom roku: $y = 2 + 0,8x - 0,2x$, $y = 2 + 0,6x$, iznosi 0,6

$$f) \quad y_t = 2 - 0,8x_t - 0,2x_{t-1} + \varepsilon_t$$

Učinak u kratkom roku: -0,8

Učinak u dugom roku: $y = 2 - 0,8x - 0,2x$, $y = 2 - x$, iznosi 1

$$g) \quad y_t = 2 - 0,8x_t + 0,2x_{t-1} + \varepsilon_t$$

Učinak u kratkom roku: -0,8

Učinak u dugom roku: $y = 2 - 0,8x + 0,2x$, $y = 2 - 0,6x$, iznosi 0,6

$$h) \quad y_t = 2 + 0,8x_t + 0,2x_{t-1} + \varepsilon_t$$

Učinak u kratkom roku: 0,8

Učinak u dugom roku: $y = 2 + 0,8x + 0,2x$, $y = 2 + x$, iznosi 1

$$i) \quad y_t = 2 + 0,8y_{t-1} + 0,8x_t + \varepsilon_t$$

Učinak u kratkom roku: 0,8

Učinak u dugom roku: $y - 0,8y = 2 + 0,8x$, $y = 10 + 4x$, iznosi 4

$$j) \quad y_t = 2 + 0,8y_{t-1} - 0,8x_t + \varepsilon_t$$

Učinak u kratkom roku: $-0,8$

Učinak u dugom roku: $y - 0,8y = 2 - 0,8x$, $y = 10 - 4x$, iznosi -4

Zadatak 12)

Panel prvi redak i prvi stupac: uočava se pozitivan učinak kroz vrijeme, pri čemu je najjači u drugome razdoblju, i vrlo sporo iščezava.

Panel prvi redak i drugi stupac: uočava se negativan učinak kroz vrijeme, pri čemu je najjači u drugome razdoblju i veoma sporo iščezava.

Panel prvi redak i treći stupac: učinak šoka u nezavisnoj na zavisnu varijablu je pozitivan, najjači je u prvome razdoblju, te nakon toga brzo nestaje.

Panel drugi redak i prvi stupac: učinak šoka u nezavisnoj na zavisnu varijablu je negativan, najjači je u prvome razdoblju, te nakon toga brzo nestaje.

Panel drugi redak i drugi stupac: učinak šoka u nezavisnoj na zavisnu varijablu je pozitivan, najjači je u drugome razdoblju, te nakon toga nestaje umjerenom brzinom.

Panel drugi redak i treći stupac: učinak šoka u nezavisnoj na zavisnu varijablu je negativan, najjači je u prvome razdoblju, te nakon toga nestaje umjerenom brzinom. Uočava se da je brzina iščezavanja ovog učinka nešto sporija u odnosu na IRF na panelu u drugom retku i prvom stupcu.

Panel treći redak i prvi stupac: uočava se da je učinak najprije pozitivan u prvome razdoblju, nakon čega postaje negativan, no puno manje jačine i potom brzo nestaje.

Panel treći redak i drugi stupac: uočava se da je učinak najprije pozitivan u prvome razdoblju, nakon čega postaje negativan, pri čemu je puno jači negativan učinak i potom umjerenom brzinom odumire.

Panel treći redak i treći stupac: uočava se da je učinak najprije pozitivan u prvome razdoblju, nakon čega postaje negativan, no puno veće jačine i potom brzo nestaje.

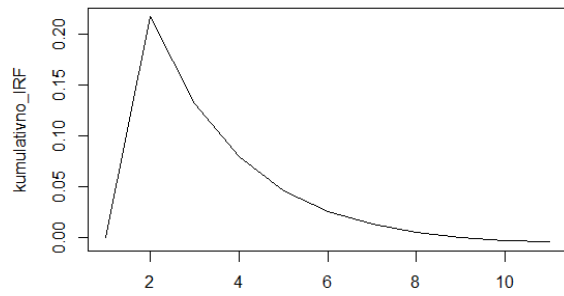
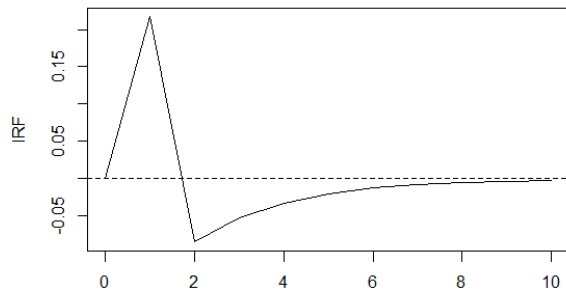
Panel četvrti redak i prvi stupac: uočava se da je učinak najprije pozitivan u prvome razdoblju, nakon čega postaje negativan i podjednake jačine, te potom brzo nestaje.

Panel treći redak i drugi stupac: uočava se da je učinak jak i negativan u prvom i drugom razdoblju, najjači upravo u drugome, no nakon toga vrlo brzo nestaje.

Panel treći redak i treći stupac: uočava se da je učinak negativan i jak, najjači u drugome razdoblju i nakon toga nestaje umjerenom brzinom.

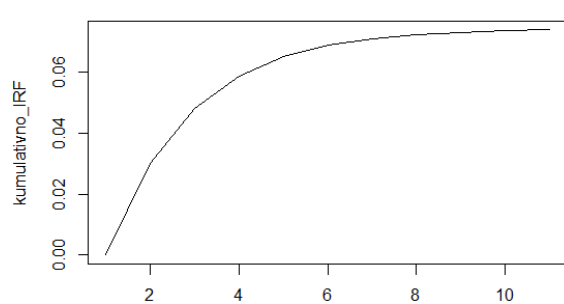
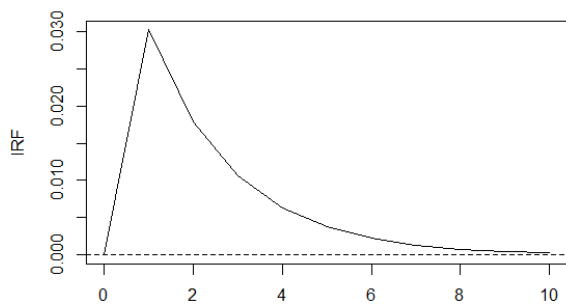
Zadatak 13)

```
dynlm(s.iip~s.nez+L(s.nez)+L(s.iip))
##
## Time series regression with "ts" data:
## Start = 1991(2), End = 2020(9)
##
## Call:
## dynlm(formula = s.iip ~ s.nez + L(s.nez) + L(s.iip))
##
## Coefficients:
## (Intercept)      s.nez      L(s.nez)      L(s.iip)
##   -0.00215      0.21688     -0.21964      0.62153
```



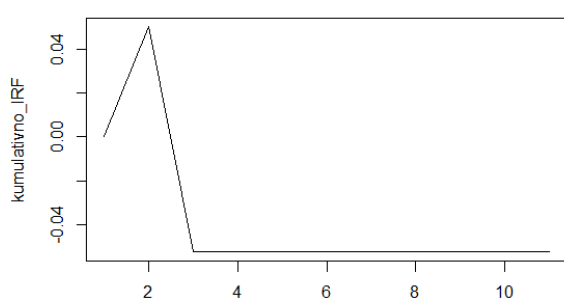
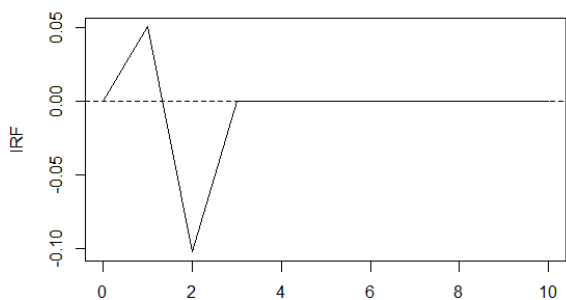
```
dynlm(s.iip~s.nez+L(s.iip))
```

```
##
## Time series regression with "ts" data:
## Start = 1991(2), End = 2020(9)
##
## Call:
## dynlm(formula = s.iip ~ s.nez + L(s.iip))
##
## Coefficients:
## (Intercept)      s.nez      L(s.iip)
## -0.002385      0.030176      0.593624
```



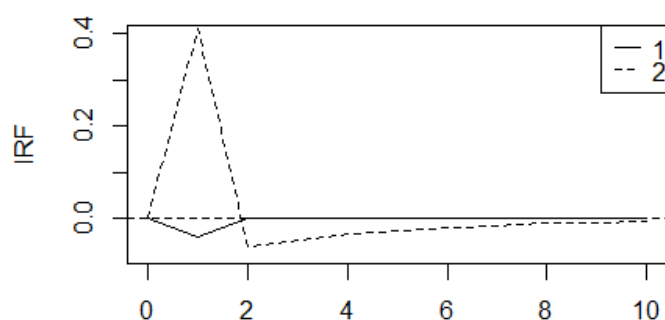
```
dynlm(s.iip~s.nez+L(s.nez))
```

```
##
## Time series regression with "ts" data:
## Start = 1991(2), End = 2020(9)
##
## Call:
## dynlm(formula = s.iip ~ s.nez + L(s.nez))
##
## Coefficients:
## (Intercept)      s.nez      L(s.nez)
## -0.005341      0.050190     -0.102739
```



```
## =====
##                               Dependent variable:
##                               -----
##                               s.iip
##                               (1)          (2)
## -----
```

	(1)	(2)
s.nez	-0.040 (0.070)	0.412*** (0.073)
L(s.iip)	0.183*** (0.070)	0.762*** (0.058)
L(s.nez)	0.006 (0.070)	-0.375*** (0.073)
Constant	0.010*** (0.003)	-0.008 (0.006)



```
grangertest(s.iip,s.nez,1)

## Granger causality test
##
## Model 1: s.nez ~ Lags(s.nez, 1:1) + Lags(s.iip, 1:1)
## Model 2: s.nez ~ Lags(s.nez, 1:1)
##   Res.Df Df      F    Pr(>F)
## 1      353
## 2      354 -1 17.557 3.528e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(s.iip,s.nez,2)

## Granger causality test
##
## Model 1: s.nez ~ Lags(s.nez, 1:2) + Lags(s.iip, 1:2)
## Model 2: s.nez ~ Lags(s.nez, 1:2)
##   Res.Df Df      F    Pr(>F)
## 1      350
## 2      352 -2 12.469 5.875e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(s.iip,s.nez,3)

## Granger causality test
##
## Model 1: s.nez ~ Lags(s.nez, 1:3) + Lags(s.iip, 1:3)
## Model 2: s.nez ~ Lags(s.nez, 1:3)
##   Res.Df Df      F    Pr(>F)
## 1      347
## 2      350 -3 8.7377 1.332e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

grangertest(s.nez,s.iip,1)

## Granger causality test
##
## Model 1: s.iip ~ Lags(s.iip, 1:1) + Lags(s.nez, 1:1)
## Model 2: s.iip ~ Lags(s.iip, 1:1)
##   Res.Df Df      F Pr(>F)
## 1     353
## 2     354 -1 1.8654 0.1729

grangertest(s.nez,s.iip,2)

## Granger causality test
##
## Model 1: s.iip ~ Lags(s.iip, 1:2) + Lags(s.nez, 1:2)
## Model 2: s.iip ~ Lags(s.iip, 1:2)
##   Res.Df Df      F Pr(>F)
## 1     350
## 2     352 -2 3.723 0.02512 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

grangertest(s.nez,s.iip,3)

## Granger causality test
##
## Model 1: s.iip ~ Lags(s.iip, 1:3) + Lags(s.nez, 1:3)
## Model 2: s.iip ~ Lags(s.iip, 1:3)
##   Res.Df Df      F Pr(>F)
## 1     347
## 2     350 -3 4.3784 0.004833 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lm(iip~nez)

##
## Call:
## lm(formula = iip ~ nez)
##
## Coefficients:
## (Intercept)          nez
## 151.48000      -0.01538

rez<-resid(lm(iip~nez))
rez<-ts(rez,start=c(1990,1),frequency = 12)
library(dynlm)
dynlm(diff(iip)~diff(nez)+L(rez))

##
## Time series regression with "ts" data:
## Start = 1990(2), End = 2020(9)
##
## Call:
## dynlm(formula = diff(iip) ~ diff(nez) + L(rez))
##
## Coefficients:
## (Intercept)  diff(nez)      L(rez)
## -0.036488    -0.002892   -0.272546

summary(ur.df(iip,type="drift",selectlags="BIC"))

##
## #####
## # Augmented Dickey-Fuller Test Unit Root #
## #####
##

```

```

## Test regression drift
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -34.829  -3.825   0.258   4.064  22.668
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  13.08607    3.41729   3.829 0.000151 ***
## z.lag.1      -0.11374    0.02948  -3.858 0.000135 ***
## z.diff.lag   -0.38113    0.04889  -7.796 6.78e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.947 on 364 degrees of freedom
## Multiple R-squared:  0.2209, Adjusted R-squared:  0.2167
## F-statistic: 51.61 on 2 and 364 DF, p-value: < 2.2e-16
##
## Value of test-statistic is: -3.8585 7.4462
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phi1  6.47  4.61  3.79

summary(ur.df(nez,type="drift",selectlags="BIC"))

##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -728.79  -73.46   5.35   85.68  450.70
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  102.74127    39.34777   2.611 0.00940 **
## z.lag.1      -0.04274    0.01639  -2.608 0.00949 **
## z.diff.lag   -0.14346    0.05186  -2.766 0.00596 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 146.3 on 364 degrees of freedom
## Multiple R-squared:  0.04504, Adjusted R-squared:  0.0398
## F-statistic: 8.585 on 2 and 364 DF, p-value: 0.0002275
##
## Value of test-statistic is: -2.6079 3.4376
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phi1  6.47  4.61  3.79

summary(ur.df(diff(iip),type="none",selectlags="BIC"))

```



```

##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -37.352  -3.330   0.039   3.819  30.165
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -1.95351    0.08320  -23.479 < 2e-16 ***
## z.diff.lag    0.36273    0.04951   7.326 1.54e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.617 on 364 degrees of freedom
## Multiple R-squared:  0.7515, Adjusted R-squared:  0.7502
## F-statistic: 550.4 on 2 and 364 DF,  p-value: < 2.2e-16
##
## Value of test-statistic is: -23.4794
##
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62

summary(ur.df(diff(nez),type="none",selectlags="BIC"))

##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -725.99  -74.98    6.23   91.37  498.53
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      -1.35860    0.07887  -17.227 < 2e-16 ***
## z.diff.lag    0.16600    0.05168   3.212 0.00144 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 145.6 on 364 degrees of freedom
## Multiple R-squared:  0.5941, Adjusted R-squared:  0.5919
## F-statistic: 266.4 on 2 and 364 DF,  p-value: < 2.2e-16
##
## Value of test-statistic is: -17.2267
##
## Critical values for test statistics:
##          1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62

rez2<-resid(lm(iip~nez))
summary(ur.df(rez2,type="none",selectlags = "BIC"))

```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -44.643  -3.872   0.111   4.350  22.760
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -0.20904    0.03822  -5.469 8.41e-08 ***
## z.diff.lag -0.27872    0.05056  -5.512 6.71e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.235 on 365 degrees of freedom
## Multiple R-squared:  0.2103, Adjusted R-squared:  0.206
## F-statistic: 48.6 on 2 and 365 DF, p-value: < 2.2e-16
##
## Value of test-statistic is: -5.469
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

AVN

2.

**VEKTORSKI
AUTOREGRESIJSKI
MODELI**

AVN

AWN

2. VEKTORSKI AUTOREGRESIJSKI MODELI

Analiza u prethodnim poglavljima odnosila se samo na jednu zavisnu varijablu koja je mogla ovisiti o jednoj ili više nezavisnih varijabli. Međutim, brojne ekonomske pojave i procesi ovisni su jedni o drugima i postoji povratna veza između njih. Na taj način je potrebno istovremeno modelirati više varijabli kao endogene. U tu svrhu razmatraju se sustavi jednadžbi koje se kompaktno zapisuju u vektore i matrice. Nadalje, kako brojne ekonomske pojave karakterizira autokorelacija, razmatraju se VAR modeli, engl. *Vector AutoRegression*, tj. vektorski autoregresijski modeli. Analizom VAR modela dobije se više informacija o dinamici pojava tijekom vremena, njihovoj povezanosti i međusobnom utjecaju. VAR modeli razvijeni su, i dalje se razvijaju od strane makroekonomista, jer su vrlo fleksibilni za analizu različitih makroekonomskih politika: Sims (1980), Blanchard i Quah (1989). VAR modeli predstavljaju generalizaciju univarijatnih $AR(p)$ modela (vidjeti odjeljak 1.9.3.). Zato se u literaturi često nalazi distinkcija između univarijatnih (uni – jedno) i multivarijatnih (multi – više) modela, pri čemu se prefiksi uni i multi odnose na broj jednadžbi u modelu.

2.1. Temeljni pojmovi

2.1.1. Oznake

Vektorski autoregresijski model može se definirati kao (Guidolin i Pedio, 2019): $VAR(p)$ je proces koji se može zapisati na sljedeći način:

$$\mathbf{y}_t = \mathbf{a}_0 + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + \dots + \mathbf{A}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t = \mathbf{a}_0 + \sum_{i=1}^p \mathbf{A}_i\mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad (2.1)$$

gdje je $\mathbf{y}_t = [y_{1t} \ y_{2t} \ \dots \ y_{Nt}]'$ N -dimenzionalni vektor koji sadrži N endogenih stacionarnih varijabli, $\mathbf{a}_0 = [a_{10} \ a_{20} \ \dots \ a_{N0}]'$ je N -dimenzionalni vektor konstanti, $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p$ su $N \cdot N$ matrice autoregresijskih koeficijenata te je $\boldsymbol{\varepsilon}_t = [\varepsilon_{1t} \ \varepsilon_{2t} \ \dots \ \varepsilon_{Nt}]'$ vektor slučajnih procesa. Vektor slučajnih procesa se još naziva i multivarijatni čisti slučajni proces. Njegove karakteristike su sljedeće:

$$E(\boldsymbol{\varepsilon}_t) = E \left(\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0}, \quad E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega}, \quad E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_s') = \mathbf{0}, \quad t \neq s \quad (2.2)$$

tj. očekivana vrijednost svakoga procesa ε_{jt} jednaka je 0, $j \in \{1, 2, \dots, N\}$, matrica varijanci-kovarijanci $\boldsymbol{\Omega}$ između svih procesa u istom vremenskom razdoblju t je pozitivno-semidefinitna matrica (vidjeti Dodatak u prvome dijelu), što znači da može postojati korelacija između varijabli ε_{jt} i ε_{kt} u istom razdoblju t , dok je matrica varijanci-kovarijanci između procesa ε_{jt} i ε_{js} nul-matrica: $\mathbf{0}$. To znači da za sve procese j i za svako razdoblje $s \neq t$ ne postoji korelacija procesa samog sa sobom: $\text{Cov}(\varepsilon_{jt}, \varepsilon_{js}) = 0$ i ne postoji korelacija procesa ε_{jt} s procesom ε_{ks} $\text{Cov}(\varepsilon_{jt}, \varepsilon_{ks}) = 0$.

Model (2.1) se može iz matričnog zapisa raspisati na sljedeći način¹²:

¹² Eksponenti kod koeficijenata a u svakoj matrici su samo oznake koje se odnose na matricu \mathbf{A}_i , ne radi se o potenciji!

$$\begin{aligned}
\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Nt} \end{bmatrix} &= \begin{bmatrix} a_{10} \\ a_{20} \\ \vdots \\ a_{N0} \end{bmatrix} + \begin{bmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1N}^1 \\ a_{21}^1 & a_{22}^1 & \cdots & a_{2N}^1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}^1 & a_{N2}^1 & \cdots & a_{NN}^1 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ \vdots \\ y_{Nt-1} \end{bmatrix} + \begin{bmatrix} a_{11}^2 & a_{12}^2 & \cdots & a_{1N}^2 \\ a_{21}^2 & a_{22}^2 & \cdots & a_{2N}^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}^2 & a_{N2}^2 & \cdots & a_{NN}^2 \end{bmatrix} \begin{bmatrix} y_{1t-2} \\ y_{2t-2} \\ \vdots \\ y_{Nt-2} \end{bmatrix} \\
&+ \dots + \begin{bmatrix} a_{11}^p & a_{12}^p & \cdots & a_{1N}^p \\ a_{21}^p & a_{22}^p & \cdots & a_{2N}^p \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}^p & a_{N2}^p & \cdots & a_{NN}^p \end{bmatrix} \begin{bmatrix} y_{1t-p} \\ y_{2t-p} \\ \vdots \\ y_{Nt-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix}
\end{aligned} \tag{2.3}$$

stoga se uočava da je kompaktan zapis pogodniji za daljnju analizu, zato se i koristi najčešće u literaturi i primjenama.

2.1.2. Strukturni i reducirani oblik modela

Bivarijatni model VAR(1) je onaj u kojemu razmatramo dvije varijable i jedan pomak ($p = 1$):

$$\begin{aligned}
y_{1t} &= a_{10} + a_{11}y_{1t-1} + a_{12}y_{2t-1} + \varepsilon_{1t} \\
y_{2t} &= a_{20} + a_{21}y_{1t-1} + a_{22}y_{2t-1} + \varepsilon_{2t}.
\end{aligned} \tag{2.4}$$

Nazvan je bivarijatni jer je $N = 2$, tj. uključene su dvije varijable, a zbog $p = 1$ se razmatra prvi pomak. Bivarijatni model može biti i drugih pomaka $p > 1$. Ovakav model naziva se **reducirani** (standardni) oblik modela (engl. *reduced form, standard form*) jer se pretpostavlja da svaka varijabla ovisi samo o svojim prethodnim vrijednostima, kao i prethodnim vrijednostima drugih varijabli u sustavu. U ekonomiji se nerijetko događa da i vrijednosti varijabli u tekućem razdoblju utječu na tekuću vrijednost drugih varijabli:

$$\begin{aligned}
y_{1t} &= b_{10} - b_{12}y_{2t} + c_{11}y_{1t-1} + c_{12}y_{2t-1} + u_{1t} \\
y_{2t} &= b_{20} - b_{21}y_{1t} + c_{21}y_{1t-1} + c_{22}y_{2t-1} + u_{2t}
\end{aligned} \tag{2.5}$$

i ovaj oblik modela se naziva **strukturni** oblik (engl. *structural form, primitive form*), u_{1t} i u_{2t} kao strukturne greške relacije su **nekorelirani čisti slučajni procesi**. Strukturni oblik modela se obično prevodi u reducirani oblik (vidjeti razloge u odjeljku 2.3), na sljedeći način:

$$\begin{aligned}
y_{1t} + b_{12}y_{2t} &= b_{10} + c_{11}y_{1t-1} + c_{12}y_{2t-1} + u_{1t} \\
y_{2t} + b_{21}y_{1t} &= b_{20} + c_{21}y_{1t-1} + c_{22}y_{2t-1} + u_{2t}
\end{aligned} \tag{2.6}$$

$$\underbrace{\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}}_B \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \underbrace{\begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}}_b + \underbrace{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}}_C \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{2.7}$$

$$\mathbf{B}y_t = \mathbf{b} + \mathbf{C}y_{t-1} + \mathbf{u}_t. \tag{2.8}$$

Problem kod strukturnog oblika VAR modela je taj što se procesi inovacija u_{1t} i u_{2t} nalaze u obje jednadžbe: ako je u drugoj jednadžbi parametar b_{21} različit od 0, to znači da je varijabla y_{1t}

značajna, a sama je definirana preko u_{1t} koji se sada prelijeva u drugu jednadžbu. Slično vrijedi i za prvu jednadžbu i ako je parametar b_{12} različit od 0. Pretpostavke linearnog regresijskog modela (vidjeti odjeljak 2.2.2. u prvom dijelu), uključivale su i pretpostavku $E(\varepsilon | X) = 0$, tj. regresori nisu smjeli biti korelirani sa slučajnom varijablom. Stoga se strukturni oblik VAR modela ne može procijeniti u tom obliku, već se prevodi u reducirani oblik:

$$\mathbf{B}^{-1} \cdot \mathbf{B}y_t = \mathbf{b} + \mathbf{C}y_{t-1} + \mathbf{u}_t \quad (2.9)$$

$$y_t = \underbrace{\mathbf{a}_0}_{=\mathbf{B}^{-1}\mathbf{b}} + \underbrace{\mathbf{A}_1}_{=\mathbf{B}^{-1}\mathbf{C}} y_{t-1} + \underbrace{\varepsilon_t}_{=\mathbf{B}^{-1}\mathbf{u}_t}. \quad (2.10)$$

Model u (2.10) se sada može procijeniti metodom najmanjih kvadrata. Greške relacije reduciranog modela ε_{1t} i ε_{2t} su kombinacija originalnih procesa u_{1t} i u_{2t} jer se računaju na sljedeći način:

$$\varepsilon_t = \mathbf{B}^{-1}\mathbf{u}_t = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} \frac{u_{1t} - b_{12}u_{2t}}{1 - b_{12}b_{21}} \\ \frac{u_{2t} - b_{21}u_{1t}}{1 - b_{12}b_{21}} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}. \quad (2.11)$$

Očekivana vrijednost grešaka u (2.11) jednaka je $E\left(\begin{bmatrix} \frac{u_{1t} - b_{12}u_{2t}}{1 - b_{12}b_{21}} \\ \frac{u_{2t} - b_{21}u_{1t}}{1 - b_{12}b_{21}} \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ jer je $E(u_{1t}) = E(u_{2t}) = 0$.

Nadalje, kako su u_{1t} i u_{2t} nekorelirane, tj. $\text{Cov}(u_{1t}, u_{2t}) = 0$, varijance od ε_{1t} i ε_{2t} jednake su:

$$\begin{bmatrix} \sigma_{\varepsilon_1}^2 \\ \sigma_{\varepsilon_2}^2 \end{bmatrix} = \text{Var} \begin{bmatrix} \frac{u_{1t} - b_{12}u_{2t}}{1 - b_{12}b_{21}} \\ \frac{u_{2t} - b_{21}u_{1t}}{1 - b_{12}b_{21}} \end{bmatrix} = \begin{bmatrix} \frac{\text{Var}(u_{1t} - b_{12}u_{2t})}{(1 - b_{12}b_{21})^2} \\ \frac{\text{Var}(u_{2t} - b_{21}u_{1t})}{(1 - b_{12}b_{21})^2} \end{bmatrix} \quad (2.12)$$

$$= \begin{bmatrix} \frac{\text{Var}(u_{1t}) + b_{12}^2 \text{Var}(u_{2t}) - 2\text{Cov}(u_{1t}, u_{2t})}{(1 - b_{12}b_{21})^2} \\ \frac{\text{Var}(u_{2t}) + b_{21}^2 \text{Var}(u_{1t}) - 2\text{Cov}(u_{1t}, u_{2t})}{(1 - b_{12}b_{21})^2} \end{bmatrix} = \begin{bmatrix} \frac{\sigma_{u_1}^2 + b_{12}^2 \sigma_{u_2}^2}{(1 - b_{12}b_{21})^2} \\ \frac{\sigma_{u_2}^2 + b_{21}^2 \sigma_{u_1}^2}{(1 - b_{12}b_{21})^2} \end{bmatrix}.$$

Dakle, uočava se da su varijance konstantne. Konačno, kovarijanca između procesa ε_{1t} i ε_{2t} jednaka je:

$$\sigma_{\varepsilon_1 \varepsilon_2} = \text{Cov}(\varepsilon_{1t}, \varepsilon_{2t}) = \text{Cov}\left(\frac{u_{1t} - b_{12}u_{2t}}{1 - b_{12}b_{21}}, \frac{u_{2t} - b_{21}u_{1t}}{1 - b_{12}b_{21}}\right) = \frac{-(b_{21}\sigma_{u_1}^2 + b_{12}\sigma_{u_2}^2)}{(1 - b_{12}b_{21})^2}, \quad (2.13)$$

i uočava se da je različita od 0, tj. **proces i ε_{1t} i ε_{2t} su korelirani**. Zato se matrica varijanci-kovarijanci procesa ε_{1t} i ε_{2t} može zapisati na sljedeći način:

$$\mathbf{\Omega} = \begin{bmatrix} \sigma_{\varepsilon 1}^2 & \sigma_{\varepsilon 1 \varepsilon 2} \\ \sigma_{\varepsilon 1 \varepsilon 2} & \sigma_{\varepsilon 2}^2 \end{bmatrix}. \quad (2.14)$$

Prethodno se može poopćiti na VAR model s više od dvije varijable, stoga je $\text{Cov}(u_{jt}, u_{kt}) = 0$ za svaki par varijabli, dok je matrica $\mathbf{\Omega}$ pozitivno semidefinitna matrica, u kojoj općenito vrijedi $\text{Cov}(\varepsilon_{jt}, \varepsilon_{kt}) \neq 0$:

$$\mathbf{\Omega} = \begin{bmatrix} \sigma_{\varepsilon 1}^2 & \sigma_{\varepsilon 1 \varepsilon 2} & \sigma_{\varepsilon 1 \varepsilon 3} & \cdots & \sigma_{\varepsilon 1 \varepsilon N} \\ \sigma_{\varepsilon 1 \varepsilon 2} & \sigma_{\varepsilon 2}^2 & \sigma_{\varepsilon 2 \varepsilon 3} & \cdots & \sigma_{\varepsilon 2 \varepsilon N} \\ \sigma_{\varepsilon 1 \varepsilon 3} & \sigma_{\varepsilon 2 \varepsilon 3} & \sigma_{\varepsilon 3}^2 & \cdots & \sigma_{\varepsilon 3 \varepsilon N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{\varepsilon 1 \varepsilon N} & \sigma_{\varepsilon 2 \varepsilon N} & \sigma_{\varepsilon 3 \varepsilon N} & \cdots & \sigma_{\varepsilon N}^2 \end{bmatrix}. \quad (2.15)$$

Model u (2.1) može sadržavati i egzogene varijable, stoga se može zapisati i u sljedećoj formi:

$$\mathbf{y}_t = \mathbf{a}_0 + \sum_{i=1}^p \mathbf{A}_i \mathbf{y}_{t-i} + \sum_{i=1}^q \mathbf{B}_i \mathbf{x}_{t-i} + \boldsymbol{\varepsilon}_t, \quad (2.16)$$

gdje je $\mathbf{x}_t = [x_{1t} \ x_{2t} \ \cdots \ x_{Mt}]'$ N -dimenzionalni vektor koji sadrži M egzogenih stacionarnih varijabli, $\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_q$ su $M \cdot M$ matrice koeficijenata.

Primjer 2.1.

Zapišimo matrično sljedeće VAR modele:

- $$\begin{aligned} y_{1t} &= 2 + 0,5y_{1t-1} - 0,1y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= 1 - 0,1y_{1t-1} + 0,9y_{2t-1} + \varepsilon_{2t}; \end{aligned}$$
- $$\begin{aligned} y_{1t} &= 2 + 0,5y_{1t-1} - 0,2y_{1t-2} - 0,1y_{2t-1} + 0,8y_{2t-2} + \varepsilon_{1t} \\ y_{2t} &= 2 - 0,5y_{1t-1} + 0,8y_{1t-2} + 0,3y_{2t-1} + 0,2y_{2t-2} + \varepsilon_{2t} \end{aligned}$$
- $$\begin{aligned} y_{1t} &= 2 + 0,5y_{1t-1} - 0,2y_{2t-1} - 0,1y_{3t-1} + \varepsilon_{1t} \\ y_{2t} &= 1 - 0,5y_{1t-1} + 0,8y_{2t-1} + 0,3y_{3t-1} + \varepsilon_{2t} \\ y_{3t} &= 4 + 0,1y_{1t-1} + 0,1y_{2t-1} + 0,5y_{3t-1} + \varepsilon_{3t} \end{aligned}$$

Pod a) se radi o VAR(1) modelu s dvije varijable: $\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0,5 & -0,1 \\ -0,1 & 0,9 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$.

U b) uočavamo da se radi o VAR(2) modelu s dvije varijable:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0,5 & -0,1 \\ -0,5 & 0,3 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} -0,2 & 0,8 \\ 0,8 & 0,2 \end{bmatrix} \begin{bmatrix} y_{1t-2} \\ y_{2t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

Konačno, pod c) se radi o VAR(1) s tri varijable:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} + \begin{bmatrix} 0,5 & -0,2 & -0,1 \\ -0,5 & 0,8 & 0,3 \\ 0,1 & 0,1 & 0,5 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}.$$

2.2. Stabilnost i stacionarnost modela

2.2.1. Stabilnost i stacionarnost VAR(1) modela

Slično kao i u slučaju univarijatnih AR(p) modela, gdje smo razmatrali njihovu stacionarnost, za VAR modele je potrebno razmotriti stabilnost i stacionarnost modela. Kada se je razmatrao AR(1) model, uvjet za stacionarnost jest bio da je parametar ϕ_1 po apsolutnoj vrijednosti manji od 1 (vidjeti odjeljak 1.9.3.).

Razmotrimo li VAR(1) model, $\mathbf{y}_t = \mathbf{a}_0 + \mathbf{A}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$, možemo ga zapisati na sljedeći način:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{a}_0 + \mathbf{A}_1 \underbrace{\mathbf{y}_{t-1}}_{=\mathbf{a}_0 + \mathbf{A}_1 \mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_{t-1}} + \boldsymbol{\varepsilon}_t = \mathbf{a}_0 + \mathbf{A}_1 \left(\mathbf{a}_0 + \mathbf{A}_1 \underbrace{\mathbf{y}_{t-2}}_{=\mathbf{a}_0 + \mathbf{A}_1 \mathbf{y}_{t-3} + \boldsymbol{\varepsilon}_{t-2}} + \boldsymbol{\varepsilon}_{t-1} \right) + \boldsymbol{\varepsilon}_t \\ &= \mathbf{a}_0 + \mathbf{A}_1 (\mathbf{a}_0 + \mathbf{A}_1 (\mathbf{a}_0 + \mathbf{A}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_{t-2}) + \boldsymbol{\varepsilon}_{t-1}) + \boldsymbol{\varepsilon}_t \\ &= \dots = (I + \mathbf{A}_1 + \mathbf{A}_1^2 + \dots + \mathbf{A}_1^n) \mathbf{a}_0 + \sum_{i=0}^n \mathbf{A}_1^i \boldsymbol{\varepsilon}_{t-i} + \mathbf{A}_1^{n+1} \mathbf{y}_{t-n-1}, \end{aligned} \quad (2.17)$$

gdje za $n \rightarrow \infty$ možemo pisati¹³:

$$\mathbf{y}_t = \underbrace{(I + \mathbf{A}_1 + \mathbf{A}_1^2 + \dots) \mathbf{a}_0}_{\boldsymbol{\mu}} + \sum_{i=1}^{\infty} \mathbf{A}_1^i \boldsymbol{\varepsilon}_{t-i} + \boldsymbol{\varepsilon}_t \quad (2.18)$$

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{i=1}^{\infty} \mathbf{A}_1^i \boldsymbol{\varepsilon}_{t-i} + \boldsymbol{\varepsilon}_t. \quad (2.19)$$

Model u (2.19) naziva se VMA(∞) reprezentacija VAR(1) modela, tj. vektorska MA(∞) reprezentacija (pogledati odjeljak 1.9.3.3. u kojem se razmatrala MA(∞) reprezentacija AR modela za usporedbu). Uočimo da je zapis u (2.19) multivarijatni slučaj Woldovog teorema dekompozicije. Kako se uočava da vrijednosti \mathbf{y}_t ovise o prethodnim vrijednostima $\boldsymbol{\varepsilon}_t$, pri čemu za svaki pomak i matrica koeficijenata \mathbf{A}_1 treba biti potencirana upravo pomakom i , šokovi koji se dogode u $\boldsymbol{\varepsilon}_t$ u nekom od prethodnih pomaka i moraju tijekom vremena nestati ako se radi o stacionarnim procesima u \mathbf{y}_t . To znači da koeficijenti u matricama \mathbf{A}_1^i moraju brzo težiti vrijednostima 0 kako bi učinci nestali tijekom vremena kako se i povećava.

Zato kažemo da je VAR(1) stabilan ako i samo ako matrica \mathbf{A}_1 ima **svojstvene vrijednosti (λ_i) koje su po apsolutnoj vrijednosti manje od 1** (o svojstvenim vrijednostima matrice vidjeti Dodatak u prvom dijelu ove serije publikacija).

¹³ Zbog svojstava koja će se opisati u nastavku teksta za svojstvene vrijednosti matrice \mathbf{A}_1 , izraz $\mathbf{A}_1^{n+1} \mathbf{y}_{t-n-1} \rightarrow \mathbf{0}$ kada $n \rightarrow \infty$.

Ako zadovoljavaju navedeno, tada za $n \rightarrow \infty$ vrijedi:

- i. $A_1^i \xrightarrow{i \rightarrow \infty} \mathbf{0}$, jer se matrica A_1 može prikazati pomoću svojstvenih vektora i dijagonalizirane matrice (čiji su elementi na glavnoj dijagonali upravo svojstvene vrijednosti od A_1 , vidjeti Dodatak¹⁴ u prvom dijelu ove serije publikacija):

$$A_1^i = P A^i P^{-1} \xrightarrow{i \rightarrow \infty} \mathbf{0}$$
- ii. Suma $(I + A_1 + A_1^2 + \dots)a_0$ u (2.18) postoji i $(I + A_1 + A_1^2 + \dots + A_1^n)a_0 \rightarrow (I - A_1)^{-1}a_0 = \mu$.
- iii. Postoje sume $\sum_{i=1}^{\infty} A_1^i$ i $\sum_{i=1}^{\infty} A_1^i \varepsilon_{t-i}$.

Kako svojstvene vrijednosti matrice A_1 zadovoljavaju jednadžbu $\det(I\lambda - A_1) = 0$, one su recipročne vrijednosti **korijena karakteristične jednadžbe**:

$$\det(I - A_1 L) = 0. \quad (2.20)$$

Ako su svojstvene vrijednosti matrice A_1 po apsolutnoj vrijednosti manje od 1, to je ekvivalentno izjavi da su **korijeni karakteristične jednadžbe (2.20) po apsolutnoj vrijednosti veći od 1** (ekvivalentno da su svojstvene vrijednosti po apsolutnoj vrijednosti manje od 1).

Rezultat se može poopćiti za VAR(p) model (vidjeti Lüketpohl, 2010): da svojstvene vrijednosti svih matrica A_i moraju po apsolutnoj vrijednosti biti manje od 1, ili da korijeni karakteristične jednadžbe

$$\det(I - A_1 \lambda - A_2 \lambda^2 - \dots - A_p \lambda^p) = 0 \quad (2.21)$$

budu po apsolutnoj vrijednosti veći od 1.

Ako je VAR model stabilan (svojstvene vrijednosti matrica A_i su manje od 1 gledajući apsolutno), tada je i stacionaran. Međutim, obrat ne vrijedi, nestabilni VAR može biti stacionaran.

Primjer 2.2

Usporedimo stabilnost dva VAR(1) modela, na način da izračunamo svojstvene vrijednosti objiju matrica A_1 , te zaključimo o stabilnostima modela:

$$\begin{aligned} y_{1t} &= 2 + 2y_{1t-1} + y_{2t-1} + \varepsilon_{1t} & y_{1t} &= 2 + 0,4y_{1t-1} + 0,6y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= 3 + y_{1t-1} + 2y_{2t-1} + \varepsilon_{2t} & y_{2t} &= 3 + 0,1y_{1t-1} + 0,3y_{2t-1} + \varepsilon_{2t} \end{aligned}$$

Za prvi model (lijevi panel) je matrica $A_1 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, čije su svojstvene vrijednosti jednake:

$$\det(I\lambda - A_1) = \det \left(\begin{bmatrix} 1-2\lambda & -1 \\ -1 & 1-2\lambda \end{bmatrix} \right) = 0 \Rightarrow (1-2\lambda)(1-2\lambda) - (-1)^2 = 0 \Rightarrow \lambda_1 = 3 \text{ i } \lambda_2 = 1.$$

Odgovarajući svojstveni vektori su redom $[1 \ 1]'$ i $[-1 \ 1]'$. Kako niti jedna svojstvena vrijednost po apsolutnoj vrijednosti nije manja od 1, model je nestabilan.

¹⁴ U dodatku se kvadratna matrica A dijagonalizira pomoću izraza $P^{-1}AP = \Lambda$, pa se A može dobiti iz navedenog izraza na sljedeći način: $P P^{-1}AP = P \Lambda$, $AP P^{-1} = P \Lambda P^{-1}$, $A = P \Lambda P^{-1}$.

Razmotrimo to u okviru zapisa u (2.19):

$$y_t = \mu + \sum_{i=1}^{\infty} A_1^i \varepsilon_{t-i} + \varepsilon_t = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^i \varepsilon_{t-i} + \varepsilon_t = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}^i \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \varepsilon_{t-i} + \varepsilon_t.$$

Uočimo da u matrici $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}^i$ za veću vrijednost i dolazi do ubrzanog povećanja vrijednosti 3^i , što uzrokuje eksplozivno ponašanje y_1 . Stoga sustav nije stabilan.

Za drugi model (desni panel) je matrica $A_1 = \begin{bmatrix} 0,4 & 0,6 \\ 0,1 & 0,3 \end{bmatrix}$, čije su svojstvene vrijednosti $\det(\lambda I - A_1) = \det \begin{pmatrix} -0,4\lambda & -0,6 \\ -0,1 & 1 - 0,3\lambda \end{pmatrix} = 0 \Rightarrow (1-0,4\lambda)(1-0,3\lambda) - (-0,6) \cdot (-0,1) = 0 \Rightarrow \lambda_1 = 0,6$ i $\lambda_2 = 0,1$. Odgovarajući svojstveni vektori su redom $[3 \ 1]'$ i $[-2 \ 1]'$. Kako su obje svojstvene vrijednosti po apsolutnoj vrijednosti manje od vrijednosti 1, model je stabilan.

Razmotrimo to u okviru zapisa u (2.19):

$$y_t = \mu + \sum_{i=1}^{\infty} A_1^i \varepsilon_{t-i} + \varepsilon_t = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} 0,4 & 0,6 \\ 0,1 & 0,3 \end{bmatrix}^i \varepsilon_{t-i} + \varepsilon_t \\ = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0,6 & 0 \\ 0 & 0,1 \end{bmatrix}^i \begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix}^{-1} \varepsilon_{t-i} + \varepsilon_t$$

Uočimo da u matrici $\begin{bmatrix} 0,6 & 0 \\ 0 & 0,1 \end{bmatrix}^i$ za veću vrijednost i dolazi do ubrzanog smanjivanja vrijednosti $0,6^i$ i $0,1^i$, što znači da je zadovoljena pretpostavka $A_1^i = P \Lambda^i P^{-1} \xrightarrow{i \rightarrow \infty} \mathbf{0}$.

2.2.2. Stacionarnost multivarijatnog procesa y_t

Napomena: kao u slučaju univarijatnih nizova kada smo definirali stacionaran proces, može se definirati i slaba stacionarnost N -dimenzionalnog niza $y_t = [y_{1t} \ y_{2t} \ \dots \ y_{Nt}]'$. Ovaj niz je (slabo) stacionaran ako su prva dva bezuvjetna momenta konačna i konstantna kroz vrijeme:

$$E(y_t) = \mu, \quad \forall t \quad (2.22)$$

$$E[(y_t - \mu)(y_t - \mu)'] = \Gamma, \quad \forall t \quad (2.23)$$

$$E[(y_t - \mu)(y_{t-h} - \mu)'] = \Gamma_h \quad \forall t \quad \forall h \quad (2.24)$$

Dakle, izraz (2.22) čitamo da je vektor očekivanja svih procesa y_{jt} jednak vektoru konstanti $\mu = [\mu_1 \ \mu_2 \ \dots \ \mu_N]'$ koje su konačne, (2.23) čitamo da je matrica varijanci-kovarijanci za procese u y_t jednaka Γ , gdje su dijagonalni elementi varijance procesa u y_t , dok su nedijagonalni elementi kovarijance (elementi u Γ su konačni), te (2.24) čitamo da kros-kovarijancijska matrica (engl. *cross-covariance matrix*) Γ_h pri pomaku h ima konačne elemente.

Kada se razmatra multivarijatni proces, možemo govoriti o njegovoj stacionarnosti kako je definirano u prethodnoj napomeni. Međutim, kada se razmatra VAR model, potrebno je ispitati

njegovu stabilnost. Ako je VAR model stabilan, ujedno je i stacionaran, jer vrijednosti u matricama A_1^i u (2.19) neće divergirati tijekom vremena.

2.2.3. Momenti VAR(1) modela

Sada kada razmatramo momente VAR(1) procesa za (slabo) stacionaran N -dimenzionalni proces $y_t = [y_{1t} \ y_{2t} \ \dots \ y_{Nt}]'$, nalazimo sljedeće:

$$E(y_t) = E(a_0 + A_1 y_{t-1} + \varepsilon_t) \Rightarrow \mu = a_0 + A_1 \mu + \mathbf{0} \Rightarrow \mu = (I - A_1)^{-1} a_0, \quad (2.25)$$

uočimo da je očekivanje jednako vektoru konstanti. Nadalje, ako se model u (2.19) zapiše na sljedeći način:

$$y_t = \mu + \sum_{i=1}^{\infty} A_1^i \varepsilon_{t-i} + \varepsilon_t = \mu + \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i} \quad (2.26)$$

gdje vrijedi $\Theta_i = A_1^i$, te se u ovome obliku uvrsti u izračun matrice varijanci-kovarijanci:

$$\begin{aligned} E[(y_t - \mu)(y_t - \mu)'] &= E\left[\left(\mu + \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i} - \mu\right)\left(\mu + \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i} - \mu\right)'\right] \\ &= E\left[\left(\sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i}\right)\left(\sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i}\right)'\right] = \sum_{i=0}^{\infty} \Theta_i \Omega \Theta_i' \end{aligned} \quad (2.27)$$

gdje je Ω matrica varijanci-kovarijanci slučajnih procesa (vidjeti (1.14) za slučaj modela s dvije varijable), dobiva se matrica varijanci-kovarijanci (2.27) čiji su elementi konstante (konačni brojevi), jer svojstva koja smo za stabilnost VAR modela razmotrili za matricu A_1 vrijede i za matricu Θ_i . Konačno, ako razmotrimo kros-kovarijančnu matricu:

$$\begin{aligned} E[(y_t - \mu)(y_{t-h} - \mu)'] &= E\left[\left(\mu + \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i} - \mu\right)\left(\mu + \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-h-i} - \mu\right)'\right] \\ &= E\left[\left(\sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i}\right)\left(\sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-h-i}\right)'\right] \\ &= E\left[\left(\sum_{i=0}^{h-1} \Theta_i \varepsilon_{t-i} + \sum_{i=0}^{\infty} \Theta_{h+i} \varepsilon_{t-h-i}\right)\left(\sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-h-i}\right)'\right] \\ &= \sum_{i=0}^{\infty} \Theta_{h+i} \Omega \Theta_i' \end{aligned} \quad , (2.28)$$

uočavamo da se također radi o matrici koja ima konačne elemente.

Primjer 2.3.

Učitajmo datoteku „rh_dinamicki.txt“ u RStudio. Datoteka sadrži mjesečne podatke o kretanju ukupne nezaposlenosti u Hrvatskoj, te IIP (indeks industrijske proizvodnje, 2015 = 100). Desezonirajmo obje varijable pomoću multiplikativne metode dekomponiranja vremenskog niza, te potom izračunajmo stope rasta objiju varijabli, koristeći formulu

$$r_t = \ln\left(\frac{y_t}{y_{t-12}}\right).$$

Neka je stopa rasta IIP-a varijabla y_{1t} , a stopa rasta nezaposlenosti y_{2t} . Početni datum je siječanj 2000. godine.

Procijenimo VAR(1) za varijable y_{1t} i y_{2t} , zapišimo reducirani oblik modela, procijenjenu matricu varijanci-kovarijanci $\hat{\Omega}$ te svojstvene vrijednosti matrice A_1 . Komentirajmo značajnost varijabli u modelu, kao i stabilnost modela.

U okviru paketa vars naredba VAR(...) procjenjuje VAR modele tako da u:

$$\text{VAR}(\text{matrica}, p=1, \text{type}=\text{"const"})$$

kao prvi argument navodimo naziv matrice podataka u kojoj se nalaze varijable za koje je potrebno procijeniti model, drugi argument je broj pomaka p , te se pod argument type navodi vrsta VAR modela s obzirom na uključivanje determinističkih regresora (bez regresora, samo konstanta, trend). Slika 2.1 predočava postupak procjene modela kao i sam ispis. Uočava se kako je ispis veći u odnosu na običan regresijski model jer je sada potrebno svaku jednadžbu modela procijeniti kao posebnu regresijsku jednadžbu, stoga je za svaku dan ispis koji je bio uobičajen za univarijatnu analizu. Posebno se kao dio ispisa prikazuju svojstvene vrijednosti matrice procijenjenih parametara, pod nazivom *Roots of characteristic polynomials*, gdje se uočava da su obje svojstvene vrijednosti po apsolutnoj vrijednosti manje od 1 (jer su $\lambda_1 = 0,989$ i $\lambda_2 = 0,443$), pa se zaključuje da je model stabilan. Sam model je moguće zapisati na sljedeći način:

$$\begin{aligned} y_{1t} &= 0,0002 + 0,465y_{1t-1} - 0,096y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= 0,0003 - 0,121y_{1t-1} + 0,966y_{2t-1} + \varepsilon_{2t}; \end{aligned}$$

što matricno možemo pisati kao: $\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 0,0002 \\ 0,0003 \end{bmatrix} + \begin{bmatrix} 0,465 & -0,096 \\ -0,121 & 0,966 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$ dok su matrica varijanci-kovarijanci, ali i korelacijska matrica za procese ε_{1t} i ε_{2t} dani na dnu samog ispisa i iznose:

$$\hat{\Omega} = \begin{bmatrix} 0,0019 & -0,0002 \\ -0,0002 & 0,0006 \end{bmatrix}, \widehat{Corr} = \begin{bmatrix} 1 & -0,198 \\ -0,198 & 1 \end{bmatrix}.$$

Ako se razmatra značajnost varijabli u modelu, uočava se da jedino konstante u oba modela nisu značajne pri uobičajenim razinama značajnosti.

```
library(vars)
matrica<-cbind(s.iip,s.nez)
m1<-VAR(matrica,p=1,type="const")
summary(m1)

##
## VAR Estimation Results:
## =====
## Endogenous variables: s.iip, s.nez
## Deterministic variables: const
```

```

## Sample size: 236
## Log Likelihood: 943.5
## Roots of the characteristic polynomial:
## 0.9886 0.443
## Call:
## VAR(y = matrica, p = 1, type = "const")
##
## Estimation results for equation s.iip:
## =====
## s.iip = s.iip.l1 + s.nez.l1 + const
##
##              Estimate Std. Error t value Pr(>|t|)
## s.iip.l1  0.4653245  0.0589947   7.888 1.19e-13 ***
## s.nez.l1 -0.0962168  0.0231725  -4.152 4.62e-05 ***
## const    0.0002102  0.0029718   0.071  0.944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0439 on 233 degrees of freedom
## Multiple R-Squared: 0.4003, Adjusted R-squared: 0.3952
## F-statistic: 77.77 on 2 and 233 DF, p-value: < 2.2e-16
##
## Estimation results for equation s.nez:
## =====
## s.nez = s.iip.l1 + s.nez.l1 + const
##
##              Estimate Std. Error t value Pr(>|t|)
## s.iip.l1 -0.1213654  0.0339972  -3.570 0.000434 ***
## s.nez.l1  0.9662333  0.0133537  72.357 < 2e-16 ***
## const    0.0002811  0.0017125   0.164 0.869770
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0253 on 233 degrees of freedom
## Multiple R-Squared: 0.9701, Adjusted R-squared: 0.9698
## F-statistic: 3780 on 2 and 233 DF, p-value: < 2.2e-16
##
## Covariance matrix of residuals:
##          s.iip      s.nez
## s.iip  0.0019270 -0.0002199
## s.nez -0.0002199  0.0006399
##
## Correlation matrix of residuals:
##          s.iip      s.nez
## s.iip  1.0000 -0.1981
## s.nez -0.1981  1.0000

```

Slika 2.1. Procjena i ispis VAR(1) modela s konstantom

Primjer 2.4.

Temeljem podataka iz prethodnog primjera (2.3), procijenimo VAR(2) model i interpretirajmo rezultate.

Sada je potrebno u naredbi VAR(...) umjesto $p = 1$ zapisati $p = 2$. Ispis je prikazan na slici 2.2. Uočava se da se povećao broj procijenjenih parametara, kao i što sada razmatramo 4 svojstvene vrijednosti jer se razmatraju po dvije za svaku matricu procijenjenih parametara.

Sada je model sljedeći: $y_{1t} = 0,001 + 0,317y_{1t-1} + 0,329y_{1t-2} - 0,292y_{2t-1} + 0,244y_{2t-2} + \varepsilon_{1t}$
 $y_{2t} = -0,0006 - 0,072y_{1t-1} + 0,003y_{1t-2} + 1,504y_{2t-1} - 0,539y_{2t-2} + \varepsilon_{2t}$,

što je matricno moguće zapisati na način:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 0,001 \\ -0,0006 \end{bmatrix} + \begin{bmatrix} 0,317 & 0,329 \\ -0,072 & 1,504 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} -0,292 & 0,244 \\ 0,003 & -0,539 \end{bmatrix} \begin{bmatrix} y_{1t-2} \\ y_{2t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

```
m2<-VAR(matrica,p=2,type="const")
summary(m2)

##
## VAR Estimation Results:
## =====
## Endogenous variables: s.iip, s.nez
## Deterministic variables: const
## Sample size: 235
## Log Likelihood: 994.285
## Roots of the characteristic polynomial:
## 0.9361 0.7804 0.5484 0.4447
## Call:
## VAR(y = matrica, p = 2, type = "const")
##
##
## Estimation results for equation s.iip:
## =====
## s.iip = s.iip.l1 + s.nez.l1 + s.iip.l2 + s.nez.l2 + const
##
##           Estimate Std. Error t value Pr(>|t|)
## s.iip.l1  0.316681   0.062721   5.049 9.03e-07 ***
## s.nez.l1 -0.291998   0.108593  -2.689 0.00769 **
## s.iip.l2  0.329149   0.062578   5.260 3.30e-07 ***
## s.nez.l2  0.243562   0.106020   2.297 0.02250 *
## const     0.001014   0.002793   0.363 0.71692
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.04105 on 230 degrees of freedom
## Multiple R-Squared: 0.4818, Adjusted R-squared: 0.4728
## F-statistic: 53.45 on 4 and 230 DF, p-value: < 2.2e-16
##
##
## Estimation results for equation s.nez:
## =====
## s.nez = s.iip.l1 + s.nez.l1 + s.iip.l2 + s.nez.l2 + const
##
##           Estimate Std. Error t value Pr(>|t|)
## s.iip.l1 -0.0723975   0.0327235  -2.212  0.0279 *
## s.nez.l1  1.5035686   0.0566568  26.538 <2e-16 ***
## s.iip.l2  0.0032895   0.0326488   0.101  0.9198
## s.nez.l2 -0.5388676   0.0553141  -9.742 <2e-16 ***
## const    -0.0005543   0.0014573  -0.380  0.7040
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.02142 on 230 degrees of freedom
## Multiple R-Squared: 0.9788, Adjusted R-squared: 0.9784
## F-statistic: 2650 on 4 and 230 DF, p-value: < 2.2e-16
##
##
## Covariance matrix of residuals:
##           s.iip      s.nez
## s.iip  0.0016852 -0.0001291
## s.nez -0.0001291  0.0004587
##
```

```
## Correlation matrix of residuals:
##      s.iip  s.nez
## s.iip 1.0000 -0.1468
## s.nez -0.1468 1.0000
```

Slika 2.2. Procjena VAR(2) modela s konstantom

Nadalje, svojstvene vrijednosti su redom $\lambda_1 = 0,936$, $\lambda_2 = 0,780$, $\lambda_3 = 0,548$ i $\lambda_4 = 0,445$ te je i ovaj model stabilan jer su sve svojstvene vrijednosti po apsolutnoj vrijednosti manje od 1. U prvoj jednadžbi jedino konstanta nije značajna pri uobičajenim razinama značajnosti, dok u drugoj jednadžbi uz konstantu nije značajna varijabla $y_{2,t-1}$ (čitatelju preostaje za vježbu komentirati pri kojim razinama značajnosti su koji pomaci varijabli značajni u obje jednadžbe). Konačno, matrica varijanci-kovarijanci i korelacijska matrica za procese ε_{1t} i ε_{2t} jednake su:

$$\hat{\Omega} = \begin{bmatrix} 0,0017 & -0,0001 \\ -0,0001 & 0,0005 \end{bmatrix}, \widehat{Corr} = \begin{bmatrix} 1 & -0,147 \\ -0,147 & 1 \end{bmatrix}.$$

2.3. Nekoliko napomena o VAR modelima

Kao i kod modeliranja ARMA modela, ideja je da se slijedi parsimonija i kod VAR modela. Odabir adekvatnog broja pomaka p kod VAR modela se temelji kao i kod ARMA modela na informacijskim kriterijima i testovima (vidjeti odjeljak 2.6.). Valja obratiti pažnju da je broj parametara koje je potrebno procijeniti u VAR modelu puno veći u odnosu na model u kojemu je samo jedna zavisna varijabla. Naime, potrebno je procijeniti N konstanti za svaku jednadžbu, potom $N \cdot N$ parametara u svakoj matrici A_i , što znači da za reducirani oblik VAR(p) treba procijeniti ukupno $p \cdot N + N^2$ parametara za sam model u (2.1), te dodatnih N varijanci i $(N^2 - N)/2$ kovarijanci za matricu Ω . Zato se često može dogoditi da mnoge varijable, tj. njihovi pomaci budu neznačajni pa je cilj pronaći važne veze između varijabli. Valja uočiti da u strukturnom modelu ima više parametara, s obzirom na trenutne učinke svih varijabli u modelu, čime se prethodno navedeni broj povećava za još $N(N-1)$ parametara trenutačnih učinaka (učinaka u razdoblju t). Već je spomenuto da strukturni oblik modela nije moguće procijeniti metodom najmanjih kvadrata, već samo reducirani oblik. No, iz reduciranog oblika modela nije moguće dobiti procjene za strukturni oblik, zbog manjeg broja parametara koji se u reduciranom modelu procjenjuju. To se naziva **problemom identifikacije**. Zato se za dobivanje procjena parametara strukturnog modela moraju nametnuti određena ograničenja na strukturni model (vidjeti odjeljak 2.5.1.2.).

Primjerice, ako se radi o VAR(1) modelu s dvije varijable koji se raspisao u prethodnom odjeljku, strukturni oblik modela sadrži 8 parametara koje je potrebno procijeniti u (2.5), te dvije varijance procesa u_{1t} i u_{2t} , što je ukupno 10 parametara (jer je $\text{Cov}(u_{1t}, u_{2t}) = 0$). No, kako je moguće procijeniti reducirani oblik modela, tj. model u (2.10), u tom modelu se procijeni 2 parametra u vektoru konstanti a_0 , 4 parametra u matrici A_1 , te 3 parametra u matrici varijanci-kovarijanci (2.14) (simetrična je pa zato 3, a ne 4). To je ukupno 9 parametara, a ne potrebnih 10.

Kako je već spomenuto, reducirani oblik VAR modela može se procijeniti metodom najmanjih kvadrata, tako da se svaka jednadžba upravo procijeni kao da se razmatra model s jednom zavisnom varijablom. Procjene dobivene metodom najmanjih kvadrata su konzistentne i asimptotski efikasne (Enders, 2004).

Posebna se napomena odnosi na stacionarnost varijabli. U nekim udžbenicima odmah se pretpostavlja da se razmatraju stacionarne varijable u okviru VAR modela. Međutim, neki

autori (vidjeti popis u Enders, 2004) nisu zagovornici korištenja stacionarnih varijabli, već razmatraju nestacionarne nizove jer im je ideja odrediti veze između varijabli, a ne fokusirati se na procjene parametara. Stoga se diferenciranjem ili detrendiranjem varijabli (da bi postale stacionarne) gube važne informacije. Neke od **prednosti VAR modeliranja** su sljedeće (Brooks, 2008). Sve varijable se razmatraju kao endogene, stoga istraživač ne mora nametati ograničenja koja varijabla je i zašto egzogena u modelu. VAR modeli su fleksibilniji od univarijatnih ARMA modela, jer bogatija struktura modela omogućava „lovljenje“ informacija u podacima, koje univarijatni modeli ne bi uspjeli obuhvatiti. U reduciranom obliku modela je jednostavno procijeniti sve jednadžbe modela, s obzirom da se primjenjuje metoda najmanjih kvadrata za procjenu. Konačno, prognoze dobivene iz VAR modela često su točnije u odnosu na druge strukturne modele.

Osim ovih prednosti, postoje i određeni **nedostaci VAR modela** (Brooks, 2008). Kako se ovi modeli ne temelje na ekonomskoj teoriji, odabir pomaka u tim modelima temelji se na dijagnostici modela, istraživač može pritom dobiti rezultate koji su ekvivalentni prividnoj regresiji (odjeljak 1.10.6.), jer se bavio postupkom rudarenja podataka (engl. *data mining*). Nadalje, javlja se problem interpretacije svih parametara u okviru modela, koji se povećava s povećanjem broja varijabli i pomaka VAR modela. Već je spomenut broj parametara koje je potrebno procijeniti, stoga dolazi do smanjenja stupnjeva slobode za odgovarajuće testove, s povećanjem broja varijabli u modelu i pomaka p . Spomenuta stacionarnost varijabli u okviru modela također predstavlja problem u analizi. Naime, ako želimo provoditi testiranje hipoteza u okviru VAR modela, varijable trebaju biti stacionarne.

2.4. Analiza uzročnosti

2.4.1. Grangerov test uzročnosti

U okviru VAR modela, razmatra se analiza Grangerove uzročnosti, kao i za slučaj ARDL modela (vidjeti odjeljak 1.11.). Kod ARDL modela se polazilo od pretpostavke da je jedna varijabla endogena, a druga egzogena, što se moglo provjeriti pomoću Grangerova testa. Sada u VAR modelu, kada se pretpostavlja da su sve varijable endogene, nakon procjene VAR(p) modela provodi se **multivarijatni Grangerov test**. Ideja ovog testa je slična kao i u odjeljku 1.12., samo što se uključuju i prethodni pomaci ostalih varijabli u modelu.

Sada se VAR(p) model može zapisati u obliku u kojem se razdvajaju varijable, tj. njihovi pomaci na sljedeći način:

$$\begin{aligned}
 y_{1t} &= a_{10} + \sum_{i=1}^p \beta_i y_{1t-i} + \sum_{i=1}^p \gamma_i y_{2t-i} + \dots + \sum_{i=1}^p \zeta_i y_{Nt-i} + \varepsilon_{1t} \\
 y_{2t} &= a_{20} + \sum_{i=1}^p b_i y_{1t-i} + \sum_{i=1}^p c_i y_{2t-i} + \dots + \sum_{i=1}^p z_i y_{Nt-i} + \varepsilon_{2t} \\
 &\vdots \\
 y_{Nt} &= a_{N0} + \sum_{i=1}^p A_i y_{1t-i} + \sum_{i=1}^p B_i y_{2t-i} + \dots + \sum_{i=1}^p Z_i y_{Nt-i} + \varepsilon_{Nt}
 \end{aligned} \quad , \quad (2.29)$$

tako da se „razdvoje“ i „grupiraju“ pomaci odgovarajuće varijable. Ako se u nultoj hipotezi za model (2.29) pretpostavlja da varijabla y_{2t} ne uzrokuje varijablu y_{1t} u Grangerovom smislu, simbolički se to piše: $H_0: \gamma_1 = \gamma_2 = \dots = \gamma_p = 0$, te se provede parcijalni F -test kako bi se odbacila

ili ne ova hipoteza. Ako bi se pretpostavilo u nultoj hipotezi da varijabla y_{2t} ne uzrokuje varijablu y_{1t} u modelu (2.29), nulta hipoteza bi glasila: $H_0: B_1 = B_2 = \dots = B_p = 0$.

Radi jednostavnosti, pretpostavimo da se radi o VAR(1) modelu za dvije varijable:

$$\begin{aligned} y_{1t} &= a_{10} + a_{11}y_{1t-1} + a_{12}y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= a_{20} + a_{21}y_{1t-1} + a_{22}y_{2t-1} + \varepsilon_{2t}. \end{aligned} \quad (2.30)$$

U tom slučaju može se testirati da varijabla y_{2t} ne uzrokuje varijablu y_{1t} u Grangerovom smislu:

$$\begin{aligned} H_0 &: a_{12} = 0 \\ H_1 &: a_{12} \neq 0 \end{aligned} \quad (2.31)$$

ili pak da varijabla y_{1t} ne uzrokuje varijablu y_{2t} u Grangerovom smislu:

$$\begin{aligned} H_0 &: a_{21} = 0 \\ H_1 &: a_{21} \neq 0, \end{aligned} \quad (2.32)$$

dok bi za VAR(2) i slučaj dvije varijable:

$$\begin{aligned} y_{1t} &= a_{10} + a_{11}^1 y_{1t-1} + a_{12}^1 y_{2t-1} + a_{11}^2 y_{1t-2} + a_{12}^2 y_{2t-2} + \varepsilon_{1t} \\ y_{2t} &= a_{20} + a_{21}^1 y_{1t-1} + a_{22}^1 y_{2t-1} + a_{21}^2 y_{1t-2} + a_{22}^2 y_{2t-2} + \varepsilon_{2t} \end{aligned} \quad (2.33)$$

može se testirati da varijabla y_{2t} ne uzrokuje varijablu y_{1t} u Grangerovom smislu:

$$H_0 : a_{12}^1 = a_{12}^2 = 0, \quad (2.34)$$

ili pak da varijabla y_{1t} ne uzrokuje varijablu y_{2t} u Grangerovom smislu:

$$H_0 : a_{21}^1 = a_{21}^2 = 0. \quad (2.35)$$

Primjer 2.5.

Za procijenjene VAR modele u prethodna dva primjera, provedimo Grangerov test uzročnosti za obje varijable u primjerima.

Ako se razmatra VAR(1) model, ispis Grangerova testa prikazan je na slici 2.3., dok je za VAR(2) model rezultat predložen na slici 2.4. Naredba `causality(...)`:

```
causality(m1,cause="s.iip")$Granger,
```

provodi testiranje uzročnosti, pri čemu je potrebno navesti naziv modela za koji se test provodi kao prvi argument, i potom drugi argument kao uzrok (`cause="..."`) navesti naziv varijable čija se uzročnost testira. Za ispis Grangerova testa, dodatno se koristi naredba `$Granger`.

```

causality(m1,cause = "s.iip")$Granger

##
## Granger causality H0: s.iip do not Granger-cause s.nez
##
## data: VAR object m1
## F-Test = 12.744, df1 = 1, df2 = 466, p-value = 0.0003942

causality(m1,cause = "s.nez")$Granger

##
## Granger causality H0: s.nez do not Granger-cause s.iip
##
## data: VAR object m1
## F-Test = 17.241, df1 = 1, df2 = 466, p-value = 3.917e-05

m2<-VAR(matrica,p=2,type="const")

```

Slika 2.3. Grangerov test uzročnosti, VAR(1)

```

causality(m2,cause = "s.iip")$Granger

##
## Granger causality H0: s.iip do not Granger-cause s.nez
##
## data: VAR object m2
## F-Test = 2.8606, df1 = 2, df2 = 460, p-value = 0.05826

causality(m2,cause = "s.nez")$Granger

##
## Granger causality H0: s.nez do not Granger-cause s.iip
##
## data: VAR object m2
## F-Test = 4.7863, df1 = 2, df2 = 460, p-value = 0.008763

```

Slika 2.4. Grangerov test uzročnosti, VAR(2)

Sada kada razmatramo sliku 2.3., za testiranje nulte hipoteze da stopa rasta IIP-a ne uzrokuje stopu rasta nezaposlenosti, odgovarajući empirijski F omjer iznosi 12,74 (stupnjevi slobode u brojniku su 1, zbog isključivanja jednog pomaka varijable stopa rasta IIP-a, dok stupnjevi slobode u nazivniku iznose 466, jer se radi o $2N-2k-2$ stupnja slobode, $N = 236$, a u svakoj jednadžbi je $k = 2$ varijable). Kako je odgovarajuća p -vrijednost jednaka 0,0004 što je manje od razine značajnosti 5%, odbacuje se hipoteza da varijabla stopa rasta IIP-a ne uzrokuje stopu rasta nezaposlenosti u VAR(1) modelu. Ako se testira hipoteza da stopa rasta nezaposlenosti ne uzrokuje stopu rasta IIP-a, odgovarajući empirijski F omjer iznosi 12,24, i pripadajuća p -vrijednost je manja od razine značajnosti 5%, te se i u ovome slučaju odbacuje nulta hipoteza.

S druge strane, ako razmatramo rezultate vezane uz VAR(2) model na slici 2.4., kod testiranja hipoteze da stopa rasta IIP-a ne uzrokuje stopu rasta nezaposlenosti, provodi se parcijalni F test na dva pomaka varijable stope rasta IIP-a. Sada je empirijski F -omjer jednak 2,86, 2 su stupnjeva slobode u brojniku (zbog nulte hipoteze testa, vidjeti(1.34)), dok su u nazivniku 460 (jer se radi o $2N-2k-2$ stupnja slobode, $N = 235$, a u svakoj jednadžbi je $k = 4$ varijabli). Pripadajuća p -vrijednost jednaka je 0,058 što je veće od razine značajnosti 5% (= 0,05). Stoga pri toj razini značajnosti ne možemo odbaciti hipotezu da stopa rasta IIP-a ne uzrokuje stopu rasta nezaposlenosti u VAR(2) modelu. Za testiranje hipoteze da stopa rasta nezaposlenosti ne uzrokuje stopu rasta IIP-a, ako se usporedi test veličina koja iznosi 4,79 s teorijskom, ili pak p -

vrijednost 0,009 s razinom značajnosti 5%, ovdje se pri toj razini značajnosti odbacuje nulta hipoteza.

Napomena: ako se razmatra VAR model s više od 2 varijable, naredba `causality(...)` testira istovremeno hipotezu da neka varijabla ne uzrokuje u Grangerovom smislu sve druge varijable u modelu. Za slučaj testiranja uzrokovanja samo jedne od preostalih varijabli potrebno je najprije u okviru paketa `dynlm` procijeniti svaku jednadžbu modela posebno, a potom u okviru paketa `car` koristiti naredbu `linearHypothesis` koja je već obrađivana u prvome dijelu ovih publikacija.

Napomena: Grangerov test uzročnosti provodi se ako su sve varijable u VAR modelu stacionarne.

2.4.2. Testovi blok-egzogenosti

Multivarijatna generalizacija Grangerove uzročnosti vodi testovima blok-egzogenosti (engl. *block-exogeneity tests*), koji se koriste u svrhu ispitivanja je li poboljšana prognostička moć modela ako se dodaju nove varijable u sam model. Testira se uzrokuje li nova varijabla za koju se pitamo dodati li ju u model ili ne u Grangerovom smislu ostale varijable koje smo već uključili u model. U testu se pretpostavlja da ta nova varijabla može uzrokovati ili prvu varijablu koja je već uključena u model, ili drugu, ..., ili posljednju. Stoga se postojeći VAR model proširi na način da se uključi nova varijabla u model i u nultoj hipotezi se testira neznačajnost njenih pomaka u svakoj od jednadžbi modela. Primjerice, ako opet razmotrimo model u (2.29), za koji smo pretpostavili da je novododana varijabla y_{2t} , nulta hipoteza bi sada glasila:

$$H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_p = B_1 = B_2 = \dots = B_p = 0, \quad (2.36)$$

dakle, u svim jednadžbama (osim vlastite) se testira neznačajnost svih pomaka varijable y_{2t} . Test veličina u ovome testu slijedi hi-kvadrat distribuciju, jer se računa sljedećim izrazom:

$$LR = (T-c)[\ln|\mathbf{\Omega}_r| - \ln|\mathbf{\Omega}|] \sim \chi^2(2p), \quad (2.37)$$

gdje $|\cdot|$ predstavlja determinantu matrice, $\mathbf{\Omega}_r$ je matrica varijanci-kovarijanci inovacija modela s ograničenjem, $\mathbf{\Omega}$ je matrica varijanci-kovarijanci inovacija modela bez ograničenja, T je broj opservacija, dok je c broj parametara u svakoj jednadžbi modela bez ograničenja. Broj stupnjeva slobode $2p$ jednak je broju pomaka koje smo isključili u modelu s ograničenjem (broju parametara u nultoj hipotezi (2.36)).

2.5. Inovacijska analiza

Jedna od svrha VAR modela jest utvrditi međusobne odnose između varijabli, na način da se konstruiraju funkcije impulsnog odziva (engl. *impulse response function*, vidjeti odjeljke 2.9.5. i 2.12.1.) i da se provodi analiza dekompozicije varijance (engl. *decomposition of variance*). Naziv inovacijska analiza se temelji na analizi posljedica kada se dogodi šok u nekoj od varijabli u modelu, koji se očituje u inovacijskom procesu ε_{jt} .

2.5.1. Funkcija impulsnog odziva

2.5.1.1. Definiranje funkcije impulsnog odziva

U okviru VAR modela, funkcija impulsnog odziva definira se kao vremenski put učinaka egzogenog šoka u jednoj ili više endogenih varijabli ili drugim varijablama u okviru modela. Prethodno je spomenuto da **stacionaran VAR(1) model ima MA(∞) reprezentaciju**¹⁵:

$$y_t = \mu + \sum_{i=0}^{\infty} A_1^i \varepsilon_{t-i} = \mu + \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i}, \quad (2.38)$$

koju je moguće za slučaj dvije varijable raspisati na sljedeći način:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \theta_{11}(i) & \theta_{12}(i) \\ \theta_{21}(i) & \theta_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-i} \\ \varepsilon_{2t-i} \end{bmatrix}. \quad (2.39)$$

U odjeljku 2.1.2. je vektor ε_t bio definiran na sljedeći način (formula (2.110)):

$$\varepsilon_t = B^{-1} u_t, \quad (2.40)$$

tj.

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} \frac{u_{1t} - b_{12}u_{2t}}{1 - b_{12}b_{21}} \\ \frac{u_{2t} - b_{21}u_{1t}}{1 - b_{12}b_{21}} \end{bmatrix}. \quad (2.41)$$

Sada, ako se supstituira (2.40) u (2.38), odnosno (2.41) u (2.39), dobivamo zapise:

$$y_t = \mu + \sum_{i=0}^{\infty} \underbrace{\Theta_i B^{-1}}_{=\Phi_i} u_{t-i} = \mu + \sum_{i=0}^{\infty} \Phi_i u_{t-i} \quad (2.42)$$

i

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \theta_{11}(i) & \theta_{12}(i) \\ \theta_{21}(i) & \theta_{22}(i) \end{bmatrix} \begin{bmatrix} \frac{u_{1t} - b_{12}u_{2t}}{1 - b_{12}b_{21}} \\ \frac{u_{2t} - b_{21}u_{1t}}{1 - b_{12}b_{21}} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} u_{1t-i} \\ u_{2t-i} \end{bmatrix}, \quad (2.43)$$

gdje je

$$\phi_{jk}(i) = \frac{\theta_{jk}(i)}{1 - b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}, \quad j, k \in \{1, 2\}. \quad (2.44)$$

Upravo se koeficijenti u matrici Φ_i , odnosno koeficijenti u (2.44) koriste za formiranje funkcije impulsnog odziva: razmotri se jedinični šok u inovacijama u_{jt} te se pomoću koeficijenata u Φ_i

¹⁵ Može se pokazati da osim obrađenog VAR(1) modela koji ima MA(∞) reprezentaciju, i VAR(p) možemo predočiti kao poseban slučaj VAR(1) modela, koji će također imati MA(∞) reprezentaciju, vidjeti Lütkepohl (2006, 2010).

razmatra reakcija varijabli u y_t na taj šok. Potrebno je bilo razmotriti zapis modela u (2.42) jer smo u odjeljku 2.1.2. pokazali da su procesi u ε_t korelirani, dok oni u u_t nisu, što je veoma važno kod razmatranja funkcije impulsnog odziva. Koeficijenti u (2.44), odnosno općenito u matrici Φ , nazivamo **multiplikatorima učinaka** (engl. *impact multipliers*). Općenito se $\phi_{jk}(i)$ interpretira na sljedeći način: učinak jedinične promjene inovacije u_k na varijablu y_j u razdoblju i . Upravo niz ovih koeficijenata čini funkciju impulsnog odziva, koja se predočava grafički radi lakše interpretacije.

Osim ovog učinka, kroz vrijeme se može razmatrati i kumulativna funkcija impulsnog odziva, na način da se zbroje učinci $\phi_{jk}(i)$ kroz vrijeme: $\sum_{i=0}^n \phi_{jk}(i)$. Ovaj zbroj naziva **multiplikator dugog roka** (engl. *long-run multiplier*) za slučaj $n \rightarrow \infty$. Dodatno, pretpostavlja se da se razmatraju stacionarne varijable u y_t , stoga za $n \rightarrow \infty$ mora vrijediti $\sum_{i=0}^n \phi_{jk}^2(i) < \infty$ (konačna je).

2.5.1.2. Cholesky dekompozicija

Problem koji nastaje za procjenu koeficijenata $\phi_{jk}(i)$, odnosno kad se razmatra model u (2.42):

$y_t = \mu + \sum_{i=0}^{\infty} \Phi_i u_{t-i}$, jest taj što temeljem reduciranog oblika modela ne možemo procijeniti sve parametre u strukturnom obliku, stoga nije moguće procijeniti $\phi_{jk}(i)$ iz reduciranog modela. Kako je spomenuto u odjeljku 2.3., za dobivanje procjena parametara strukturnog modela moraju se nametnuti određena ograničenja na strukturni model.

Jedno od mogućih rješenja je **Cholesky dekompozicija**. Razmotrimo najprije problem VAR(1) modela za dvije varijable, u strukturnom i reduciranom obliku (vidjeti odjeljak 2.1.2.), usporedimo parametre modela kako bi nakon toga definirali Cholesky dekompoziciju koju je potrebno napraviti da bismo mogli dobiti parametre strukturnog modela.

U VAR(1) strukturnom obliku modela je potrebno u dvije jednadžbe procijeniti 8 parametara, s dodatna 2, koja se odnose na varijance u_{1t} i u_{2t} (jer je $\text{Cov}(u_{1t}, u_{2t}) = 0$), što je ukupno 10. Reducirani oblik modela rezultira s procjenom 6 parametara u dvije jednadžbe, s dodatna 3, koja se odnose na varijance ε_{1t} i ε_{2t} i $\text{Cov}(\varepsilon_{1t}, \varepsilon_{2t})$, dakle ukupno 9. Zato moramo nametnuti neko ograničenje na strukturni oblik modela, na način da neki parametar izjednačimo s 0 i na taj način ćemo imati ukupno 9 parametara koje je moguće procijeniti iz reduciranog oblika modela. Postavlja se pitanje kako nametnuti to ograničenje?

Pretpostavimo slučaj da u strukturnom modelu:

$$\begin{aligned} y_{1t} &= b_{10} - b_{12}y_{2t} + c_{11}y_{1t-1} + c_{12}y_{2t-1} + u_{1t} \\ y_{2t} &= b_{20} - b_{21}y_{2t} + c_{21}y_{1t-1} + c_{22}y_{2t-1} + u_{2t} \end{aligned}$$

vrijedi $b_{21} = 0$:

$$\begin{aligned} y_{1t} &= b_{10} - b_{12}y_{2t} + c_{11}y_{1t-1} + c_{12}y_{2t-1} + u_{1t} \\ y_{2t} &= b_{20} + c_{21}y_{1t-1} + c_{22}y_{2t-1} + u_{2t}. \end{aligned} \quad (2.45)$$

To znači da pretpostavljamo da samo varijabla y_2 ima trenutni učinak na varijablu y_1 u istom razdoblju t ($-b_{12}$), dok promjene u varijabli y_1 nemaju trenutni učinak na varijablu y_2 .

Izvođenjem reduciranog modela kao u odjeljku 2.1.2., rezultiralo bi s:

$$\underbrace{\begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}}_{\mathbf{B}} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \underbrace{\begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}}_{\mathbf{b}} + \underbrace{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}}_{\mathbf{C}} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \quad (2.46)$$

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \quad (2.47)$$

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} b_{10} - b_{12}b_{20} \\ b_{20} \end{bmatrix} + \begin{bmatrix} c_{11} - b_{12}c_{21} & c_{12} - b_{12}c_{22} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} - b_{12}u_{2t} \\ u_{2t} \end{bmatrix} \quad (2.48)$$

Model (2.48) se procijeni metodom najmanjih kvadrata:

$$\begin{aligned} y_{1t} &= a_{10} + a_{11}y_{1t-1} + a_{12}y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= a_{20} + a_{21}y_{1t-1} + a_{22}y_{2t-1} + \varepsilon_{2t}, \end{aligned} \quad (2.49)$$

gdje je $a_{10} = b_{10} - b_{12}b_{20}$, $a_{20} = b_{20}$, $a_{11} = c_{11} - b_{12}c_{21}$, $a_{12} = c_{12} - b_{12}c_{22}$, $a_{21} = c_{21}$, $a_{22} = c_{22}$, te za izračun varijanci ε_{1t} i ε_{2t} i $\text{Cov}(\varepsilon_{1t}, \varepsilon_{2t})$, prema formulama (2.12) i (2.13) vrijedi:

$$\begin{bmatrix} \sigma_{\varepsilon_1}^2 \\ \sigma_{\varepsilon_2}^2 \end{bmatrix} = \begin{bmatrix} \sigma_{u_1}^2 + b_{12}^2 \sigma_{u_2}^2 \\ \sigma_{u_2}^2 \end{bmatrix} \quad \text{i} \quad \sigma_{\varepsilon_1 \varepsilon_2} = \text{Cov}(\varepsilon_{1t}, \varepsilon_{2t}) = -b_{12} \sigma_{u_2}^2. \quad (2.50)$$

I sada je temeljem 9 procijenjenih parametara reduciranog modela moguće izračunati 9 parametara strukturnog modela, s obzirom na 9 nepoznanica i 9 jednadžbi.

S obzirom na ograničenje $b_{21} = 0$, pretpostavlja se da šokovi u u_{1t} i u_{2t} imaju trenutni učinak u y_{1t} , ali samo šok u u_{2t} ima trenutni učinak na y_{2t} . Uočimo da je zato matrica \mathbf{B} u (2.46) gornja trokutasta matrica:

$$\mathbf{B} = \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}, \quad (2.51)$$

kao i njezin inverz:

$$\mathbf{B}^{-1} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix}, \quad (2.52)$$

i tehnika dekomponiranja reziduala u VAR modelu na ovaj način **naziva se dekompozicija Choleskyog** (engl. *Cholesky decomposition, triangularization*). Pretpostavka u (2.51), odnosno (2.52) je važna za matricu varijanci-kovarijanci $\mathbf{\Omega}$ u (2.14):

$$\mathbf{\Omega} = \mathbf{B}^{-1} \begin{bmatrix} \sigma_{u_1}^2 & 0 \\ 0 & \sigma_{u_2}^2 \end{bmatrix} (\mathbf{B}^{-1})', \quad (2.53)$$

odnosno u (2.15) za slučaj N varijabli:

$$\mathbf{\Omega} = \mathbf{B}^{-1}\mathbf{\Omega}_u(\mathbf{B}^{-1})' = \mathbf{\Omega}_u^{1/2}(\mathbf{\Omega}_u^{1/2})', \quad (2.54)$$

gdje je $\mathbf{\Omega}_u = \begin{bmatrix} \sigma_{u1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{u2}^2 & \dots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{uN}^2 \end{bmatrix}$, a $\mathbf{\Omega}_u^{1/2}$ se interpretira kao drugi korijen pozitivno

semidefinitne matrice. Matrica $\mathbf{\Omega}_u^{1/2}$ je **faktor Choleskyog** (engl. *Cholesky factor*). Za slučaj dvije varijable možemo (2.53) raspisati kao:

$$\mathbf{\Omega} = \begin{bmatrix} \sigma_{\varepsilon 1}^2 & \sigma_{\varepsilon 1 \varepsilon 2} \\ \sigma_{\varepsilon 1 \varepsilon 2} & \sigma_{\varepsilon 2}^2 \end{bmatrix} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_{u1}^2 & 0 \\ 0 & \sigma_{u2}^2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -b_{12} & 1 \end{bmatrix} = \begin{bmatrix} \sigma_{u1}^2 + b_{12}^2 \sigma_{u2}^2 & -b_{12} \sigma_{u2}^2 \\ -b_{12} \sigma_{u2}^2 & \sigma_{u2}^2 \end{bmatrix}, \quad (2.55)$$

što smo upravo dobili u (2.50). Stoga se izraz u (2.53), odnosno u (2.54) naziva dekompozicija Choleskyog simetrične matrice $\mathbf{\Omega}$. U praksi želimo iz procijenjene matrice $\mathbf{\Omega}$ doći do $\mathbf{\Omega}_u$, pa ako vrijedi u (2.54):

$$\mathbf{\Omega} = \mathbf{B}^{-1}\mathbf{\Omega}_u(\mathbf{B}^{-1})', \quad (2.56)$$

slijedi:

$$\mathbf{\Omega}_u = \mathbf{B}\mathbf{\Omega}(\mathbf{B}')^{-1}. \quad (2.57)$$

Zato je a VAR s N varijabli, \mathbf{B} je $N \cdot N$ matrica, **potrebno je nametnuti $(N^2 - N)/2$ ograničenja da bi se identificirao strukturni model iz procijenjenog reduciranog modela.**

Primjer 2.6.

Prikažimo faktorizaciju Choleskyog za sljedeću matricu:

$$\mathbf{\Omega} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} \sigma_{z_1}^2 & Cov(z_1, z_2) \\ Cov(z_1, z_2) & \sigma_{z_2}^2 \end{bmatrix},$$

uz pretpostavku ograničenja kao u (2.51).

Potrebno je odrediti matricu $\mathbf{\Omega}_u^{1/2} = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$, $a, b, c \in \mathbb{R}$ takvu da je:

$$\mathbf{\Omega} = \mathbf{\Omega}_u^{1/2}(\mathbf{\Omega}_u^{1/2})' \Rightarrow \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} = \begin{bmatrix} a^2 + b^2 & bc \\ bc & c^2 \end{bmatrix}$$

Pa vrijedi: $a^2 + b^2 = 2$, $bc = 1$, $c^2 = 4$. Uz pretpostavku da su svi elementi na glavnoj dijagonali matrice $\Omega_u^{1/2}$ pozitivni, slijedi da je $a = \frac{\sqrt{7}}{2}$, $c = 2$ pa je $b = \frac{1}{2}$ i matrica $\Omega_u^{1/2}$ je sljedeća: $\Omega_u^{1/2}$

$$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{7}}{2} & \frac{1}{2} \\ 0 & 2 \end{bmatrix}. \text{ Dakle, } \Omega = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \sqrt{2} & -1 \\ 0 & 2 \end{bmatrix}.$$

Napomena: S druge strane, da smo pretpostavili u matrici B ograničenje $b_{12} = 0$, tj.

$$B = \begin{bmatrix} 1 & 0 \\ b_{21} & 1 \end{bmatrix}, \quad (2.58)$$

to bi značilo da pretpostavljamo da samo varijabla y_1 ima trenutni učinak na varijablu y_2 u istom razdoblju t ($-b_{21}$), dok promjene u varijabli y_2 nemaju trenutni učinak na varijablu y_1 . U tom slučaju bi matrica Ω bila sljedeća:

$$\Omega = \begin{bmatrix} \sigma_{\varepsilon 1}^2 & \sigma_{\varepsilon 1 \varepsilon 2} \\ \sigma_{\varepsilon 1 \varepsilon 2} & \sigma_{\varepsilon 2}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \sigma_{u1}^2 & 0 \\ 0 & \sigma_{u2}^2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \sigma_{u1}^2 & -b_{21}\sigma_{u1}^2 \\ -b_{21}\sigma_{u1}^2 & \sigma_{u2}^2 - b_{21}^2\sigma_{u1}^2 \end{bmatrix}, \quad (2.59)$$

i razlikuje se od one u (2.55).

Primjer 2.2.

Prikažimo faktorizaciju Choleskyog za sljedeću matricu, u kojoj će varijable zamijeniti mjesta:

$$\Omega = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} \sigma_{z2}^2 & Cov(z_1, z_2) \\ Cov(z_1, z_2) & \sigma_{z1}^2 \end{bmatrix},$$

$$\text{odnosno } \Omega = \Omega_u^{1/2} (\Omega_u^{1/2})', : \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}', \Rightarrow \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} a^2 + b^2 & bc \\ cb & c^2 \end{bmatrix}$$

dakle je $a^2 + b^2 = 4$, $bc = 1$, $c^2 = 2$. Uz pretpostavku da su svi elementi na glavnoj dijagonali matrice $\Omega_u^{1/2}$ pozitivni, slijedi da je $c = \sqrt{2}$ pa je $b = \frac{\sqrt{2}}{2}$ i $a = \frac{\sqrt{7}}{2}$. Matrica $\Omega_u^{1/2}$ je sljedeća:

$$\Omega_u^{1/2} = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{7}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \sqrt{2} \end{bmatrix}.$$

Napomena: Zato se rezultati dekompozicije razlikuju u ovisnosti o poretku varijabli u VAR modelu. To znači da odabir poretka varijabli rezultira s različitim Cholesky dekompozicijama, koje se koriste za konstrukciju funkcija impulsnog odziva u (2.42) i rezultati se mogu razlikovati, o čemu ovisi interpretacija tih funkcija.

Napomena: U slučaju konstrukcije funkcija impulsnog odziva je potrebno ispravno odrediti pomak varijabli, za korektnu provedbu Cholesky dekompoziciju. Varijabla koja nema prediktivnog utjecaja na ostale varijable se stavlja posljednja u poretku.

Napomena: S obzirom na prethodnu napomenu, preporuča se provesti testiranje uzročnosti u Grangerovom smislu (odjeljak 2.4.), i temeljem rezultata tog testa napraviti poredak varijabli za konstrukciju funkcija impulsnih odziva. Također se može temeljem ekonomske teorije odrediti poredak varijabli.

Primjer 2.8.

Iz primjera 2.3. gdje smo procijenili VAR(1) model za varijable stopa rasta IIP-a i stopa rasta nezaposlenosti, i temeljem rezultata Grangerova testa uzročnosti u primjeru 2.5., odredimo funkcije impulsnog odziva uz dekompoziciju Choleskyog.

Kako smo u primjeru 2.5. (slika 2.3.) utvrdili da se odbacuje nulta hipoteza da pojedina varijabla ne uzrokuje drugu, ovdje nije moguće strogo odrediti poredak temeljem tog rezultata. Ovdje bi istraživač temeljem ekonomske teorije trebao odrediti poredak varijabli. U nastavku ćemo prikazati da dolazi do značajnih promjena u ponašanju funkcija impulsnog odziva u ovisnosti o promjeni poretka varijabli.

Za slučaj već procijenjenog modela u primjeru 2.5. (vidjeti poredak varijabli u matrici, da je jednak s.iip, s.nez) naredba `irf(...)` u RStudiju računa koeficijente u (2.44) za slučaj dvije varijable koje razmatramo:

```
irf(m1,n.ahead=4,ortho=T,ci=0.95,impulse="s.iip",response="s.nez")
```

U samoj naredbi je potrebno upisati sljedeće argumente: naziv modela za koji se procjenjuju koeficijenti funkcije impulsnog odziva (`m1`), broj razdoblja za koje želimo procijeniti (`n.ahead`) te koeficijente, želimo li provesti dekompoziciju Choleskyog ili ne (`ortho = T` ili `F`, `T` za točno, `F` za netočno), te potom želimo li procijeniti i granice pouzdanosti i pri kojoj razini (CI – engl. *confidence intervals*, uz razinu pouzdanosti, u primjeru je 95%). Konačno, potrebno je odrediti u kojoj varijabli se provodi šok (`impulse`), i za koju varijablu promatramo funkciju impulsnog odziva kroz vrijeme (`response`). Za primjer koji se ovdje razmatra su sljedeće naredbe (vidjeti sliku 2.5.): naziv modela je `m1`, broj razdoblja je 4, želimo provesti dekompoziciju Choleskyog (`ortho=T`), vršimo šok u varijabli stopa rasta IIP-a, a razmatramo funkciju impulsnog odziva u stopi rasta nezaposlenosti (naziv `irf1`). Slično je definirano i za `irf2`, funkciju impulsnog odziva stope rasta IIP-a.

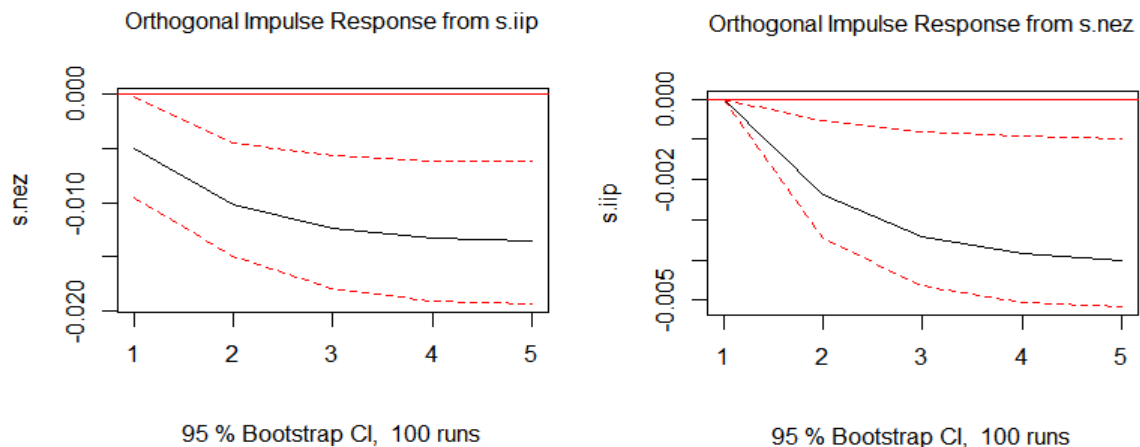
```
irf1<-irf(m1,n.ahead=4,ortho=T,ci=0.95,impulse="s.iip",response="s.nez")
irf2<-irf(m1,n.ahead=4,ortho=T,ci=0.95,impulse="s.nez",response="s.iip")
plot(irf1)
plot(irf2)
```

Slika 2.5. Definiranje i grafičko predočavanje funkcija impulsnog odziva u VAR modelu

Slika 2.6. predočava funkcije impulsnih odziva temeljem definiranih naredbi na slici 2.5., pri čemu je lijevi panel `irf1`, dok je desni `irf2`. Crvene iscrtane linije su granice pouzdanosti¹⁶ (engl.

¹⁶ Kako CI granice predstavljaju 95%-tni interval pouzdanosti procijenjenih koeficijenata funkcije impulsnih odziva, u analizama se razmatra je li vrijednost 0 obuhvaćena tim intervalom. Naime, ako je 0 obuhvaćena, onda se rezultati interpretiraju na način da nisu značajni, jer šok u jednoj varijabli vodi do reakcije 0 druge varijable, odnosno interpretacije da druga varijabla ne reagira na šok u prvoj varijabli, pri razini pouzdanosti od 5%.

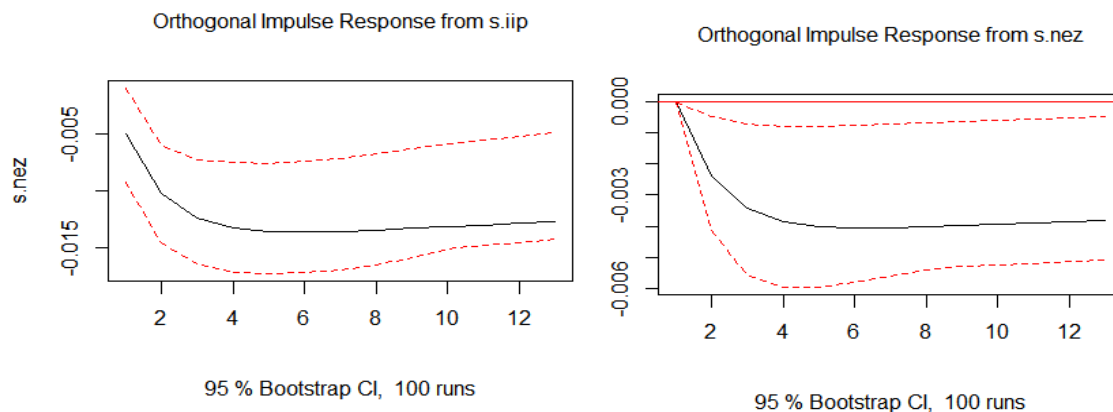
CI – *confidence intervals*) pri razini 95%. Funkcije impulsnih odziva interpretiraju se slično kao i u slučaju univarijantnih modela razmatranih u poglavlju 5. Lijevi panel interpretiramo na sljedeći način. Pretpostavljeno je da je učinjen šok u varijabli stopa rasta IIP-a i razmatramo reakciju stope rasta nezaposlenosti. Kako varijable u VAR modelu mogu biti u različitim mjernim jedinicama, uobičajena procedura u paketima koji procjenjuju i računaju koeficijente funkcije impulsnog odziva jest standardizacija varijabli u kojima se dogodi šok. **Pa se interpretacija vrši na način da se pretpostavi šok od jedne standardne devijacije u nekoj varijabli, a razmatra se reakcija drugih varijabli u njihovim mjernim jedinicama.** Naravno, moguće je i reakcije (tj. procijenjene koeficijente funkcije impulsnog odziva) također standardizirati kako bi se dobila potpunija slika.



Slika 2.6. Funkcije impulsnog odziva u VAR(1) modelu, poredak varijabli: stopa rasta IIP-a, stopa rasta nezaposlenosti

Lijevi panel na slici 2.6. govori da šok od jedne standardne devijacije u varijabli stopa rasta IIP-a rezultira sa smanjenjem stope rasta nezaposlenosti u idućim razdobljima, pri čemu se taj učinak stabilizira oko vrijednosti $-0,015$. Kako se radi o stopi rasta, vrijednosti predočene crnom linijom se interpretiraju kao promjene stope rasta nezaposlenosti u postotnim bodovima (stopa rasta računata je kao prirodni logaritam omjera nezaposlenosti u mjesecu t i $t-12$, pa je potrebno pomnožiti sa 100% ako se želi vršiti interpretacija u postotnim bodovima). Slično tome, desni panel na slici 2.6. treba interpretirati da šok od jedne standardne devijacije u stopi rasta nezaposlenosti rezultira sa smanjenjem stope rasta IIP-a, pri čemu se to smanjenje stabilizira oko vrijednosti $-0,004$ u posljednjem razdoblju.

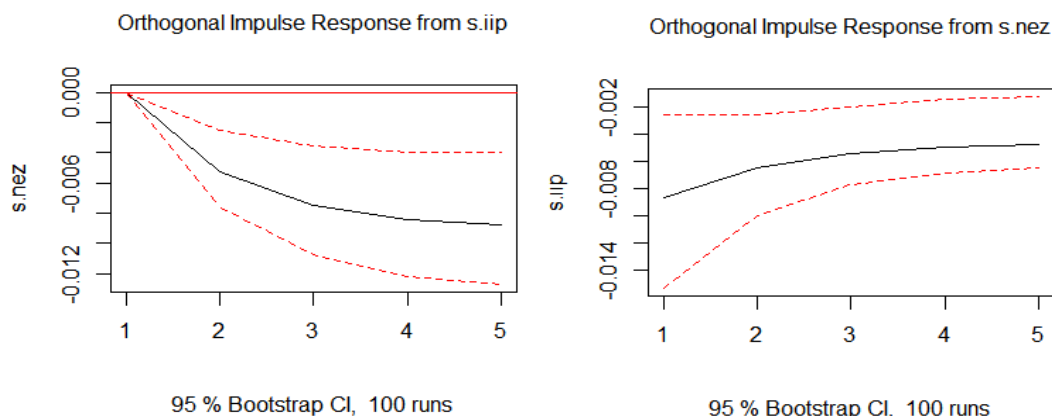
Predočimo funkcije impulsnog odziva sa slike 2.6. za idućih 12 razdoblja, umjesto 4, kako bismo lakše uočili reakciju na šokove u obje varijable. Slika 2.2. predočava rezultate, dobivene temeljem naredbi sa slike 2.5. (samo je izmijenjena vrijednost 4 u 12). Obično se za mjesečne podatke u analizama odabire 12 ili 24 mjeseca za predočavanje funkcija impulsnog odziva, dok se za tromjesečne podatke odabire 8, 12 ili 16 tromjesečja.



Slika 2.7. Funkcije impulsnog odziva u VAR(1) modelu, poredak varijabli: stopa rasta IIP-a, stopa rasta nezaposlenosti

```
matrica2<-cbind(s.nez,s.iip)
m3<-VAR(matrica2,p=1,type="const")
irf3<-irf(m3,n.ahead=4,ortho=T,ci=0.95,impulse="s.iip",response="s.nez")
irf4<-irf(m3,n.ahead=4,ortho=T,ci=0.95,impulse="s.nez",response="s.iip")
plot(irf3)
plot(irf4)
```

Slika 2.8. Definiranje i grafičko predočavanje funkcija impulsnog odziva u VAR modelu, promijenjen redoslijed varijabli



Slika 2.9. Funkcije impulsnog odziva u VAR(1) modelu, poredak varijabli: stopa rasta nezaposlenosti, stopa rasta IIP-a

Nadalje, ako izmijenimo poredak varijabli u matrici za procjenu VAR modela (vidjeti sliku 2.8., gdje je poredak s.nez pa potom s.iip), dolazi do određenih promjena. Iako će se procijeniti identični parametri za svaku jednadžbu, jer poredak nije bitan za procjenu parametara u VAR modelu, poredak utječe na dekompoziciju Choleskyog, što se uočava na slici 2.9. gdje su predočene funkcije impulsnog odziva za ovaj novi poredak. Sada je reakcija stope rasta nezaposlenosti manji u odnosu na prethodni poredak (usporediti lijeve panele na slikama 2.9. i 2.6.), te je i drugačija reakcija stope rasta IIP-a na šok u varijabli stopa rasta nezaposlenosti: najjači učinak je u prvome mjesecu te se potom ublažava na slici 2.9. (desni panel), dok je prvotna reakcija na slici 2.6. bila 0 (odnosno nije bilo trenutne reakcije), koja se nakon toga smanjivala prema $-0,004$. Zato valja biti oprezan pri definiranju poretka varijabli u VAR modelu.

Primjer 2.9.

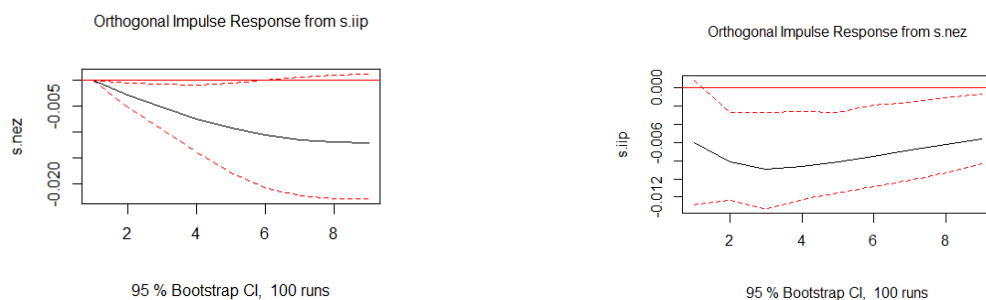
Temeljem procijenjenog VAR(2) modela u primjeru 2.4. i rezultata Grangerova testa uzročnosti u primjeru 2.5., predočimo i komentirajmo funkcije impulsnog odziva za idućih 12 mjeseci, uz dekompoziciju Choleskyog.

Kako smo u primjeru 2.5. zaključili da odbacujemo hipotezu da stopa rasta IIP-a ne uzrokuje stopu rasta nezaposlenosti pri 5% značajnosti, dok odbacujemo hipotezu da stopa rasta nezaposlenosti ne uzrokuje stopu rasta IIP-a, ovdje biramo poredak varijabli tako da na prvo mjesto biramo varijablu stopu rasta nezaposlenosti, a na drugo stopu rasta IIP-a. Slike 2.10. i 2.11. predočavaju potrebne naredbe i prikaz funkcija impulsnog odziva. Sada se uočava da šok od jedne standardne devijacije u varijabli stopa rasta IIP-a rezultira s negativnom reakcijom stope rasta nezaposlenosti, koja već u trećem mjesecu postaje neznačajna (s obzirom da je vrijednost 0 uključena u 95%-tni interval pouzdanosti). S druge strane, šok od jedne standardne devijacije u stopi rasta nezaposlenosti rezultira sa smanjenjem stope rasta IIP-a, pri čemu učinak u prvome mjesecu nije značajan, a za razdoblje od drugog do osmog mjeseca od šoka je značajan, nakon čega ponovno postaje neznačajan.

```
mat2<-cbind(s.nez,s.iip)
var2<-VAR(mat2,p=2,type="const")
irf5<-irf(var2,n.ahead=8,ortho=T,ci=0.95,impulse="s.iip",response="s.nez")
irf6<-irf(var2,n.ahead=8,ortho=T,ci=0.95,impulse="s.nez",response="s.iip")
plot(irf5)
plot(irf6)
```

Slika 2.10. Definiranje i grafičko predočavanje funkcija impulsnog odziva u VAR(2) modelu

Napomena: ako se radi o stabilnom VAR modelu, funkcije impulsnog odziva će predočavati kratkoročan učinak šokova u jednoj varijabli na druge, nakon čega će taj šok nestati. Nadalje, razmatranjem kumulativnog učinka pomoću kumulativne funkcije impulsnog odziva (u naredbi irf dodaje se cumulative=T naredba), uočit će se u stabilnim modelima stabiliziranje učinka u dugome roku.



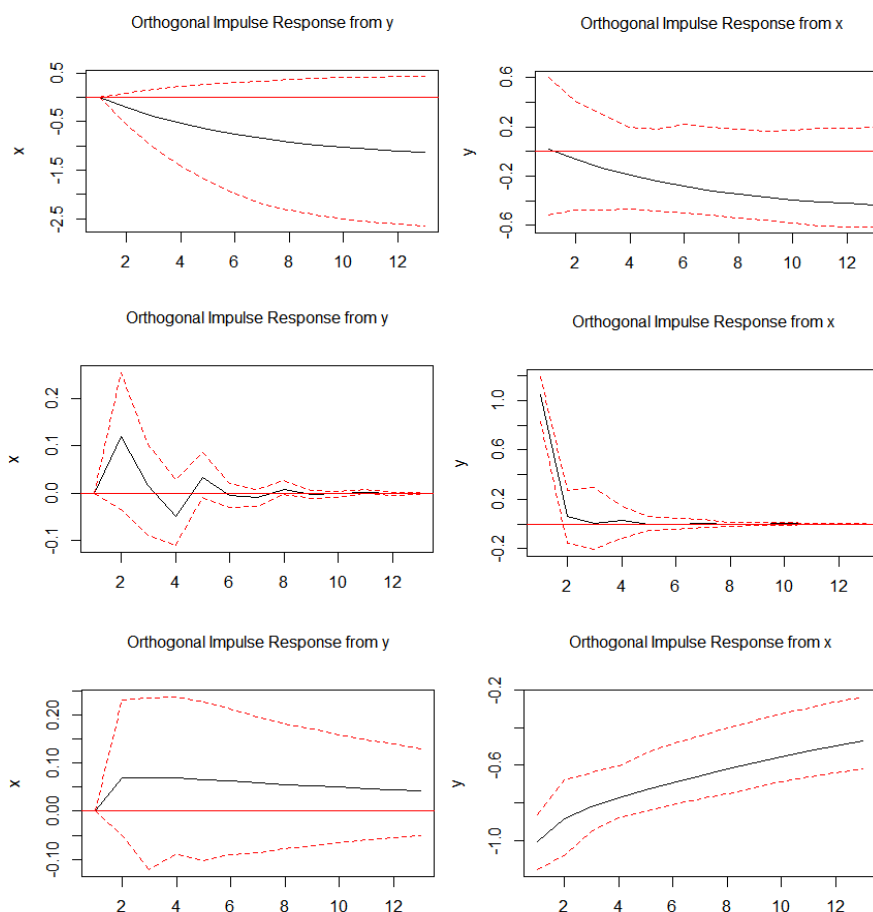
Slika 2.10. Funkcije impulsnog odziva u VAR(2) modelu

Primjer 2.10.

Interpretirajmo funkcije impulsnih odziva za parove varijabli x i y predočene na panelima na slici 2.11.

Prvi panel na slici 2.11. predočava sljedeće: šok od 1 standardne devijacije u varijabli y rezultira sa smanjenjem varijable x u svim razdobljima, no taj učinak nije značajan. Šok od 1 standardne devijacije u varijabli x rezultira sa smanjenjem varijable y , no također učinci nisu značajni. Drugi panel predočava neznačajne učinke šoka u y u reakciji x , dok šok u varijabli x rezultira s pozitivnom reakcijom varijable y u prvom razdoblju, no već u drugom postaje neznačajan. Treći

panel predočava neznačajnu reakciju x na šok u varijabli y , dok s druge strane šok od 1 standardne devijacije u varijabli x rezultira s negativnom reakcijom varijable y , koja se smanjuje u svim idućim razdobljima, te su jedino koeficijenti funkcije impulsnog odziva ovdje značajni u svim razdobljima koja se razmatraju.



Slika 2.11. Funkcije impulsnog odziva za parove varijabli x i y

2.5.2. Dekompozicija varijance

Osim funkcije impulsnog odziva, u inovacijsku analizu uključujemo i dekompoziciju varijance (engl. *variance decomposition*), u okviru koje se **razmatra koliki udio prognostičkih pogrešaka su objašnjeni šokovima u svim varijablama**. Varijanca svake varijable u modelu može se raščlaniti na onaj dio koji je uzrokovan šokom u samoj toj varijabli, te dijelovima koji su zbog šokova u drugim varijablama. Polazi se ponovno od modela (2.42):

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\Phi}_i \mathbf{u}_{t-i}, \quad (2.60)$$

temeljem kojeg se mogu predviđati vrijednosti \mathbf{y}_{t+h} :

$$\mathbf{y}_{t+h} = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\Phi}_i \mathbf{u}_{t+h-i}, \quad h \in \{1, 2, \dots\}. \quad (2.61)$$

Tako je predviđena vrijednost u razdoblju $\tau \in \{1, 2, \dots\}$ jednaka

$$\mathbf{y}_{t+\tau} = \boldsymbol{\mu} + \sum_{i=0}^{\tau-1} \boldsymbol{\Phi}_i \mathbf{u}_{t+\tau-i}, \quad (2.62)$$

i pogreška predviđanja $\varepsilon_t(\tau)$ za razdoblje τ iznosi:

$$\varepsilon_t(\tau) = \mathbf{y}_{t+\tau} - E(\mathbf{y}_{t+\tau}) = \sum_{i=0}^{\tau-1} \boldsymbol{\Phi}_i \mathbf{u}_{t+\tau-i}. \quad (2.63)$$

Za lakše razumijevanje, vratimo se na bivarijatni VAR(1) model. **Pogreška predviđanja za varijablu y_{1t} u razdoblju τ** (engl. *the τ -step-ahead forecast error*) iznosi:

$$\begin{aligned} \varepsilon_{1t}(\tau) &= y_{1t+\tau} - E(y_{1t+\tau}) = \\ &= \phi_{11}(0)u_{1t+\tau} + \phi_{11}(1)u_{1t+\tau-1} + \dots + \phi_{11}(\tau-1)u_{1t+1} + \phi_{12}(0)u_{2t+\tau} + \phi_{12}(1)u_{2t+\tau-1} + \dots + \phi_{12}(\tau-1)u_{2t+1} \end{aligned} \quad (2.64)$$

Sada, kada računamo varijancu izraza (2.64), odnosno **varijancu pogreške predviđanja u razdoblju τ** (engl. *the τ -step-ahead variance of the forecast of $y_{1,t+\tau}$*), ona je jednaka:

$$\begin{aligned} \text{Var}[\varepsilon_{1t}(\tau)] &= \sigma_{y_1}^2(\tau) \\ &= \phi_{11}^2(0)\sigma_{u_1}^2 + \phi_{11}^2(1)\sigma_{u_1}^2 + \dots + \phi_{11}^2(\tau-1)\sigma_{u_1}^2 + \phi_{12}^2(0)\sigma_{u_2}^2 + \phi_{12}^2(1)\sigma_{u_2}^2 + \dots + \phi_{12}^2(\tau-1)\sigma_{u_2}^2 \\ &= \sigma_{u_1}^2 [\phi_{11}^2(0) + \phi_{11}^2(1) + \dots + \phi_{11}^2(\tau-1)] + \sigma_{u_2}^2 [\phi_{12}^2(0) + \phi_{12}^2(1) + \dots + \phi_{12}^2(\tau-1)] \end{aligned} \quad (2.65)$$

Uočimo da zbog toga što su varijance u (2.65) pozitivne, kao i što su vrijednosti koeficijenata $\phi_{jk}^2(\tau)$ nenegativne, varijanca pogreške predviđanja je veća ako se povećava horizont predviđanja (odnosno vrijednost τ).

Sada je moguće **dekomponirati varijancu $\sigma_{y_1}^2(\tau)$** u (2.65) na dio koji je objašnjen šokovima u samoj varijabli y_{1t} i dio koji je objašnjen šokovima u varijabli y_{2t} :

$$\sigma_{y_1}^2(\tau) = \underbrace{\sigma_{u_1}^2 [\phi_{11}^2(0) + \phi_{11}^2(1) + \dots + \phi_{11}^2(\tau-1)]}_{\text{rezultat zbog šokova u } y_{1t}} + \underbrace{\sigma_{u_2}^2 [\phi_{12}^2(0) + \phi_{12}^2(1) + \dots + \phi_{12}^2(\tau-1)]}_{\text{rezultat zbog šokova u } y_{2t}} \quad (2.66)$$

pa je udio u varijanci pogreške predviđanja u razdoblju τ zbog šokova u y_{1t} jednak:

$$\frac{\sigma_{u_1}^2 [\phi_{11}^2(0) + \phi_{11}^2(1) + \dots + \phi_{11}^2(\tau-1)]}{\sigma_{y_1}^2(\tau)}, \quad (2.67)$$

dok je udio u varijanci pogreške predviđanja u razdoblju τ zbog šokova u y_{2t} jednak:

$$\frac{\sigma_{u_2}^2 [\phi_{12}^2(0) + \phi_{12}^2(1) + \dots + \phi_{12}^2(\tau-1)]}{\sigma_{y_1}^2(\tau)}. \quad (2.68)$$

Upravo dekompozicija varijance govori koliki je udio varijance pogreške predviđanja neke varijable u VAR modelu uzrokovan šokom u toj istoj varijabli, a koliki je udio zbog šokova u

drugim varijablama sustava. Ovaj rezultat se može poopćiti za slučaj s N varijabli u VAR modelu. Pogreška predviđanja za svaku varijablu n u VAR modelu u razdoblju τ iznosi:

$$\varepsilon_{nt}(\tau) = y_{nt+\tau} - E(y_{nt+\tau}) = \sum_{i=1}^N (\phi_{ni}(0)u_{it+\tau} + \phi_{ni}(1)u_{it+\tau-1} + \dots + \phi_{ni}(\tau-1)u_{it+1}), \quad (2.69)$$

dok je varijanca te pogreške u razdoblju τ jednaka

$$\begin{aligned} Var(\varepsilon_{nt}(\tau)) &= \sigma_{yn}^2(\tau) = \sum_{i=1}^N (\phi_{ni}^2(0)\sigma_{ui}^2 + \phi_{ni}^2(1)\sigma_{ui}^2 + \dots + \phi_{ni}^2(\tau-1)\sigma_{ui}^2) \\ &= \sum_{i=1}^N \sigma_{ui}^2 \sum_{t=0}^{\tau-1} (\phi_{ni}^2(t)) \end{aligned} \quad (2.70)$$

Udio u varijanci pogreške predviđanja za varijablu y_{nt} u razdoblju τ s obzirom na šok u varijabli i jednak je

$$\varpi_{yn}(\tau) = \frac{\sigma_{un}^2 [\phi_{ni}^2(0) + \phi_{ni}^2(1) + \dots + \phi_{ni}^2(\tau-1)]}{\sigma_{yn}^2(\tau)}. \quad (2.71)$$

Napomena: kao i u slučaju procjene funkcija impulsnog odziva, i za izračun udjela u varijanci pogreške predviđanja bitan je poredak varijabli u modelu..

Primjer 2.11.

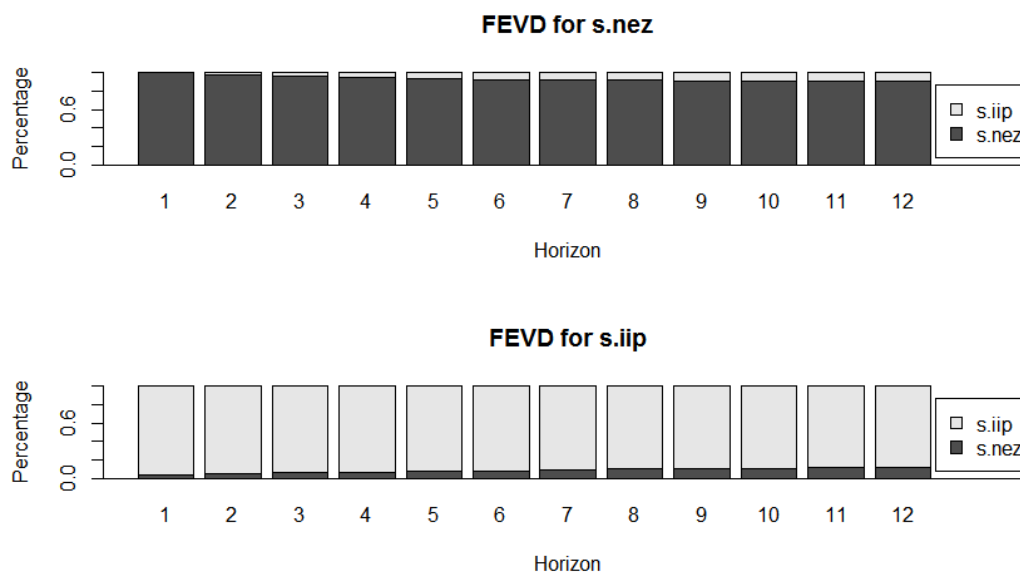
Za VAR(1) model procijenjen u primjeru 2.3. izvršimo dekompoziciju varijanci obiju varijabli u modelu za iduća 4 mjeseca i interpretirajmo rezultate. Poredak varijabli neka bude s.iip, s.nez, a potom obrnut.

## \$s.iip			## \$s.nez		
##	s.iip	s.nez	##	s.iip	s.nez
## [1,]	1.0000000	0.000000000	## [1,]	0.03922820	0.9607718
## [2,]	0.9975982	0.002401810	## [2,]	0.09755006	0.9024499
## [3,]	0.9930467	0.006953344	## [3,]	0.13931147	0.8606885
## [4,]	0.9873926	0.012607417	## [4,]	0.16751151	0.8324885

Slika 2.12. Dekompozicije varijanci za VAR model, $\tau = 4$, poredak varijabli: stopa rasta IIP-a, stopa rasta nezaposlenosti

## \$s.nez			## \$s.iip		
##	s.nez	s.iip	##	s.nez	s.iip
## [1,]	1.0000000	0.000000000	## [1,]	0.03922820	0.9607718
## [2,]	0.9792993	0.02070074	## [2,]	0.04961444	0.9503856
## [3,]	0.9588250	0.04117500	## [3,]	0.05908382	0.9409162
## [4,]	0.9435368	0.05646322	## [4,]	0.06762972	0.9323703

Slika 2.13. Dekompozicije varijanci za VAR model, $\tau = 4$, poredak varijabli: stopa rasta nezaposlenosti, stopa rasta IIP-a



Slika 2.14. Grafički prikaz dekompozicije varijanci u VAR modelu

Naredba fevd(...):

```
fevd(m1,n.ahead=4)
```

odnosi se na procjenu izraza (2.67) i (2.68) za slučaj dviju varijabli koje razmatramo u modelu. Kao prvi argument se navodi naziv VAR modela kojeg smo prethodno procijenili ($m1$), a potom kao drugi argument pišemo vrijednost τ ($n.ahead=4$). Na slici 2.12. predočene su procijenjene vrijednosti udjela u varijanci pogrešaka objiju varijabli za prvi poredak, dok je za obrnut poredak rezultat prikazan na slici 2.13. Interpretacija na slici 2.12. je sljedeća: varijanca pogreške predviđanja stope rasta IIP-a u prvome mjesecu je objašnjena u potpunosti šokovima u toj istoj varijabli (varijabla stopa rasta IIP-a u prvome mjesecu objašnjava 100% svoje varijacije). U drugome mjesecu šokovi u stopi rasta IIP-a pojašnjavaju ukupno 99,76% varijacije, dok šokovi u stopi rasta nezaposlenosti objašnjavaju 0,24% ukupne varijance varijable stopa rasta IIP-a. S druge strane, varijanca pogreške predviđanja stope rasta nezaposlenosti u prvome mjesecu objašnjena u iznosu 3,92% zbog šokova u stopi rasta IIP-a, dok je 96,08% objašnjeno šokovima u samoj stopi rasta nezaposlenosti, u drugom mjesecu se smanjuje udio zbog šokova u stopi rasta nezaposlenosti (90,24%), dok se povećava udio zbog šokova u stopi rasta IIP-a (9,76%). Konačno, ako se promijeni poredak varijabli, rezultati su nešto drugačiji (vidjeti sliku 2.13). Čitatelju ostaju interpretacije u tom slučaju za vježbu.

Dekompozicija varijanci se može i grafički predočiti na način da se najprije spremi naredba kojom se procjenjuju izrazi u (2.71) i potom se pomoću naredbe `plot(...)` nacrtaju udjeli kao što je predočeno na slici 2.14. za $\tau = 12$.

2.6. Kriteriji odabira adekvatnog pomaka p

Prilikom procjene $VAR(p)$ modela, jedno od važnih pitanja je odabir pomaka p . Želimo odabrati pomak p takav da model dobro opisuje ponašanje varijabli koje se razmatraju, no valja imati na umu da iako veći pomak p rezultira s točnijim procjenama unutar uzorka, to ide na štetu predviđanja izvan uzorka. U svrhu određivanja optimalnog pomaka p možemo provesti test omjera vjerodostojnosti, kao i usporediti informacijske kriterije.

2.6.1. Test omjera vjerodostojnosti

Test omjera vjerodostojnosti (engl. *likelihood ratio test*) temelji se na usporedbi determinanti procijenjenih matrica varijanci-kovarijanci reziduala VAR modela s brojem pomaka p_0 i VAR modela s brojem pomaka p_1 (vidjeti detalje u odjeljku 2.9. vezano uz procjenu VAR modela), pri čemu je $p_1 > p_0$. Ideja je slična kao i u odjeljku 2.1.9.5., gdje su se uspoređivale optimalne vrijednosti funkcije cilja u slučaju metode najveće vjerodostojnosti. U nultoj hipotezi se pretpostavlja da je optimalan broj pomaka p_0 , dok u alternativnoj p_1 . Test veličina za usporedbu dva VAR modela s brojem pomaka p_0 i p_1 računa se na sljedeći način:

$$LR = T (\ln |\hat{\Omega}_{p_0}| - \ln |\hat{\Omega}_{p_1}|), \quad (2.72)$$

gdje je $|\hat{\Omega}_{p_0}|$ determinanta matrice varijanci-kovarijanci reziduala modela s p_0 pomaka, a $|\hat{\Omega}_{p_1}|$ determinanta matrice varijanci-kovarijanci reziduala modela s p_1 pomaka. Kada se radi o uzorcima s malim brojem opservacija, Sims (1980) je predložio korekciju veličine u (2.72):

$$LR' = (T-c) (\ln |\hat{\Omega}_{p_0}| - \ln |\hat{\Omega}_{p_1}|), \quad (2.73)$$

gdje je c broj parametara koji se procjenjuju u svakoj jednadžbi u modelu koji ima veći broj parametara, $c = Np+1$. Test veličina u (2.73) slijedi hi-kvadrat distribuciju s brojem stupnjeva slobode koji su jednaki broju ograničenja u sustavu, $N^2(p_1-p_0)$. Ako je test veličina u (2.73) veća od kritične granice za zadanu razinu značajnosti i broj stupnjeva slobode $N^2(p_1-p_0)$, odbacuje se nulta hipoteza.

Primjer 2.12.

Temeljem podataka učitanih u primjeru 2.3. za VAR(1) i VAR(2) model i varijable stopa rasta IIP-a i stopa rasta nezaposlenosti, provedimo LR test kojim se testiraju sljedeće hipoteze o broju pomaka u VAR modelu:

$$H_0 : p_0 = 1$$

$$H_1 : p_1 = 2$$

odnosno, u nultoj hipotezi se pretpostavlja da je optimalan broj pomaka u VAR modelu jednak 1, dok u alternativnoj da je broj pomaka 2.

Temeljem naredbi na slici 2.15., izračunata je test veličina u (2.73) i pripadajuća p -vrijednost. lr je test veličina, a naredba `pchisq(...)` računa pripadajuću p -vrijednost, gdje je

$$(\text{length}(s.iip)-2*2+1)*(\log(\det(\text{summary}(m1)\$covres))-\log(\det(\text{summary}(m2)\$covres)))$$

naredba za izračun vrijednosti u (2.73), pri čemu se prva zagrada odnosi na izračun duljine vremenskog niza `s.iip` (= T u formuli (2.73)), te je broj parametara koji se procjenjuju c jednak $2*2+1$ (prema $Np + 1 = 2*2+1$), te je druga zagrada razlika logaritmiranih determinanti matrica varijanci-kovarijanci za model `m1` i `m2`. Naredba `summary(m1)$covres` se odnosi na matricu varijanci-kovarijanci iz prvog modela, `det(...)` je naredba za determinantu spomenute matrice, a slično vrijedi i za model `m2`.

Ispis rezultata temeljem slike 2.15. je sljedeći: $LR' = 105,0223$, dok je pripadajuća p -vrijednost približno 0, što znači da se uz razinu značajnosti od 5% odbacuje hipoteza da je optimalan broj pomaka 1.

```
lr<-(length(s.iip)-2*2+1)*(log(det(summary(m1)$covres))-log(det(summary(m2)$covres)))
lr
## [1] 105.0223
pchisq(lr,2^2*(2+1),lower.tail = F)
## [1] 5.745193e-17
```

Slika 2.15. LR test o optimalnom broj u pomaka u VAR (p) modelu

2.6.2. Informacijski kriteriji

Problem korištenja testa omjera vjerodostojnosti jest što se test veličine temelje na asimptotskoj teoriji. Stoga se kao alternativni kriteriji odabira pomaka p koriste informacijski kriteriji. **To su veličine koje istovremeno uzimaju u obzir optimalnu vrijednost funkcije cilja koja se optimizira prilikom procjene nepoznatih parametara, ali i broj parametara koje je potrebno procijeniti u modelu.** Veći broj parametara u VAR modelu kao rezultat povećanja broja pomaka p rezultira s boljom¹⁷ optimalnom vrijednošću funkcije cilja, no potrebno je i penalizirati prevelik broj parametara koji mogu biti nepotrebno uvršteni u model. Zato se informacijski kriteriji računaju na način da se veća optimalna vrijednost funkcije cilja „nagrađuje“, dok se povećanje broja parametara penalizira. Tri najpoznatija i najčešće korištena informacijska kriterija su AIC (Akaikeov), SIC (Schwartzov) te HQC (Hannan-Quinnov) informacijski kriterij (engl. *IC – information criterion*):

$$AIC = T \ln |\hat{\Omega}| + 2M, \quad (2.74)$$

$$SIC = T \ln |\hat{\Omega}| + M \ln(T), \quad (2.75)$$

$$HQC = T \ln |\hat{\Omega}| + 2M \ln(\ln(T)), \quad (2.76)$$

gdje je M broj parametara procijenjenih u VAR modelu. Dakle, uočimo da kao i u slučaju univarijatnih modela (vidjeti odjeljak 1.17.) manja vrijednost ovih kriterija, jer veći broj parametara M rezultira smanjenjem vrijednosti $\ln |\hat{\Omega}|$ („nagrađivanje“), no povećava se M u svakome kriteriju (penaliziranje).

Napomena: Potrebno je usporediti modele na ispravan način da se koristi jednak broj opservacija u svakome. Naime, kod modela s većim brojem pomaka p se smanjuje broj dostupnih opservacija, zbog dodatnih pomaka u modelu, pa je u modelu s manjem brojem pomaka potrebno koristiti jednak broj opservacija za izračun informacijskih kriterija i ispravnu usporedbu s modelima s većim brojem pomaka.

Napomena: Kao i u slučaju univarijatnih modela, Schwartzov informacijski kriterij je najstroži (zbog penala $\ln(T)$ za svaki dodatni parametar).

¹⁷ Bolja optimalna vrijednost funkcije cilja ovisi o metodi kojom se procjenjuje model, može se primijeniti multivarijanta metoda najmanjih kvadrata ili metoda najveće vjerodostojnosti.

Primjer 2.13.

Temeljem podataka učitanih u primjeru 2.3. za VAR model i varijable stopa rasta IIP-a i stopa rasta nezaposlenosti, odaberimo adekvatan broj pomaka u VAR modelu temeljem informacijskih kriterija u (2.74)-(2.76).

Naredba za izračun informacijskih kriterija je VARselect(...), gdje kao prvi argument potrebno navesti naziv matrice u kojoj se nalaze podaci, maksimalan broj pomaka p do kojeg se žele procijeniti VAR(p) modeli (argument lag.max=12) i usporediti prema tim kriterijima, te se navede tip modela (samo uključena konstanta, uključen i trend, bez konstante i trenda, argument type="const"):

VARselect(matrica,lag.max=12,type="const")\$selection.

U ovome primjeru je odabran maksimalan broj pomaka $p = 12$, te slika 2.16. predočava ispis optimalnog broja pomaka. AIC, HQ i SC redom označavaju Akaikeov, Hannan-Quinnov te Schwartzov informacijski kriterij, dok je FPE konačna pogreška predviđanja (vidjeti odjeljak 2.6.3.). Dva kriterija, Akaikeov i Hannan-Quinnov informacijski kriteriji ukazuju da je optimalan broj pomaka $p = 5$, dok Schwartzov upućuje na 3.

AIC(n)	HQ(n)	SC(n)	FPE(n)
5	5	3	5

Slika 2.16. Optimalan broj pomaka prema informacijskim kriterijima

Same vrijednosti svih kriterija moguće je dobiti naredbom VARselect(...) bez dijela \$selection, ako su potrebne u samoj analizi.

Konačna pogreška predviđanja

Dodatan kriterij koji se učestalo koristi uz informacijske kriterije je konačna pogreška predviđanja (engl. *final prediction error*, FPE), koju je predložio Akaike (1969). FPE mjera računa se na sljedeći način:

$$FPE = \left(\frac{T + Np + 1}{T - Np - 1} \right)^N \ln |\hat{\Omega}|. \quad (2.77)$$

Naravno, što je greška manja, model je bolji po tom kriteriju, kad se uspoređuje više modela istovremeno.

Primjer 2.14.

Temeljem podataka učitanih u primjeru 2.3. za VAR model i varijable stopa rasta IIP-a i stopa rasta nezaposlenosti, odaberimo adekvatan broj pomaka u VAR modelu temeljem FPE kriterija.

U ispisu na slici 2.16. uočava se da je optimalan broj pomaka za odabrane varijable prema FPE kriteriju jednak 5.

2.7. Multivarijantni testovi vezani uz inovacijske procese

U odjeljku 2.1.1. navedeno je da je vektor $\boldsymbol{\varepsilon}_t = [\varepsilon_{1t} \ \varepsilon_{2t} \ \dots \ \varepsilon_{Nt}]'$ vektor slučajnih procesa (multivarijantni čisti slučajni proces) u okviru VAR(p) modela. Navedene su njegove karakteristike:

$$E(\boldsymbol{\varepsilon}_t) = E \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{pmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0}, \quad E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega}, \quad E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s') = \mathbf{0}, \quad t \neq s \quad (2.78)$$

tj. očekivana vrijednost svakoga procesa ε_{jt} jednaka je 0, $j \in \{1, 2, \dots, N\}$, matrica varijanci-kovarijanci $\boldsymbol{\Omega}$ između svih procesa u istom vremenskom razdoblju t je pozitivno-semidefinitna matrica (vidjeti Dodatak u prvom dijelu), što znači da može postojati korelacija između varijabli ε_{jt} i ε_{kt} u istom razdoblju t , dok je matrica varijanci-kovarijanci između procesa ε_{jt} i ε_{js} nulmatrica: $\mathbf{0}$. To znači da za sve procese j i da za ni jedno razdoblje $s \neq t$ ne postoji korelacija procesa samog sa sobom: $\text{Cov}(\varepsilon_{jt}, \varepsilon_{js}) = 0$ i ne postoji korelacija procesa ε_{jt} s procesom ε_{kt} $\text{Cov}(\varepsilon_{jt}, \varepsilon_{ks}) = 0$.

Lütkepohl (2006) navodi da ispitivanje ovih karakteristika procesa $\boldsymbol{\varepsilon}_t$ nije od primarne važnosti ako se VAR koristi u svrhu prognoziranja (predviđanja). No, korisno je provesti testove nad procjenom procesa $\boldsymbol{\varepsilon}_t$, pogotovo kada se model konstruira temeljem ekonomske teorije, a ne vjerojatnosnih svojstava varijabli koje se razmatraju.

2.7.1. Multivarijantni testovi autokorelacije

Za multivarijantni vektorski proces \mathbf{y}_t može se provesti test o auto i kros-korelacijama u \mathbf{y}_t . U nultoj hipotezi se pretpostavlja $H_0: \rho_1 = \dots = \rho_M = 0$, dok u alternativnoj $H_1: \exists \rho_i \neq 0, i \in \{1, \dots, M\}$. Multivarijantno proširenje Ljung-Boxova testa svodi se na izračun sljedeće test veličine:

$$Q(M) = T \sum_{h=1}^M \text{tr}(\hat{\boldsymbol{\Gamma}}_h \hat{\boldsymbol{\Gamma}}^{-1} \hat{\boldsymbol{\Gamma}}_h \hat{\boldsymbol{\Gamma}}^{-1}) \sim \chi^2(N^2 M), \quad (2.79)$$

gdje je M maksimalni pomak za koji se testira nepostojanje autokorelacije, T je veličina uzorka, N je broj varijabli u vektoru \mathbf{y}_t , a matrice $\hat{\boldsymbol{\Gamma}}_h$ i $\hat{\boldsymbol{\Gamma}}$ su procjene matrica u (2.23) i (2.24) i $\text{tr}(\dots)$ je trag matrice. Ako je nulta hipoteza istinita, test veličina $Q(M)$ asimptotski slijedi hi-kvadrat distribuciju s $N^2 M$ stupnjeva slobode. Ako se ovaj test provodi nad vektorom reziduala $\hat{\boldsymbol{\varepsilon}}_t$ iz VAR modela, test veličina je

$$Q(M) = T \sum_{h=1}^M \text{tr}(\hat{\boldsymbol{\Omega}}_h \hat{\boldsymbol{\Omega}}^{-1} \hat{\boldsymbol{\Omega}}_h \hat{\boldsymbol{\Omega}}^{-1}) \sim \chi^2(N^2(M-p)), \quad (2.80)$$

gdje je $\hat{\boldsymbol{\Omega}} = \frac{\sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t'}{T}$, $\hat{\boldsymbol{\Omega}}_h = \frac{\sum_{t=1+h}^T \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_{t-h}'}{T}$, te ako je nulta hipoteza istinita, test veličina u (2.80) asimptotski slijedi hi-kvadrat distribuciju s $N^2(M-p)$ stupnjeva slobode. Problem s ovim testom

nastaje za male uzorke¹⁸ (mala vrijednost T), pa se test veličine u (2.79) i (2.80) korigiraju u sljedeće:

$$Q^*(M) = T^2 \sum_{h=1}^M \frac{1}{T-h} \text{tr}(\hat{\Gamma}_h \hat{\Gamma}_h^{-1} \hat{\Gamma}_h \hat{\Gamma}_h^{-1}), \quad (2.81)$$

i

$$Q^*(M) = T^2 \sum_{h=1}^M \frac{1}{T-h} \text{tr}(\hat{\Omega}_h \hat{\Omega}_h^{-1} \hat{\Omega}_h \hat{\Omega}_h^{-1}). \quad (2.82)$$

Osim ovoga, provodi se i LM (engl. *Lagrange multiplier*) test, koji se temelji na pomoćnim regresijskim jednadžbama:

$$\hat{\boldsymbol{\varepsilon}}_t = \sum_{i=1}^p \mathbf{A}_i \mathbf{y}_{t-i} + \sum_{i=1}^M \mathbf{B}_i \hat{\boldsymbol{\varepsilon}}_{t-i} + \mathbf{u}_t, \quad (2.83)$$

gdje se u nultoj hipotezi pretpostavlja da su sve \mathbf{B}_i matrice nul-matrice:

$$H_0 : \mathbf{B}_1 = \mathbf{B}_2 = \dots = \mathbf{B}_M = \mathbf{0}, \quad (2.84)$$

dok alternativna pretpostavlja da je neka od tih matrica nenul-matrice:

$$H_1 : \exists \mathbf{B}_i \neq \mathbf{0}, i \in \{1, 2, \dots, M\}. \quad (2.85)$$

Test veličina jednaka je

$$LM = T \left(N - \text{tr}(\hat{\Omega}_R^{-1} \hat{\Omega}_U) \right) \sim \chi^2(MN^2), \quad (2.86)$$

gdje su $\hat{\Omega}_R$ i $\hat{\Omega}_U$ matrice varijanci-kovarijanci reziduala iz modela s ograničenjem (R – engl. *restricted*) i bez ograničenja (U – engl. *unrestricted*). Ponovno, postoji korekcija veličine u (2.96) za slučaj malog uzorka (Edgerton i Shukur, 1999):

$$LM^* = \frac{1 - (1 - R_r^2)^{1/r}}{(1 - R_r^2)^{1/r}} \frac{Kr - q}{Nh} \sim F(MN^2; \text{int}(Kr - q)), \quad (2.87)$$

gdje je $r = ((N^2 h^2 - 4) / (N^2 + h^2 - 5))^{0.5}$, $q = 1/2Nh - 1$, $K = T - N - h - 1/2(N - h + 1)$, $h = NM$, $R_r^2 = 1 - |\hat{\Omega}_U| / |\hat{\Omega}_R|$. $\text{int}(\dots)$ je naredba koja pretvara vrijednost u zagradi u cijeli broj (najveći cijeli broj koji je manji ili jednak zadanom broju). Ako je nulta hipoteza istinita, test veličina u (2.87) slijedi F distribuciju s MN^2 stupnjeva slobode u brojniku i $\text{int}(Kr - q)$ u nazivniku. Izvode test veličina u ovome odjeljku moguće je vidjeti u Lütkepohl (2006).

Napomena: sljedeće naredbe je potrebno upisati u RStudiju kako bi se ispravno proveo neki od prethodno navedenih testova, u okviru naredbe `serial.test(...)`:

¹⁸ Vidjeti Davies, Triggs i Newbold (1977), Ljung i Box (1978), Hosking (1980).

Test	Formula	Tip testa	Broj pomaka
Multivarijantno proširenje Ljung-Boxova testa	(2.80)	PT.asymptotic	lags.pt = ...
Korigirana vrijednost multivarijantnog proširenja Ljung-Boxova testa	(2.82)	PT.adjusted	lags.pt = ...
LM test	(2.86)	BG	lags.bg = ...
Korigirani LM test	(2.87)	ES	lags.bg = ...

Primjer 2.14.

Za VAR(1) model koji je procijenjen u primjeru 2.3. provedimo multivarijantne testove autokorelacije reziduala pomoću 4 prethodno obrađena testa. Odaberimo pomak $M = 1$ za svaki test.

Za spremljeni model (naziv m1) je potrebno koristiti sljedeće naredbe kako bi se testovi mogli provesti (slika 2.16.): `serial.test(...)` je naredba za provođenje multivarijantnog testa autokorelacije, pri čemu se u naredbi na prvo mjesto navodi naziv VAR modela za čije rezidualne provodimo sam test, potom se zapisuje broj pomaka (`lags.pt` ili `lags.bg = broj pomaka`), i pod tip testa (`type`) navodi se naziv samog testa koji se provodi.

Uočava se da su sve test veličine relativno velike vrijednosti pa su pripadajuće p -vrijednosti veoma male (p -value u ispisima), manje od razine značajnosti i 1%, stoga se zaključuje da se pri odabranoj razini značajnosti odbacuju nulte hipoteze svakoga testa. U ovome primjeru postoji problem multivarijantne autokorelacije.

```
serial.test(m1, lags.pt=1, type="PT.asymptotic")
serial.test(m1, lags.pt=1, type="PT.adjusted")
serial.test(m1, lags.bg=1, type="BG")
serial.test(m1, lags.bg=1, type="ES")
```

Slika 2.16. Naredbe za provođenje multivarijantnih testova autokorelacije

```
serial.test(m1, lags.pt=1, type="PT.asymptotic")
##
## Portmanteau Test (asymptotic)
##
## data: Residuals of VAR object m1
## Chi-squared = 69.771, df = 0, p-value < 2.2e-16
serial.test(m1, lags.pt=1, type="PT.adjusted")
##
## Portmanteau Test (adjusted)
##
## data: Residuals of VAR object m1
## Chi-squared = 70.068, df = 0, p-value < 2.2e-16

serial.test(m1, lags.bg=1, type="BG")
##
## Breusch-Godfrey LM test
##
## data: Residuals of VAR object m1
## Chi-squared = 87.651, df = 4, p-value < 2.2e-16

serial.test(m1, lags.bg=1, type="ES")
##
## Edgerton-Shukur F test
```

```
##
## data: Residuals of VAR object m1
## F statistic = 27.45, df1 = 4, df2 = 460, p-value < 2.2e-16
```

Slika 2.17. Rezultati testova temeljem naredbi na slici 2.16.

2.7.2. Multivarijatni test heteroskedastičnosti

Za testiranje multivarijatne heteroskedastičnosti, ponovno se mogu koristiti LM testovi. Ideja je procijeniti pomoćni model:

$$vech(\hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t') = \boldsymbol{\beta}_0 + \sum_{i=1}^q \boldsymbol{\beta}_i vech(\hat{\boldsymbol{\varepsilon}}_{t-i} \hat{\boldsymbol{\varepsilon}}_{t-i}') + \mathbf{n}_t, \quad (2.88)$$

gdje operator $vech(\dots)$ predstavlja operator koji konstruira vektor na način da komponente tog vektora čine stupci donjeg trokuta neke kvadratne matrice sukcesivnim slaganjem, $\boldsymbol{\beta}_0$ je $\frac{1}{2}N(N+1)$ -dimenzionalni vektor, a $\boldsymbol{\beta}_i$ su kvadratne matrice formata $\frac{1}{2}N(N+1)$. U nultoj hipotezi testa se pretpostavlja:

$$H_0: \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \dots = \boldsymbol{\beta}_q = \mathbf{0}, \quad (2.89)$$

dok alternativna pretpostavlja da je neka od tih matrica nenul-matrica:

$$H_2: \exists \boldsymbol{\beta}_i \neq \mathbf{0}, i \in \{1, 2, \dots, q\}. \quad (2.90)$$

Odgovarajuća test veličina računa se izrazom:

$$LM_q = \frac{1}{2}TN(N+1) - Ttr(\hat{\boldsymbol{\Omega}}_{vech} \hat{\boldsymbol{\Omega}}_0^{-1}) \sim \chi^2(qN^2(N+1)^2/4), \quad (2.91)$$

gdje su $\hat{\boldsymbol{\Omega}}_{vech}$ i $\hat{\boldsymbol{\Omega}}_0$ matrica varijanci-kovarijanci reziduala iz modela (2.88) i matrica varijanci-kovarijanci reziduala iz tog istog modela kada se uvrsti da je $q = 0$. Pod pretpostavkom da je nulta hipoteza istinita, test veličina u (2.91) asimptotski slijedi hi-kvadrat distribuciju s $qN^2(N+1)^2/4$ stupnjeva slobode.

Primjer 2.15.

Za VAR(1) model koji je procijenjen u primjeru 2.3. provedimo multivarijatni test heteroskedastičnosti reziduala za pomak $q = 1$.

U RStudiju je u okviru naredbe `arch.test(...)` moguće istovremeno provesti i multivarijatni test heteroskedastičnosti, kao i pojedinačne (univarijatne testove):

```
arch.test(m1,lags.single=1,lags.multi=1,multivariate.only=F),
```

gdje je prvi argument naziv modela za koji se test provodi, `lags.single=` je argument za koji navodimo za koliko pomaka u pojedinačnim jednadžbama treba provesti test, `lags.multi=` je argument koji se odnosi na pomake u multivarijatnom slučaju testiranja, te se posljednji argument, `multivariate.only=` odnosi na izjavu T (engl. *true* – točno) ili F (engl. *false* – netočno), s kojom se definira hoće li se test provesti samo za multivarijatni slučaj (T) ili i za pojedinačne slučajeve (F).

Naredba na slici 2.18. predočava provođenje testa u ovome primjeru, za pomak $q = 1$, te istovremenu provedbu univarijantnih i multivarijantnog testa, kao i sam ispis rezultata.

```
arch.test(m1,lags.single=1,lags.multi=1,multivariate.only=F)

## $s.iip
##
## ARCH test (univariate)
##
## data: Residual of s.iip equation
## Chi-squared = 0.29584, df = 1, p-value = 0.5865
##
##
## $s.nez
##
## ARCH test (univariate)
##
## data: Residual of s.nez equation
## Chi-squared = 40.322, df = 1, p-value = 2.154e-10
##
##
## ARCH (multivariate)
##
## data: Residuals of VAR object m1
## Chi-squared = 54.974, df = 9, p-value = 1.232e-08
```

Slika 2.18. Naredbe za provođenje multivarijantnog testa heteroskedastičnosti, zajedno s rezultatima

U naredbi `arch.test(...)` se najprije navodi naziv VAR modela za koji se test provodi, potom se određuju pomaci do kojih se provode testovi (`lags.single` ili `lags.multi`), i zadnji dio naredbe određuje želimo li provesti samo multivarijantni test (`multivariate.only=T`) ili uključiti i univarijantne (`multivariate.only=F`). Uočava se da za slučaj reziduala iz jednadžbe koja se odnosi na stopu rasta IIP-a (u ispisu `$s.iip`) ne možemo odbaciti hipotezu o homoskedastičnosti reziduala pri uobičajenim razinama značajnosti. Nadalje, za drugu varijablu u VAR modelu (stopu rasta nezaposlenosti, u ispisu `$s.nez`) se nulta hipoteza odbacuje i pri razini značajnosti od 1%, kao i za slučaj multivarijantnog testa koji je posljednji u ispisu.

2.7.3. Multivarijantni test normalnosti

Kako bi se mogle raditi intervalne procjene¹⁹ u okviru VAR modela, posebice vezano uz funkcije impulsnog odziva, potrebno je testirati i normalnost vektora reziduala iz VAR modela. U odjeljku 4.4.2. u prvome dijelu ovih publikacija obrađivan je univarijantni Jarque-Beraov test koji se temeljio na koeficijentu asimetrije i zaobljenosti normalne distribucije. Poopćenje na multivarijantni slučaj vektorskog procesa ε_t je sljedeće. Ako je normalno distribuiran proces, tada je $\varepsilon_t \sim N(\mu, \Omega)$, gdje su μ i Ω vektor očekivanih vrijednosti i matrica varijanci-kovarijanci. Temeljem procijenjenog VAR modela, procijeni se vektor reziduala $\hat{\varepsilon}_t$, koji se potom centriraju, te se nad matricom varijanci-kovarijanci provodi dekompozicije Choleskyog za centrirane rezidualne (izvod spomenutih postupaka vidjeti u Lütkepohl (2006)).

¹⁹ Vidjeti Hamilton (1994), Lütkepohl (1991), ili Canova (2007). U posljednje spomenutoj referenci se prikazuju i *bootstrap* metode koje se temelje na pretpostavci normalnosti reziduala.

Testiranje je moguće provesti na način da se razmatraju samo koeficijenti asimetrije pojedinog procesa u $\hat{\varepsilon}_t$ i uspoređuju s nul-vektorom (vektor koeficijenata asimetrije), samo koeficijenti zaobljenosti pojedinog procesa u $\hat{\varepsilon}_t$ i uspoređuju s vektorom 3_N (N -komponentni vektor čije su sve komponente jednake 3), ili da se istovremeno uspoređuju koeficijenti asimetrije i koeficijenti zaobljenosti s onima koji vrijede za normalno distribuiran vektorski proces. Dodatno, moguće je provesti univarijatne testove, kao i sam multivarijatni test normalnosti. Bez obzira na tip testa koji se provode, sve test veličine asimptotski slijede hi-kvadrat distribuciju pod pretpostavkom istinite nulte hipoteze. No, broj stupnjeva slobode varira. Ako se provode univarijatni testovi za asimetriju ili zaobljenost, u tom slučaju se radi o 1 stupnju slobode za pojedinu varijablu, ako se provode multivarijatni testovi za asimetriju ili zaobljenost, u tom slučaju se radi o N stupnjeva slobode. Za istovremeno razmatranje asimetrije i zaobljenosti u univarijatnim testovima, broj stupnjeva slobode je 2 po pojedinoj varijabli, dok za multivarijatni slučaj test veličina ima $2N$ stupnjeva slobode.

Napomena: za provođenje testova normalnosti je također bitan poredak varijabli u VAR modelu o kojem je već bilo riječi, jer se provodi dekompozicija Choleskyog nad matricom varijanci-kovarijanci rezidua. U okviru provođenja multivarijatnih testova koji se ovdje obrađuju, korisno je razmotriti i univarijatne inačice prilikom samog provođenja testa, jer se na taj način može lakše uočiti u kojoj jednadžbi VAR modela dolazi do problema koji utječe na multivarijatnu inačicu testa.

Primjer 2.16.

Za VAR(1) model koji je procijenjen u primjeru 2.3. provedimo multivarijatni test normalnosti reziduala.

Naredba za provođenje multivarijatnog testa normalnosti prikazana je na slici 2.19., kao i sam ispis rezultata:

```
normality.test(m1,multivariate.only=T),
```

gdje je prvi argument naziv modela za koji se test provodi, a drugi argument je pitanje hoće i se provesti samo multivarijatno testiranje ili i univarijatno (multivariate.only=), tj. T ili F.

Ponovno se u okviru naredbe na prvo mjesto navodi naziv VAR modela koji smo procijenili, a potom se određuje želimo li provesti samo multivarijatni slučaj testa ili ne (multivariate.only=T ili F). Uočimo da u slučaju univarijatnih testova (lijevi panel) ne odbacujemo hipotezu o normalnosti distribucije reziduala iz jednadžbe gdje je zavisna varijabla stopa rasta IIP-a, dok se nulta hipoteza odbacuje u slučaju zavisne varijable stopa rasta nezaposlenosti (p -vrijednost je manja od uobičajenih razina značajnosti). Na desnome panelu provedeno je multivarijatno testiranje normalnosti. Prvi dio ispisa (JB) provodi istovremenu usporedbu koeficijenata asimetrije i zaobljenosti s onima koji karakteriziraju multivarijatnu normalnu distribuciju. Test veličina iznosi 6122,7, s veoma malom pripadajućom p -vrijednosti, što znači da odbacujemo nultu hipotezu o multivarijatnoj normalnosti reziduala u VAR modelu. Ako se još posebno testiraju multivarijatna asimetričnost (Skewness u ispisu) i zaobljenost (Kurtosis) vektora reziduala, uočava se da se i u tim testovima odbacuju nulte hipoteze.

```
normality.test(m1,multivariate.only=T)
```

```
## $JB
##
## JB-Test (multivariate)
##
```

```
## data: Residuals of VAR object m1
## Chi-squared = 6122.7, df = 4, p-value < 2.2e-16
##
## $Skewness
##
## Skewness only (multivariate)
##
## data: Residuals of VAR object m1
## Chi-squared = 396.39, df = 2, p-value < 2.2e-16
##
## $Kurtosis
##
## Kurtosis only (multivariate)
##
## data: Residuals of VAR object m1
## Chi-squared = 5726.3, df = 2, p-value < 2.2e-16
```

Slika 2.19. Naredbe za provođenje multivarijantnog testa normalnosti reziduala

2.8. Primjer

U RStudio učitajte datoteku „var.txt“ koja sadrži mjesečne podatke o 4 varijable: x , y , z i w , s početnim datumom siječanj 2000. Definirajte vremenske nizove x i y .

- a) Pomoću ADF testa ispitajte stacionarnost oba niza, pri čemu ćemo u DF jednadžbu uključiti konstantu kao deterministički regresor.

```
library(urca)
summary(ur.df(x,type="drift",selectlags="BIC"))

##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.51008 -0.14848 -0.00402  0.15670  0.46363
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.52018    0.06153   8.454 3.00e-15 ***
## z.lag.1      -0.62887    0.07277  -8.642 8.65e-16 ***
## z.diff.lag   0.01040    0.06529   0.159  0.874
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1974 on 235 degrees of freedom
## Multiple R-squared:  0.3107, Adjusted R-squared:  0.3048
## F-statistic: 52.96 on 2 and 235 DF, p-value: < 2.2e-16
##
## Value of test-statistic is: -8.6419 37.3416
##
## Critical values for test statistics:
##      1pct 5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

summary(ur.df(y,type="drift",selectlags="BIC"))

##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.137550 -0.055996  0.008972  0.054913  0.136819
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.59153    0.05474  10.807 <2e-16 ***
## z.lag.1      -0.94017    0.08661 -10.856 <2e-16 ***
## z.diff.lag   0.03442    0.06195   0.556  0.579
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07151 on 235 degrees of freedom
## Multiple R-squared:  0.4544, Adjusted R-squared:  0.4497
## F-statistic: 97.84 on 2 and 235 DF,  p-value: < 2.2e-16
##
## Value of test-statistic is: -10.8558 58.9316
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

Slika 2.20. Testiranje (ne)stacionarnosti nizova x i y

Temeljem slike 2.20. zaključuje se kako se na razini značajnosti od 5% odbacuje hipoteza da niz x nije stacionaran, i da niz y nije stacionaran. Dalje, u analizi koristimo varijable u razinama.

- b) Procijenite VAR(1) model između varijabli x i y , pri čemu neka poredak bude x i potom y . Zapišimo u matričnom obliku procijenjeni model, potom komentirajmo stabilnost modela i značajnost varijabli u pojedinoj jednadžbi.

Na slici 2.21. predočene su naredbe za procjenu i ispis procijenjenog VAR(1) modela. Matrični zapis modela je sljedeći:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 0,56 \\ 0,58 \end{bmatrix} + \begin{bmatrix} 0,38 & -0,07 \\ 0,01 & 0,06 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \end{bmatrix}$$

U prvoj jednadžbi je značajna konstanta i pomak varijable x , dok je u drugoj jednadžbi značajna samo konstanta. Svojsvene vrijednosti matrice A_1 iznose redom: 0,38 i 0,06. Kako su obje po apsolutnoj vrijednosti manje od 1, model je stabilan. Matrica varijanci-kovarijanci reziduala, zajedno s korelacijskom matricom reziduala je sljedeća:

$$\hat{\Omega} = \begin{bmatrix} 0,039 & 0,001 \\ 0,001 & 0,005 \end{bmatrix}, \widehat{Corr} = \begin{bmatrix} 1 & 0,062 \\ 0,062 & 1 \end{bmatrix}.$$

```
library(vars)
matrica<-cbind(x,y)
var1<-VAR(matrica,1,type="const")
summary(var1)

##
## VAR Estimation Results:
## =====
## Endogenous variables: x, y
## Deterministic variables: const
## Sample size: 239
## Log Likelihood: 343.224
```

```

## Roots of the characteristic polynomial:
## 0.3799 0.06187
## Call:
## VAR(y = matrica, p = 1, type = "const")
##
##
## Estimation results for equation x:
## =====
## x = x.l1 + y.l1 + const
##
##           Estimate Std. Error t value Pr(>|t|)
## x.l1    0.38232    0.06057   6.312 1.35e-09 ***
## y.l1   -0.07248    0.17050  -0.425   0.671
## const   0.55726    0.11522   4.837 2.38e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.1971 on 236 degrees of freedom
## Multiple R-Squared: 0.1445, Adjusted R-squared: 0.1373
## F-statistic: 19.93 on 2 and 236 DF, p-value: 1.004e-08
##
##
## Estimation results for equation y:
## =====
## y = x.l1 + y.l1 + const
##
##           Estimate Std. Error t value Pr(>|t|)
## x.l1    0.01075    0.02202   0.488   0.626
## y.l1    0.05944    0.06199   0.959   0.339
## const   0.58244    0.04189  13.904 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.07168 on 236 degrees of freedom
## Multiple R-Squared: 0.005275, Adjusted R-squared: -0.003155
## F-statistic: 0.6258 on 2 and 236 DF, p-value: 0.5357
##
##
## Covariance matrix of residuals:
##           x           y
## x 0.0388660 0.0008745
## y 0.0008745 0.0051375
##
## Correlation matrix of residuals:
##           x           y
## x 1.00000 0.06189
## y 0.06189 1.00000

```

Slika 2.21. Procjena VAR(1) modela

- c) Procijenite VAR(2) model između varijabli x i y , pri čemu neka poredak bude x i potom y . Zapišite u matricnom obliku procijenjeni model, potom komentirajmo stabilnost modela i značajnost varijabli u pojedinoj jednadžbi.

Na slici 2.22. predočene su naredbe za procjenu i ispis procijenjenog VAR(2) modela. Matricni zapis modela je sljedeći:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 0,46 \\ 0,57 \end{bmatrix} + \begin{bmatrix} 0,38 & -0,15 \\ 0,002 & 0,09 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} -0,02 & 0,27 \\ 0,03 & -0,04 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \end{bmatrix}$$

U prvoj jednadžbi je značajna konstanta i prvi pomak varijable x , dok je u drugoj jednadžbi značajna samo konstanta.

```
var2<-VAR(matrica,2,type="const")
summary(var2)

##
## VAR Estimation Results:
## =====
## Endogenous variables: x, y
## Deterministic variables: const
## Sample size: 238
## Log Likelihood: 344.486
## Roots of the characteristic polynomial:
## 0.4166 0.3172 0.3172 0.1469
## Call:
## VAR(y = matrica, p = 2, type = "const")
##
##
## Estimation results for equation x:
## =====
## x = x.l1 + y.l1 + x.l2 + y.l2 + const
##
##           Estimate Std. Error t value Pr(>|t|)
## x.l1    0.38863    0.06545   5.938 1.04e-08 ***
## y.l1   -0.14594    0.18003  -0.811  0.4184
## x.l2   -0.01965    0.06546  -0.300  0.7643
## y.l2    0.26753    0.17144   1.560  0.1200
## const  0.44515    0.15814   2.815  0.0053 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.197 on 233 degrees of freedom
## Multiple R-Squared: 0.153, Adjusted R-squared: 0.1384
## F-statistic: 10.52 on 4 and 233 DF, p-value: 7.503e-08
##
##
## Estimation results for equation y:
## =====
## y = x.l1 + y.l1 + x.l2 + y.l2 + const
##
##           Estimate Std. Error t value Pr(>|t|)
## x.l1    0.001635    0.023786   0.069  0.945
## y.l1    0.091397    0.065425   1.397  0.164
## x.l2    0.026013    0.023790   1.093  0.275
## y.l2   -0.040774    0.062304  -0.654  0.513
## const  0.574458    0.057469   9.996 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.07159 on 233 degrees of freedom
## Multiple R-Squared: 0.01589, Adjusted R-squared: -0.001008
## F-statistic: 0.9403 on 4 and 233 DF, p-value: 0.4413
##
##
##
## Covariance matrix of residuals:
##           x           y
## x 0.038808 0.001039
## y 0.001039 0.005125
##
##
## Correlation matrix of residuals:
##           x           y
```

```
## x 1.00000 0.07369
## y 0.07369 1.00000
```

Slika 2.22. Procjena VAR(2) modela

Svojevrsne vrijednosti matrice A_1 iznose redom: 0,42 i 0,32, dok za matricu A_2 iznose: 0,32 i 0,15. Kako su sve svojevrsne vrijednosti po apsolutnoj vrijednosti manje od 1, i ovaj model je stabilan. Matrica varijanci-kovarijanci reziduala, zajedno s korelacijskom matricom reziduala je sljedeća:

$$\hat{\Omega} = \begin{bmatrix} 0,039 & 0,001 \\ 0,001 & 0,005 \end{bmatrix}, \quad \widehat{Corr} = \begin{bmatrix} 1 & 0,074 \\ 0,074 & 1 \end{bmatrix}$$

- d) Provjerimo koji vremenski pomak p u VAR modelu je prikladan s obzirom na pogrešku FPE te informacijske kriterije AIC, SIC, HQC.

Na slici 2.23. uočava se da su svi informacijski kriteriji, kao i pogreška FPE ukazuju da je optimalan broj pomaka jednak 1. Stoga se ostatak analize vrši nad VAR(1) modelom.

```
VARselect(matrica, lag.max = 12, type="const")$selection
```

```
## AIC(n) HQ(n) SC(n) FPE(n)
##      1      1      1      1
```

Slika 2.23. Adekvatan broj pomaka u VAR modelu

- e) Za odabrani pomak u postupku e) provedite Grangerovo testiranje uzročnosti i donesimo zaključak o ishodima testova uz razinu značajnosti 5%.

Za nultu hipotezu „varijabla x ne uzrokuje varijablu y “ u Grangerovom smislu u VAR(1) modelu, testira se značajnost varijable x_{t-1} u jednadžbi gdje je zavisna varijabla y . Test veličina iznosi 0,24, s pripadajućom p -vrijednošću 0,63, što je veće od razine značajnosti 0,05 (=5%). Pri razini značajnosti od 5%, ne odbacuje se hipoteza da varijabla x ne uzrokuje varijablu y u Grangerovom smislu. Sličan je zaključak i za varijablu y : za nultu hipotezu „varijabla y ne uzrokuje varijablu x “ u VAR(1) modelu, test veličina iznosi 0.18, uz p -vrijednost 0,67, što znači da se također ne odbacuje nulta hipoteza i ovoga testa (slika 2.24).

```
causality(var1, cause="x")$Granger
```

```
## Granger causality H0: x do not Granger-cause y
##
## data: VAR object var1
## F-Test = 0.23814, df1 = 1, df2 = 472, p-value = 0.6258
```

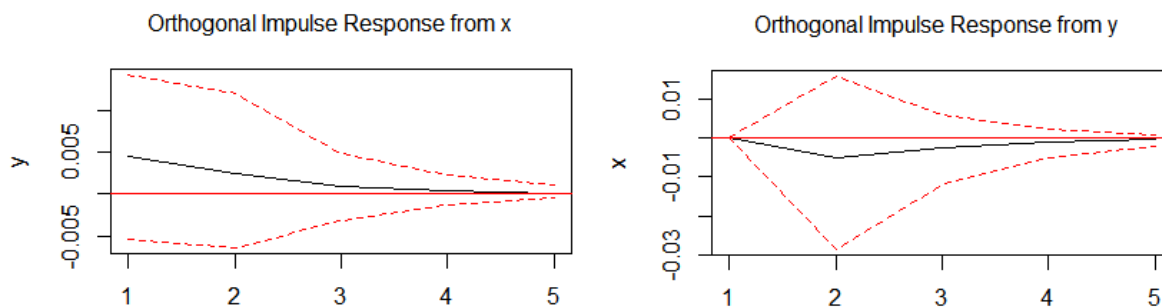
```
causality(var1, cause="y")$Granger
```

```
## Granger causality H0: y do not Granger-cause x
##
## data: VAR object var1
## F-Test = 0.18069, df1 = 1, df2 = 472, p-value = 0.671
```

Slika 2.24. Multivarijatni Grangerov test uzročnosti u VAR(1) modelu

- f) S obzirom na procijenjeni VAR(1) model procijenimo funkcije impulsnog odziva pomoću dekompozicije Choleskyog i interpretirajmo rezultate za $\tau = 4$.

```
plot(irf(var1,n.ahead = 4,ortho=T,ci=.95,impulse="x",response="y"))
plot(irf(var1,n.ahead = 4,ortho=T,ci=.95,impulse="y",response="x"))
```



Slika 2.25. Funkcije impulsnog odziva u VAR(1) modelu

Temeljem slike 2.25. zaključuje se da šok od jedne standardne devijacije u varijabli x ne rezultira sa značajnom reakcijom varijabli y (iako je vrijednost koeficijenata IRF funkcije pozitivna, nalaze se unutar granica neznačajnosti), kao i što šok od jedne standardne devijacije u varijabli y ne rezultira sa značajnom reakcijom varijable x .

- g) S obzirom na procijenjen VAR(1) model provedimo dekompoziciju varijanci obiju varijabli i interpretirajmo rezultate za $\tau = 4$. Poredak varijabli neka ostane kao kod procjene samog modela (x i potom y).

Slika 2.26. predočava rezultat naredbe `fevd(var1,n.ahead=4)`, čija je interpretacija sljedeća. U prvome mjesecu je 100% varijance varijable x objašnjeno šokovima u toj istoj varijabli, dok u drugom mjesecu je 99,94% varijance od x objašnjeno šokovima u istoj varijabli i 0,06% šokovima u varijabli y (nastavak interpretacija za preostale mjesece je sličan). Uočava se da šokovi u varijabli x najvećim dijelom pojašnjavaju varijacije te iste varijable. Slično vrijedi i za varijablu y . U prvome mjesecu je 99,62% varijance varijable y objašnjeno šokovima u toj istoj varijabli, a preostalih 0,38% je pojašnjeno šokovima u varijabli x . Zbog neznačajnosti varijabli u VAR(1) modelu, ovakav ishod za dekompoziciju varijanci i koeficijenata funkcije impulsnih odziva je očekivan.

## \$x			## \$y		
##	x	y	##	x	y
## [1,]	1.0000000	0.0000000000	## [1,]	0.003829971	0.9961700
## [2,]	0.9993962	0.0006037627	## [2,]	0.004911857	0.9950881
## [3,]	0.9992914	0.0007085612	## [3,]	0.005085095	0.9949149
## [4,]	0.9992759	0.0007241413	## [4,]	0.005110467	0.9948895

Slika 2.26. Dekompozicije varijanci varijabli x i y u VAR(1) modelu

- h) Provedimo multivarijatne testove autokorelacije za procijenjeni VAR(1) model, pri čemu ćemo provesti sva 4 testa obrađena u odjeljku 2.2.1. Testira se multivarijatna autokorelacija do četvrtog reda.

Temeljem ispisa na slici 2.27. uočava se da se nulta hipoteza svakog testa ne može odbaciti (uz razinu značajnosti od 5%, ne odbacuju se hipoteze svakog testa da ne postoji multivarijatna autokorelacija reziduala VAR(1) modela do zaključno četvrtog reda).


```

serial.test(var1, lags.pt=4, type="PT.asymptotic")
##
## Portmanteau Test (asymptotic)
##
## data: Residuals of VAR object var1
## Chi-squared = 15.964, df = 12,
p-value = 0.1929
serial.test(var1, lags.pt=4, type="PT.adjusted")
##
## Portmanteau Test (adjusted)
##
## data: Residuals of VAR object var1
## Chi-squared = 16.152, df = 12, p-value = 0.1844

serial.test(var1, lags.bg=4, type="BG")
##
## Breusch-Godfrey LM test
##
## data: Residuals of VAR object var1
## Chi-squared = 21.561, df = 16,
p-value = 0.1579
serial.test(var1, lags.bg=4, type="ES")
##
## Edgerton-Shukur F test
##
## data: Residuals of VAR object var1
## F statistic = 1.3421, df1 = 16, df2 = 454,
p-value = 0.1672

```

Slika 2.27. Multivarijantni testovi autokorelacije reziduala u VAR(1) modelu

- i) Provedimo multivarijantni test heteroskedastičnosti reziduala za pomak $q = 1$ u VAR(1) modelu.

Proveden je samo multivarijantni test (slika 2.28.), gdje se uočava da je test veličina jednaka 3.79, s pripadajućom p -vrijednošću 0.92, što je veće od uobičajenih razina značajnosti. Dakle, pri uobičajenim razinama značajnosti, ne odbacuje se hipoteza o multivarijantnoj homoskedastičnosti reziduala u VAR(1) modelu.

```

arch.test(var1, lags.single=1, lags.multi=1, multivariate.only=F)

## $x
##
## ARCH test (univariate)
##
## data: Residual of x equation
## Chi-squared = 0.39548, df = 1, p-value = 0.5294
##
##
## $y
##
## ARCH test (univariate)
##
## data: Residual of y equation
## Chi-squared = 1.1683, df = 1, p-value = 0.2798
##
##
## ARCH (multivariate)
##
## data: Residuals of VAR object var1
## Chi-squared = 3.7906, df = 9, p-value = 0.9246

```

Slika 2.28. Multivarijantni test heteroskedastičnosti reziduala u VAR(1) modelu

- j) Provedimo multivarijantni test normalnosti reziduala u VAR(1) modelu.

Za testiranje nulte hipoteze o multivarijantnoj normalnoj distribuciji reziduala VAR(1) modela, test veličina iznosi 14,58, s pripadajućom p -vrijednosti 0,006 (slika 2.29.). Uz razinu značajnosti od 5%, odbacuje se hipoteza da reziduali VAR(1) modela slijede multivarijantnu normalnu distribuciju.

```

normality.test(var1, multivariate.only = T)

## $JB
##

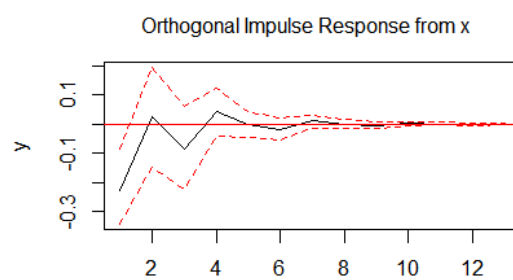
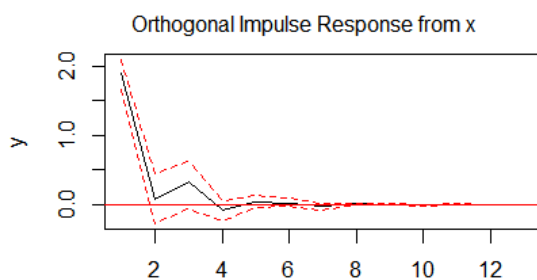
```

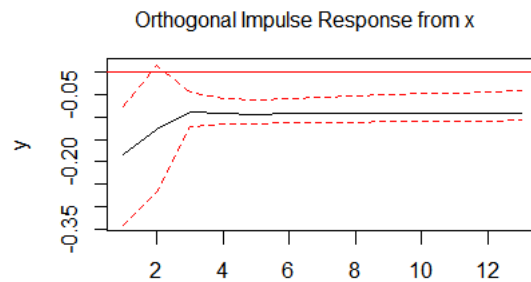
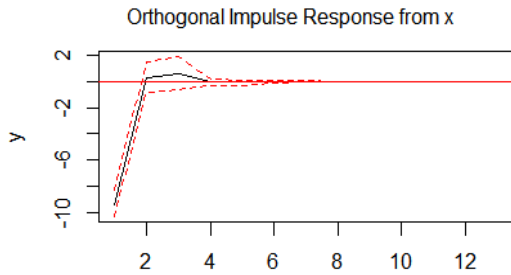
```
## JB-Test (multivariate)
##
## data: Residuals of VAR object var1
## Chi-squared = 14.581, df = 4, p-value = 0.005654
##
##
## $Skewness
##
## Skewness only (multivariate)
##
## data: Residuals of VAR object var1
## Chi-squared = 0.038716, df = 2, p-value = 0.9808
##
##
## $Kurtosis
##
## Kurtosis only (multivariate)
##
## data: Residuals of VAR object var1
## Chi-squared = 14.542, df = 2, p-value = 0.0006953
```

Slika 2.29. Multivarijantni test normalnosti reziduala u VAR(1) modelu

2.9. Pitanja za ponavljanje

1. Zašto se razmatraju VAR (vektorski autoregresijski) modeli?
2. Zapišite VAR(1) model u matricnom zapisu. Definirajte oznake koje pritom koristite.
3. Zapišite VAR(p) model u matricnom zapisu. Definirajte oznake koje pritom koristite.
4. Koje pretpostavke vrijede za vektor slučajnih procesa u VAR modelu?
5. Što je to strukturni oblik VAR modela?
6. Što je to reducirani oblik VAR modela?
7. Zašto se strukturni oblik VAR modela prevodi u reducirani oblik?
8. Zapišite sljedeće VAR modele matricno:
 - a) $y_{1t} = 1 - 0,5y_{1t-1} + 0,2y_{2t-1} + \varepsilon_{1t}$
 $y_{2t} = 3 + 0,1y_{1t-1} - 0,9y_{2t-1} + \varepsilon_{2t}$;
 - b) $y_{1t} = 20 + 5y_{1t-1} - 2y_{1t-2} - y_{2t-1} + 8y_{2t-2} + \varepsilon_{1t}$
 $y_{2t} = 10 - 5y_{1t-1} + 8y_{1t-2} + 3y_{2t-1} + 2y_{2t-2} + \varepsilon_{2t}$
 - c) $y_{1t} = 0,88 - 0,5y_{1t-1} + 0,4y_{2t-1} - 0,1y_{3t-1} + \varepsilon_{1t}$
 $y_{2t} = 0,27 + 0,5y_{1t-1} + 0,2y_{2t-1} - 0,4y_{3t-1} + \varepsilon_{2t}$
 $y_{3t} = 0,33 - 0,1y_{1t-1} - 0,1y_{2t-1} + 0,2y_{3t-1} + \varepsilon_{3t}$
9. Kada kažemo da je VAR model stabilan?
10. Zašto je bitna stabilnost VAR modela?
11. Što je to problem identifikacije kod VAR modela?
12. Koje su prednosti, a koji su nedostaci VAR modela?
13. Što je to multivarijatan Grangerov test uzročnosti? Po čemu se razlikuje od običnog Grangerova testa uzročnosti?
14. Zapišite nultu hipotezu multivarijatnog testa Grangerove uzročnosti za slučaj dvije varijable i dva pomaka u VAR modelu.
15. Zapišite nultu hipotezu multivarijatnog testa Grangerove uzročnosti za slučaj tri varijable i jedan pomak u VAR modelu.
16. Što su funkcije impulsnog odziva u VAR modelu? Kako se koeficijenti tih funkcija interpretiraju?
17. Što je to faktorizacija Choleskyog? Zašto se koristi? Što je pritom važno paziti kod provođenja te faktorizacije?
18. Interpretirajte sljedeće funkcije impulsnog odziva:





19. Što je to dekompozicija varijance u VAR modelu? Kako se vrši interpretacija?
 20. Interpretirajte sljedeće dekompozicije varijanci za odabrani VAR model:

a)

\$`x`			\$y		
	x	y		x	y
[1,]	1.0000000	0.000000000	[1,]	0.02779532	0.9722047
[2,]	0.9987965	0.001203485	[2,]	0.04098479	0.9590152
[3,]	0.9946909	0.005309068	[3,]	0.04727027	0.9527297
[4,]	0.9929008	0.007099180	[4,]	0.05393423	0.9460658

b)

\$`x`			\$y		
	x	y		x	y
[1,]	1.0000000	0.000000000	[1,]	0.3775670	0.6224330
[2,]	0.9924391	0.007560898	[2,]	0.3074415	0.6925585
[3,]	0.9835947	0.016405309	[3,]	0.3080549	0.6919451
[4,]	0.9827795	0.017220546	[4,]	0.2966338	0.7033662

21. Kako se odabire adekvatan broj pomaka p u VAR modelu?
 22. Što su to informacijski kriteriji? Čemu služe? Što se uzima u obzir kod izračuna informacijskih kriterija?
 23. Što je to konačna pogreška predviđanja (FPE)? Čemu služi?
 24. Nabrojite multivarijatne testove vezane uz inovacijski proces u VAR modelu.
 25. Učitajte datoteku „var.txt“ u RStudio. Datoteka sadrži mjesečne podatke o 4 varijable: x , y , z i w , s početnim datumom siječanj 2000. Definirajte vremenske nizove z i w i provedite sve postupke s interpretacijama kao u sveobuhvatnom primjeru 2.8.

Rješenja

Zadatak 8)

a)

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} -0,5 & 0,3 \\ 0,1 & -0,9 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

b)

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 20 \\ 10 \end{bmatrix} + \begin{bmatrix} 5 & -2 \\ -5 & 8 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} -1 & 8 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} y_{1t-2} \\ y_{2t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

c)

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} 0,88 \\ 0,27 \\ 0,33 \end{bmatrix} + \begin{bmatrix} -0,5 & 0,4 & -0,1 \\ 0,5 & 0,2 & -0,4 \\ -0,1 & -0,1 & 0,2 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

Zadatak 14)

$H_0: a_{12}^1 = a_{12}^2 = 0$ (varijabla y_2 ne uzrokuje varijablu y_1)

$H_0: a_{21}^1 = a_{21}^2 = 0$ (varijabla y_1 ne uzrokuje varijablu y_2)

Zadatak 15)

$H_0: a_{12} = 0$ (varijabla y_2 ne uzrokuje varijablu y_1)

$H_0: a_{13} = 0$ (varijabla y_3 ne uzrokuje varijablu y_1)

$H_0: a_{21} = 0$ (varijabla y_1 ne uzrokuje varijablu y_2)

$H_0: a_{23} = 0$ (varijabla y_3 ne uzrokuje varijablu y_2)

$H_0: a_{31} = 0$ (varijabla y_1 ne uzrokuje varijablu y_3)

$H_0: a_{32} = 0$ (varijabla y_2 ne uzrokuje varijablu y_3)

Zadatak 18)

Panel prvi redak i prvi stupac: šok od jedne standardne devijacije u varijabli x ima pozitivan učinak na varijablu y u prvom razdoblju, nakon čega vrlo brzo nestaje i postaje neznačajan.

Panel prvi redak i drugi stupac: šok od jedne standardne devijacije u varijabli x ima negativan učinak na varijablu y u prvom razdoblju, te alternira reakcija u predznacima i vrlo brzo nestaje.

Panel drugi redak i prvi stupac: šok od jedne standardne devijacije u varijabli x ima veoma jak negativan učinak na varijablu y u prvom razdoblju, nakon čega odmah postaje neznačajan.

Panel drugi redak i drugi stupac: šok od jedne standardne devijacije u varijabli x ima negativan učinak na varijablu y u cijelom promatranom razdoblju, pri čemu je s izuzećem za drugo razdoblje u svim drugima statistički značajan.

Zadatak 20)

- Uočavamo da je najveći udio varijacije u varijabli x upravo objašnjen šokovima u toj istoj varijabli, dok je veoma mali udio objašnjen šokovima u varijabli y . S druge strane, iako je i udio varijacije u varijabli y najviše objašnjen šokovima u toj samoj varijabli, ipak je nakon 4. razdoblja 5,39% varijacija od y pojašnjeno šokovima u varijabli x .
- Slično vrijedi i u postupku b), gdje je ponovno najveći udio varijacija u x pojašnjen šokovima u toj istoj varijabli, iako je sada 4. razdoblja nakon šoka u y ukupno 1,72% varijacija varijable x pojašnjeno tim šokovima. No, nakon 4. razdoblja je čak 29,66% varijacija varijable y pojašnjeno šokovima u varijabli x .

Zadatak 25)

```
nizevi<-read.table("var.txt",sep="\t",header = T)
z<-ts(nizevi$z,start=c(2000,1),frequency = 12)
w<-ts(nizevi$w,start=c(2000,1),frequency = 12)

library(urca)
summary(ur.df(z,type="drift",selectlags="BIC"))
```

```

##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.47488 -0.14789  0.01018  0.15145  0.50318
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.51073    0.06420   7.955 7.54e-14 ***
## z.lag.1      -0.59539    0.07313  -8.142 2.29e-14 ***
## z.diff.lag   -0.05164    0.06527  -0.791   0.43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2011 on 235 degrees of freedom
## Multiple R-squared:  0.3156, Adjusted R-squared:  0.3098
## F-statistic: 54.19 on 2 and 235 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic is: -8.1415 33.1455
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

summary(ur.df(w,type="drift",selectlags="BIC"))

##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.3903 -0.1295 -0.0065  0.1373  0.3076
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.38134    0.05379   7.090 1.56e-11 ***
## z.lag.1      -0.47791    0.06614  -7.226 6.92e-12 ***
## z.diff.lag   -0.07828    0.06461  -1.212   0.227
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1628 on 235 degrees of freedom
## Multiple R-squared:  0.2648, Adjusted R-squared:  0.2585
## F-statistic: 42.31 on 2 and 235 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic is: -7.2258 26.1062

```

```

##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

library(vars)
matrica<-cbind(z,w)
var1<-VAR(matrica,1,type="const")
summary(var1)

##
## VAR Estimation Results:
## =====
## Endogenous variables: z, w
## Deterministic variables: const
## Sample size: 239
## Log Likelihood: 142.311
## Roots of the characteristic polynomial:
## 0.4557 0.3906
## Call:
## VAR(y = matrica, p = 1, type = "const")
##
##
## Estimation results for equation z:
## =====
## z = z.l1 + w.l1 + const
##
##      Estimate Std. Error t value Pr(>|t|)
## z.l1  0.37211    0.06000   6.202 2.48e-09 ***
## w.l1  0.12247    0.06957   1.760  0.0796 .
## const 0.44089    0.07697   5.728 3.08e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.1996 on 236 degrees of freedom
## Multiple R-Squared: 0.1495, Adjusted R-squared: 0.1423
## F-statistic: 20.75 on 2 and 236 DF, p-value: 5.001e-09
##
##
## Estimation results for equation w:
## =====
## w = z.l1 + w.l1 + const
##
##      Estimate Std. Error t value Pr(>|t|)
## z.l1 -0.01259    0.04930  -0.255  0.799
## w.l1  0.47412    0.05717   8.294 8.39e-15 ***
## const 0.42923    0.06325   6.786 9.21e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.164 on 236 degrees of freedom
## Multiple R-Squared: 0.2259, Adjusted R-squared: 0.2193
## F-statistic: 34.43 on 2 and 236 DF, p-value: 7.558e-14
##
##
## Covariance matrix of residuals:
##      z      w
## z  0.039855 -0.001989
## w -0.001989  0.026912
##
## Correlation matrix of residuals:
##      z      w

```

```

## z 1.00000 -0.06073
## w -0.06073 1.00000

var2<-VAR(matrica,2,type="const")
summary(var2)

##
## VAR Estimation Results:
## =====
## Endogenous variables: z, w
## Deterministic variables: const
## Sample size: 238
## Log Likelihood: 144.178
## Roots of the characteristic polynomial:
## 0.6209 0.414 0.1211 0.1211
## Call:
## VAR(y = matrica, p = 2, type = "const")
##
## Estimation results for equation z:
## =====
## z = z.l1 + w.l1 + z.l2 + w.l2 + const
##
##      Estimate Std. Error t value Pr(>|t|)
## z.l1  0.35150    0.06553   5.364 1.96e-07 ***
## w.l1  0.11859    0.07990   1.484  0.139
## z.l2  0.05411    0.06517   0.830  0.407
## w.l2  0.01279    0.08026   0.159  0.873
## const 0.40513    0.09018   4.492 1.11e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2006 on 233 degrees of freedom
## Multiple R-Squared: 0.1523, Adjusted R-squared: 0.1377
## F-statistic: 10.46 on 4 and 233 DF, p-value: 8.228e-08
##
## Estimation results for equation w:
## =====
## w = z.l1 + w.l1 + z.l2 + w.l2 + const
##
##      Estimate Std. Error t value Pr(>|t|)
## z.l1 -0.03889    0.05327  -0.730  0.466
## w.l1  0.44219    0.06495   6.808 8.30e-11 ***
## z.l2  0.05969    0.05298   1.127  0.261
## w.l2  0.08383    0.06525   1.285  0.200
## const 0.36025    0.07331   4.914 1.68e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1631 on 233 degrees of freedom
## Multiple R-Squared: 0.2416, Adjusted R-squared: 0.2285
## F-statistic: 18.55 on 4 and 233 DF, p-value: 2.985e-13
##
## Covariance matrix of residuals:
##          z          w
## z  0.040236 -0.002228
## w -0.002228  0.026589
##
## Correlation matrix of residuals:
##          z          w

```



```
## z 1.0000 -0.0681
## w -0.0681 1.0000

VARselect(matrica,lag.max = 12,type="const")$selection

## AIC(n) HQ(n) SC(n) FPE(n)
##      1      1      1      1

library(lmtest)

causality(var1,cause="z")$Granger

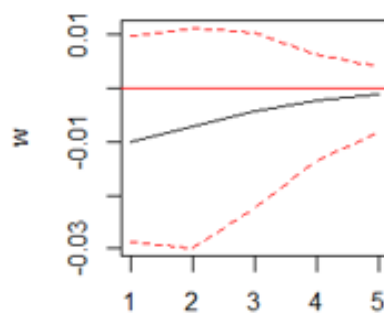
##
## Granger causality H0: z do not Granger-cause w
##
## data: VAR object var1
## F-Test = 0.065192, df1 = 1, df2 = 472, p-value = 0.7986

causality(var1,cause="w")$Granger

##
## Granger causality H0: w do not Granger-cause z
##
## data: VAR object var1
## F-Test = 3.0991, df1 = 1, df2 = 472, p-value = 0.07898

plot(irf(var1,n.ahead = 4,ortho=T,ci=.95,impulse="z",response="w"))
```

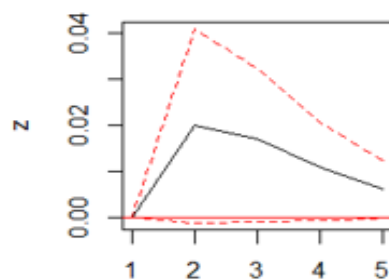
Orthogonal Impulse Response from z



95 % Bootstrap CI, 100 runs

```
plot(irf(var1,n.ahead = 4,ortho=T,ci=.95,impulse="w",response="z"))
```

Orthogonal Impulse Response from w



95 % Bootstrap CI, 100 runs

```

fevd(var1,n.ahead=4)

## $z
##           z           w
## [1,] 1.0000000 0.000000000
## [2,] 0.9911803 0.008819668
## [3,] 0.9851829 0.014817084
## [4,] 0.9827573 0.017242655
##
## $w
##           z           w
## [1,] 0.003687650 0.9963124
## [2,] 0.004595303 0.9954047
## [3,] 0.004965078 0.9950349
## [4,] 0.005087667 0.9949123

serial.test(var1, lags.pt=4, type="PT.asymptotic")

##
## Portmanteau Test (asymptotic)
##
## data: Residuals of VAR object var1
## Chi-squared = 10.086, df = 12, p-value = 0.6084

serial.test(var1, lags.pt=4, type="PT.adjusted")

##
## Portmanteau Test (adjusted)
##
## data: Residuals of VAR object var1
## Chi-squared = 10.202, df = 12, p-value = 0.5983

serial.test(var1, lags.bg=4, type="BG")

##
## Breusch-Godfrey LM test
##
## data: Residuals of VAR object var1
## Chi-squared = 12.609, df = 16, p-value = 0.7011

serial.test(var1, lags.bg=4, type="ES")

##
## Edgerton-Shukur F test
##
## data: Residuals of VAR object var1
## F statistic = 0.7724, df1 = 16, df2 = 454, p-value = 0.7174

arch.test(var1, lags.single=1, lags.multi=1, multivariate.only=F)

## $z
##
## ARCH test (univariate)
##
## data: Residual of z equation
## Chi-squared = 1.4643, df = 1, p-value = 0.2262
##
##
## $w
##
## ARCH test (univariate)
##
## data: Residual of w equation
## Chi-squared = 0.065674, df = 1, p-value = 0.7977
##
##
##

```

```
## ARCH (multivariate)
##
## data: Residuals of VAR object var1
## Chi-squared = 7.7129, df = 9, p-value = 0.5633

normality.test(var1,multivariate.only = T)

## $JB
##
## JB-Test (multivariate)
##
## data: Residuals of VAR object var1
## Chi-squared = 11.879, df = 4, p-value = 0.01827
##
##
## $Skewness
##
## Skewness only (multivariate)
##
## data: Residuals of VAR object var1
## Chi-squared = 0.538, df = 2, p-value = 0.7641
##
##
## $Kurtosis
##
## Kurtosis only (multivariate)
##
## data: Residuals of VAR object var1
## Chi-squared = 11.341, df = 2, p-value = 0.003446
```

AWN

AVN

3.

**KOINTEGRACIJA
U VEKTORSKIM
MODELIMA**

AVN

AWN

3. KOINTEGRACIJA U VEKTORSKIM MODELIMA

3.1. Temeljni pojmovi

U prethodnom poglavlju razmatrali su se stacionarni procesi u okviru VAR modeliranja. Često ekonomske analize iziskuju razmatranje odnosa između varijabli u dugome roku, pri čemu se promatraju varijable u razinama, što znači da su najčešće nestacionarni nizovi u pitanju. U odjeljcima 1.14. i 1.15. obrađivali su se model korekcije pogreške i kointegracija između dviju varijabli. U ovome poglavlju obrađuju se vektorski modeli korekcije pogreške (VECM – engl. *vector error correction model*) i **kointegracija u vektorskim modelima**. Ako postoji kointegracija u vektorskom modelu u dugome roku, to znači da promjene u varijablama uzrokovane odstupanjima od tog dugoročnog ravnotežnog stanja trebaju biti takve da ih neka sila „vuče“ prema ravnotežnom stanju. Takvo prilagođavanje nedostaje u VAR modelu koji je obrađen u prethodnom poglavlju. Međutim, VAR model može se zapisati u obliku vektorskog modela korekcije pogreške, što će se učiniti u ovome poglavlju.

Vektorski autoregresijski model (bez konstanti, vidjeti odjeljak 2.3. za konstante) može se zapisati na sljedeći način:

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{A}_2 \mathbf{y}_{t-2} + \dots + \mathbf{A}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad (3.1)$$

gdje je $\mathbf{y}_t = [y_{1t} \ y_{2t} \ \dots \ y_{Nt}]'$ N -dimenzionalni vektor koji sadrži N endogenih varijabli, $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p$ su $N \cdot N$ matrice autoregresijskih koeficijenata te je $\boldsymbol{\varepsilon}_t = [\varepsilon_{1t} \ \varepsilon_{2t} \ \dots \ \varepsilon_{Nt}]'$ vektor slučajnih procesa. Vektor slučajnih procesa se još naziva i multivarijatni čisti slučajni proces. Njegove karakteristike su obrađene u prethodnom poglavlju detaljno.

Sada razmotrimo model (3.1) na način da od lijeve i desne strane oduzmemo i dodamo istovremeno vrijednost $\mathbf{A}_p \mathbf{y}_{t-p+1}$ kako bi se dobio izraz:

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{A}_2 \mathbf{y}_{t-2} + \dots + \mathbf{A}_{p-2} \mathbf{y}_{t-p+2} + (\mathbf{A}_{p-1} + \mathbf{A}_p) \mathbf{y}_{t-p+1} - \mathbf{A}_p \mathbf{y}_{t-p+1} + \boldsymbol{\varepsilon}_t. \quad (3.2)$$

Sada istovremeno dodamo i oduzmemo izraz $(\mathbf{A}_{p-1} + \mathbf{A}_p) \mathbf{y}_{t-p+2}$, kako bi model u (3.2) u tom slučaju bio sljedeći:

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{A}_2 \mathbf{y}_{t-2} + \dots - \underbrace{(\mathbf{A}_{p-1} + \mathbf{A}_p)}_{=\Gamma_{p-2}} \Delta \mathbf{y}_{t-p+2} - \underbrace{\mathbf{A}_p}_{=\Gamma_{p-1}} \Delta \mathbf{y}_{t-p+1} + \boldsymbol{\varepsilon}_t, \quad (3.3)$$

te ponavljamo postupak za daljnje pomake sve dok ne dobijemo zapis:

$$\Delta \mathbf{y}_t = \Gamma \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad (3.4)$$

gdje je $\Gamma = \sum_{i=1}^p A_i - I$ i $\Gamma_i = -\sum_{j=i+1}^p A_j$. Model u (3.3) naziva se **vektorski model korekcije**

pogreške (engl. VEC – *vector error correction*), a član $\Gamma = \sum_{i=1}^p A_i - I$ član korekcije pogreške (engl. *error correction term*). Uočimo u zapisu (3.4) da se radi o VEC($p-1$) modelu, jer je zadnji autoregresijski član s pomakom $p-1$. Zato se kaže da **VAR(p) model ima VEC($p-1$) reprezentaciju**.

Napomena: U literaturi se koristi i oznaka Π za matricu Γ .

Ako razmatramo VAR(1) model:

$$\mathbf{y}_t = A_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (3.5)$$

kojemu dodamo i oduzmemo istovremeno \mathbf{y}_{t-1} :

$$\mathbf{y}_t = A_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t + \mathbf{y}_{t-1} - \mathbf{y}_{t-1}, \quad (3.6)$$

$$\mathbf{y}_t - \mathbf{y}_{t-1} = A_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t - \mathbf{y}_{t-1}, \quad (3.7)$$

$$\Delta \mathbf{y}_t = A_1 \mathbf{y}_{t-1} - \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t = (A_1 - I) \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (3.8)$$

$$\Delta \mathbf{y}_t = \Gamma \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (3.9)$$

te je model u (3.9) VEC(0) reprezentacija VAR(1) modela (uočimo da nema pomaka od $\Delta \mathbf{y}_t$ u izrazu (3.9)).

Analogno, VAR(2) model može se zapisati kao:

$$\Delta \mathbf{y}_t = \Gamma \mathbf{y}_{t-1} + \Gamma_1 \Delta \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t. \quad (3.10)$$

pa se u (3.10) radi o VEC(1) modelu.

3.2. Testiranje kointegriranosti u VEC modelu

Kointegriranost varijabli razmatra se fokusiranjem na matricu Γ u modelu (3.4):

$$\Delta \mathbf{y}_t = \Gamma \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t. \quad (3.11)$$

Naime, kako je $\boldsymbol{\varepsilon}_t$ stacionaran proces, ostali članovi u modelu moraju biti stacionarni da bi model bio korektno definiran.

Napomena: Model u (3.9) je specijalan slučaj VEC($p-1$) modela za kojeg vrijedi da je $p-1 = 0$, tj. uočimo da nema sume koja se nalazi u formuli (3.11).

Postoje tri slučaja koji se mogu dogoditi:

1. Vektorski proces y_t je stacionaran

U tom slučaju nije potrebno diferencirati varijable i razmatrati kointegraciju između njih (jer su sve stacionarne. Vidjeti odjeljak 2.16.). Razmatra se samo VAR(p) model između stacionarnih varijabli. Kada bi se razmotrio model u (3.11), rang matrice Γ bio bi jednak broju varijabli u modelu, tj. $r(\Gamma) = N$. Drugim riječima, ako je $r(\Gamma) = N$, vrijedilo bi: $\det(\Gamma) \neq 0$, odnosno proces y_t ne sadrži jedinične korijene (nije nestacionaran).

Napomena: Razmatranje determinante matrice Γ je multivarijatna analogija razmatranju parametra λ u univarijatnom testiranju nestacionarnosti procesa y_t u ADF jednadžbi (vidjeti odjeljak 2.10.7.): $\Delta y_t = \phi_0 + \lambda y_{t-1} + \varepsilon_t$, gdje se u nultoj hipotezi za nestacionarnost testiralo je li parametar $\lambda = 0$. U multivarijatnom slučaju, gdje se razmatra matrica Γ u (3.11), analiza se svodi na razmatranje determinante matrice Γ .

2. Vektorski proces y_t nije stacionaran i postoji kointegracija

Analizira se model (3.11). Za nestacionaran proces y_t su Δy_t i $\sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i}$ stacionarni, kao i sam proces ε_t . Ostaje razmatranje izraza Γy_{t-1} , tj. matrice Γ . Njegov rang iznosio bi $r(\Gamma) = k < N$, $k > 0$, tj. rang matrice manji je od broja varijabli u samome modelu, što znači da je $\det(\Gamma) = 0$ (proces y_t sadrži jedinične korijene).

Kako se radi o reduciranom rangu kvadratne matrice, moguće ju je prikazati kao umnožak dviju matrica α i β , koje su formata $(N \cdot k)$:

$$\Gamma = \alpha\beta', \quad (3.12)$$

gdje su **koeficijenti u β kointegracijski parametri** za N kointegriranih varijabli, a $\beta'y_{t-1}$ je linearna kombinacija tih varijabli, koja je stacionarna: $\beta'y_{t-1} \sim I(0)$.

Napomena: Vidjeti detaljnije odjeljak 2.15., gdje je općenito definirana kointegracija u dugome roku između varijabli:

$$\beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} = \beta x_t = 0, \quad (3.13)$$

gdje je β vektor parametara, a x'_t vektor razmatranih varijabli.

Sada se model u (3.11) može pisati na sljedeći način:

$$\Delta y_t = \alpha\beta y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t. \quad (3.14)$$

Dakle, koeficijenti u matrici β su parametri dugog roka, dok se elementi u matrici α interpretiraju kao **brzine korekcije pogreške**. Znači, ako je $0 < r(\Gamma) = k < N$, postoji k kointegracijskih relacija. **Elementi u matrici α tumače se na sljedeći način:** α_{ij} je utjecaj komponente j u vektoru pogreške u korekciji neravnoteže varijable i u vektoru y_t .

Napomena: Matrice α i β nisu jedinstvene: naime, moguće je odabrati regularnu matricu \mathbf{W} reda k takvu da je $\alpha\mathbf{W}'(\beta\mathbf{W}^{-1})' = \alpha\mathbf{W}'\mathbf{W}^{-1}\beta = \alpha I\beta = \alpha\beta = \Gamma$. Zato se za procjenu koeficijenata u tim matricama u okviru programske podrške odabiru elementi matrice \mathbf{W} takvi da su svi dijagonalni elementi jednaki 1. To ima za posljedicu prikaz svake varijable redom kao linearnu kombinaciju preostalih $N-1$ varijabli. Ovaj postupak naziva se **normalizacija**.

3. Vektorski proces y_t nije stacionaran i ne postoji kointegracija

Analizira se model (3.11). Za nestacionaran proces y_t su Δy_t i $\sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i}$ stacionarni, kao i sam proces ε_t . Ostaje razmatranje izraza Γy_t , tj. matrice Γ . Njezin rang iznosio bi $r(\Gamma) = 0$, jer ne postoji kointegracija među varijablama pa je 0 kointegracijskih relacija prisutno u modelu. U tom slučaju radi se o nul-matrici.

Analiza se može dalje nastaviti pomoću VAR modela nad diferenciranim varijablama koje su u tom slučaju stacionarne.

Upravo se matrica Γ razmatra u formalnom testu – **Johanesovom pristupu testiranja kointegracije** (Johansen 1988, Johansen i Juselius 1990). Johansenov pristup sastoji se od procjene matrice Γ i razmatranja ranga matrice $\hat{\Gamma}$, s obzirom na prethodno tri iznesena slučaja. Kako se u testu rang matrice razmatra pomoću svojstvenih vrijednosti matrice (vidjeti Dodatak u prvom dijelu), definirajmo još sljedeće. Ako su za kvadratnu matricu $\hat{\Gamma}$ definirane svojstvene vrijednosti $\lambda_1, \dots, \lambda_N$, te svojstveni vektori v_1, \dots, v_N , iz dijagonalizacije matrice $\hat{\Gamma}: \mathbf{P}\hat{\Gamma}\mathbf{P}^{-1} = \Lambda$, gdje je $\mathbf{P} = [v_1 \dots v_N]$, i Λ je dijagonalna matrica čiji su elementi na glavnoj dijagonali svojstvene vrijednosti $\lambda_1, \dots, \lambda_N$, slijedi: $\hat{\Gamma} = \mathbf{P}^{-1}\Lambda\mathbf{P}$, te je $r(\hat{\Gamma}) = r(\Lambda)$.

Kako je Λ dijagonalna matrica, njezin rang je jednak broju elemenata na glavnoj dijagonali koji nisu jednaki 0. S obzirom na prethodno obrađivan rang matrice Γ koji je jednak broju kointegracijskih jednadžbi, u Johanesovom pristupu testiranja se pronadu svojstvene vrijednosti matrice $\hat{\Gamma}$ da bi se testirao broj kointegracijskih relacija koji je sada jednak broju elemenata na glavnoj dijagonali matrice Λ koji su različiti od 0.

Dakle, procijeni se matrica Γ te se izračunaju njezine svojstvene vrijednosti $\lambda_1, \dots, \lambda_N$ i poredaju se od najveće do najmanje. Neka je poredak $\lambda_1, \dots, \lambda_N$ upravo takav: $\lambda_1 > \lambda_2 > \dots > \lambda_N$. Ako varijable nisu kointegrirane, tj. matrica Γ je nul-matrica, procjene svojstvenih vrijednosti će biti približno 0. Ako postoji k kointegracijskih relacija, tada postoji k svojstvenih vrijednosti koje su različite od 0.

Stoga se mogu definirati dvije test veličine (Johansen 1988):

$$\lambda_{trace}(k) = -T \sum_{i=k+1}^N \ln(1 - \hat{\lambda}_i), \quad (3.15)$$

za test pomoću traga matrice svojstvenih vrijednosti i

$$\lambda_{max}(k, k+1) = -T \ln(1 - \hat{\lambda}_{k+1}), \quad (3.16)$$

za test pomoću najveće svojstvene vrijednosti. T je broj opažanja, a $\hat{\lambda}_i$ je procijenjena i -ta svojstvena vrijednost matrice $\hat{\Gamma}$. Ako su sve svojstvene vrijednosti $\hat{\lambda}_i$ približno 0, i test veličine u (3.15) i (3.16) su približno 0, a veće vrijednosti $\hat{\lambda}_i$ vode većim vrijednostima test veličina.

Nulta hipoteza testa traga matrice pretpostavlja da je broj kointegracijskih vektora manji ili jednak k , dok alternativna pretpostavlja broj veći od k (ne nužno $k+1$). Ako je broj kointegracijskih vektora jednak k , upravo je k svojstvenih vrijednosti različito od 0, dok su $k+1$ -a, $k+2$ -a, itd. do N -te jednake 0. Zato simbolički možemo pisati nultu hipotezu na sljedeći način:

$$H_0 : \lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_N = 0 . \quad (3.17)$$

Nulta hipoteza testa pomoću najveće svojstvene vrijednosti pretpostavlja da je broj kointegracijskih vektora jednak k , dok alternativna pretpostavlja da je jednak $k+1$.

Kritične granice oba testa dobivene su Monte Carlo simulacijama u Johansen i Juselius (1990) i ovise o veličini $N-k$, broju nestacionarnih komponenti u modelu, o uključivanju ili ne determinističkih regresora (konstanta, trend) u svakoj jednadžbi. Ako su test veličine veće od kritičnih granica, odbacuje se nulta hipoteza pojedinog testa:

- U slučaju testa traga matrice odbacuje se nulta hipoteza da postoji maksimalno k kointegracijskih vektora u korist alternativne hipoteze
- U slučaju testa pomoću najveće svojstvene vrijednosti odbacuje se nulta hipoteza da postoji k kointegracijskih vektora

Primjer 3.1.

Učitajmo datoteku „vec.txt“ u RStudio. Datoteka sadrži mjesečne podatke o 4 varijable: y_1 , y_2 , y_3 i y_4 , s početnim datumom siječanj 2002. Provedimo najprije testiranje nestacionarnosti svih varijabli pomoću ADF testa, uz uključenu konstantu u DF jednadžbu.

- Provjerimo koji je optimalni pomak p u VAR(p) modelu za varijable y_1 i y_4 . U okviru VAR modela isključimo sve determinističke regresore. Temeljem odabranog pomaka p , provedimo Johansenov pristup testiranja kointegracije između odabranog para varijabli pomoću testa ranga matrice i testa najveće svojstvene vrijednosti. Razina značajnosti je 5%.
- Provjerimo koji je optimalni pomak p u VAR(p) modelu za varijable y_2 i y_4 . U okviru VAR modela isključimo sve determinističke regresore. Temeljem odabranog pomaka p , provedimo Johansenov pristup testiranja kointegracije između odabranog para varijabli pomoću testa ranga matrice i testa najveće svojstvene vrijednosti. Razina značajnosti je 5%.
- Provjerimo koji je optimalni pomak p u VAR(p) modelu za varijable y_1 i y_3 . U okviru VAR modela isključimo sve determinističke regresore. Temeljem odabranog pomaka p , provedimo Johansenov pristup testiranja kointegracije između odabranog para varijabli pomoću testa ranga matrice i testa najveće svojstvene vrijednosti. Razina značajnosti je 5%.
- Temeljem zaključka u b), procijenimo VEC($p-1$) model. Zapišimo procijenjene koeficijente matrica α (vektor brzine prilagodbe) i β (kointegracijski vektor), te izračunajmo matricu Γ . Zapišimo VEC model u matričnoj formi, te kointegracijsku jednadžbu.

- e) Interpretirajmo brzine prilagodbe iz svake jednačbe VEC modela.
- f) S obzirom na procijenjen VEC model procijenimo funkcije impulsnog odziva pomoću dekompozicije Choleskyog te interpretirajmo rezultate.
- g) Procijenimo VEC(1) model uz pretpostavku 2 kointegracijske jednačbe između varijabli y_1 , y_2 i y_3 . Zapišimo procijenjene koeficijente matrica α i β , te izračunajmo matricu $\hat{\Gamma}$. Zapišimo VEC model u matricnoj formi, te kointegracijske jednačbe. Interpretirajmo brzine prilagodbe iz svake jednačbe VEC modela.
- h) Provedimo proceduru u g) uz uključenu konstantu u dio kointegracijski dio modela (vidjeti idući odjeljak).
- i) S obzirom na procijenjen VEC model u postupku g) procijenimo funkcije impulsnog odziva pomoću dekompozicije Choleskyog te interpretirajmo rezultate.

Slika 3.1. predočava učitavanje podataka u RStudio te naredbe za ispis test veličina u ADF testu (zadnji retci na slici), koje redom iznose za razmatrane varijable: $-2,76$, $1,08$, $-2,39$ i $-2,43$. Kritična granica za 1% iznosi $-3,46$ i kako su sve test veličine veće od kritične granice, ne odbacuje se nulta hipoteza o nestacionarnosti svih nizova. Dakle, kako su varijable nestacionarne, može se razmotriti postoji li kointegracija među njima.

```
vec<-read.table("vec.txt",header = T, sep="\t")
y1<-ts(vec$y1,start=c(2002,1),frequency = 12)
y2<-ts(vec$y2,start=c(2002,1),frequency = 12)
y3<-ts(vec$y3,start=c(2002,1),frequency = 12)
y4<-ts(vec$y4,start=c(2002,1),frequency = 12)
library(urca)
summary(ur.df(y1,type="drift",selectlags="BIC"))@teststat[1,1];
summary(ur.df(y2,type="drift",selectlags="BIC"))@teststat[1,1];
summary(ur.df(y3,type="drift",selectlags="BIC"))@teststat[1,1];
summary(ur.df(y4,type="drift",selectlags="BIC"))@teststat[1,1]
```

Slika 3.1. Učitavanje podataka i ADF testovi

- a) Za varijable y_1 i y_4 su naredbe predočene na slici 3.2. Optimalan pomak u VAR modelu je $p = 1$, što znači da je za VEC reprezentaciju pomak $p - 1 = 0$. Napomenimo da se ovdje radi o specijalnom slučaju (vidjeti formulu (3.9)), u kojem nema pomaka Δy_t s desne strane jednačbi).

Zato je u okviru paketa tsDyn u naredbi VECM(...):

$$\text{VECM}(m1, \text{lag}=0, \text{include}="none", \text{LRinclude}="none", \text{estim}="ML")$$

kao prvi argument navedena matrica podataka ($m1$), za drugi argument je odabran pomak jednak 0 ($\text{lag}=0$), te su u kointegracijsku jednačbu ($\text{LRinclude}="none"$) kao i model korekcije pogreške ($\text{include}="none"$) isključeni svi deterministički regresori. Za spremljeni VEC model (naziv $vec1$) je provedeno testiranje kointegracije između odabranog para varijabli pomoću testa ranga matrice (u ispisu trace i trace_pval) i testa najveće svojstvene vrijednosti (eigen i eigen_pval).

```
m1<-cbind(y1,y4)
VARselect(m1,12,type="none")$selection

## AIC(n)  HQ(n)  SC(n)  FPE(n)
##      1      1      1      1

vec1<-VECM(m1,lag=0,include="none",LRinclude="none",estim="ML")
```

```
summary(rank.test(vec1))
##      r      trace trace_pval trace_pval_T      eigen eigen_pval
## 1 0 11.39036655  0.07075  0.07227 11.33785471  0.04653
## 2 1  0.05251183  0.87588  0.86742  0.05251183  0.86785
```

Slika 3.2. Testiranje kointegracije pomoću testa ranga matrice i testa najveće svojstvene vrijednosti, varijable y_1 i y_4

U nultoj hipotezi testova u prvome retku pretpostavlja se u nultoj hipotezi da je broj kointegracijskih relacija jednak 0 (u ispisu r , prvi redak 0). Test veličine iznose redom 11,39 i 11,34, dok odgovarajuće p -vrijednosti 0,071 i 0,043. Temeljem ispisa prvog testa uz razinu značajnosti od 5% odbacuje se nulta hipoteza o nepostojanju kointegracije između razmatranih varijabli, dok se za drugi test ona odbacuje (p -vrijednost u ovome slučaju manja je od razine značajnosti).

U praksi se može dogoditi da testovi daju različite ishode, stoga je potrebno napraviti dodatne analize, procijeniti različite specifikacije VAR i VEC modela te temeljem informacijskih kriterija i testiranja hipotezi donijeti konačnu odluku.

- b) Za varijable y_2 i y_4 su naredbe predočene na slici 3.3. Optimalan pomak u VAR modelu je ponovno $p = 1$, što znači da je za VEC reprezentaciju pomak $p - 1 = 0$. Sada su test veličine redom: 21,60 i 20,73, te pripadajuće p -vrijednosti: manje od 0,001. Stoga se uz razinu značajnosti od 5% odbacuje nulta hipoteza o nepostojanju kointegracije između odabranog para varijabli.

Nakon toga se prelazi na drugi redak ispisa testiranja kointegracije (zadnji redak na slici), gdje je nulta hipoteza testova sada da postoji 1 kointegracijski vektor. Test veličine iznose 0,87 svaka, s pripadajućim p -vrijednostima 0,42 i 0,40, što je veće od razine značajnosti 5%. Stoga se sada ne odbacuje nulta hipoteza postojanja 1 kointegracijskog vektora (ili ekvivalentno, da je rang matrice Γ jednak 1). Stoga u ovome slučaju zaključujemo da postoji jedna kointegracijska relacija između odabranog para varijabli i razmatra se VEC model u postupcima d) i e).

```
m2<-cbind(y2,y4)
VARselect(m2,12,type="none")$selection
## AIC(n) HQ(n) SC(n) FPE(n)
##      1      1      1      1

vec2<-VECM(m2,lag=0,include="none",LRinclude="none",estim="ML")
summary(rank.test(vec2))
##      r      trace trace_pval trace_pval_T      eigen eigen_pval
## 1 0 21.6002427  <0.001  <0.001 20.7344632  <0.001
## 2 1  0.8657795  0.4085  0.404  0.8657795  0.4042
```

Slika 3.3. Testiranje kointegracije pomoću testa ranga matrice i testa najveće svojstvene vrijednosti, varijable y_2 i y_4

- c) Za varijable y_1 i y_3 su naredbe predočene na slici 3.3. Optimalan pomak u VAR modelu je ponovno $p = 1$, što znači da je za VEC reprezentaciju pomak $p - 1 = 0$. Sada su test veličine redom: 9,89 i 9,82, te pripadajuće p -vrijednosti: 0,12 i 0,09. Stoga se uz razinu značajnosti od 5% ne odbacuje nulta hipoteza o nepostojanju kointegracije između odabranog para varijabli.

```

m3<-cbind(y3,y1)
VARselect(m3,12,type="none")$selection

## AIC(n)  HQ(n)  SC(n)  FPE(n)
##      1      1      1      1

vec3<-VECM(m3,lag=0,include="none",LRinclude="none",estim="ML")

summary(rank.test(vec3))

##   r      trace trace_pval trace_pval_T      eigen eigen_pval
## 1 0 9.89331842   0.1238   0.1266 9.82016071   0.08757
## 2 1 0.07315771   0.8485   0.8395 0.07315771   0.83994
    
```

Slika 3.4. Testiranje kointegracije pomoću testa ranga matrice i testa najveće svojstvene vrijednosti, varijable y_1 i y_3

- d) i e) Procijenjen je VEC($p-1$) model, gdje je $p = 1$, procijenjen je bio u okviru VECM(...) naredbe u c), a `summary(vec2)` daje potreban ispis za razmatranje i zapis samog modela, što je predočeno na slici 3.5.

Kointegracijska jednadžba dana je u ispisu na način da su svi članovi jednadžbe prebačeni na desnu stranu jednakosti, pa je sljedeći zapis jednadžbe (vidjeti napomenu o nejedinstvenosti matrica α i β , pa se stoga provodi postupak normalizacije):

$$y_{2t} + 48,15y_{4t} = 0 \rightarrow y_{2t} = -48,15y_{4t}.$$

Dakle, u dugome roku je učinak varijable y_4 na varijablu y_2 jednak $-48,15$. Sada možemo zapisati sljedeći matricni zapis VEC modela. Elementi matrice α dani su u dijelu ispisa ECT (engl. *error correction*), dok su kointegracijski parametri dani u dijelu ispisa „kointegracijski vektor“ (*cointegrating vector*), koji je raspisan preko kointegracijske jednadžbe. Stoga pišemo redom:

$$\alpha = \begin{bmatrix} 0,001 \\ 2,7 \cdot 10^{-5} \end{bmatrix}, \beta = \begin{bmatrix} 1 \\ 48,15 \end{bmatrix} \Rightarrow \Gamma = \alpha\beta' = \begin{bmatrix} 0,001 & 0,057 \\ 2,7 \cdot 10^{-5} & 0,001 \end{bmatrix}$$

Dodatno, naredbe `coefA(vec2)`, `coefB(vec2)` i `coefPI(vec2)` daju ispis matrica α , β i Γ .

Model je sljedeći:

$$\begin{bmatrix} \Delta y_{2t} \\ \Delta y_{4t} \end{bmatrix} = \begin{bmatrix} 0,001 & 0,057 \\ 2,7 \cdot 10^{-5} & 0,001 \end{bmatrix} \begin{bmatrix} y_{2t-1} \\ y_{4t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{2t} \\ \varepsilon_{4t} \end{bmatrix}$$

```

summary(vec2)

## #####
## ###Model VECM
## #####
## Full sample size: 220   End sample size: 219
## Number of variables: 2   Number of estimated slope parameters 2
## AIC 150.8868   BIC 161.0541   SSR 868.1138
## Cointegrating vector (estimated by ML):
##   y2   y4
## r1  1 48.15125
##
##
##           ECT
## Equation y2 0.0012(0.0003)***
## Equation y4 2.7e-05(0.0001)
    
```

```
summary(vec2)$coefMat
##           Estimate Std. Error t value Pr(>|t|)
## y2:ECT 1.190482e-03 0.0002568956 4.634106 6.167974e-06
## y4:ECT 2.694515e-05 0.0001054931 0.255421 7.986388e-01
```

Slika 3.5. Ispis VEC modela

Kako je vektor α s elementima brzine prilagodbe prema ravnotežnom stanju u dugome roku (slično kao parametar γ u EC modelu u odjeljku 5.14.), vrijednosti se interpretiraju na sljedeći način: Svakog mjeseca se 0,1% neravnoteže kratkog roka varijable y_2 korigira prema ravnotežnom stanju dugog roka zbog utjecaja te iste varijable, dok se $2,7 \cdot 10^{-3}\%$ neravnoteže kratkog roka varijable y_2 korigira prema ravnotežnom stanju dugog roka zbog utjecaja varijable y_4 . Konačno, za razmatranje značajnosti varijabli u EC relaciji, naredba u zadnjem dijelu ispisa na slici 3.5. predočava tablicu za provođenje t -testova.

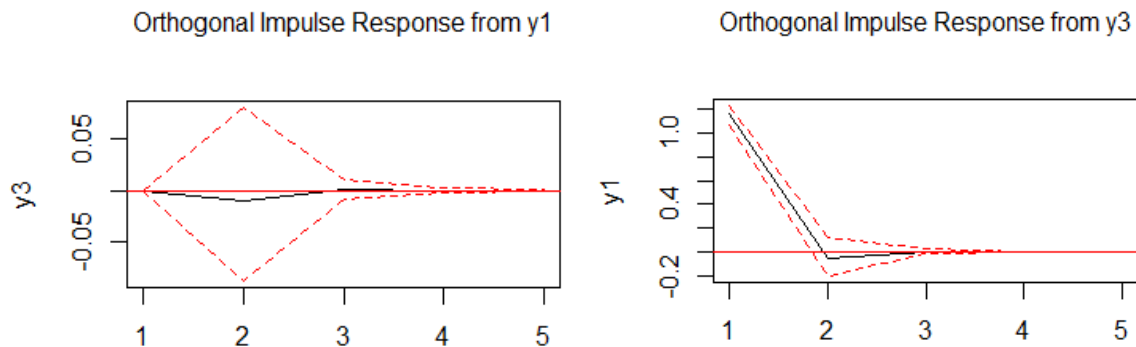
f) Za interpretaciju funkcija impulsnih odziva u okviru VEC modela iz prethodnog postupka, potrebno je uvažiti sljedeće napomene.

Napomena: ako se u okviru paketa tsDyn razmatra VAR(1) model, odnosno VEC(0) model, naredbe za IRF funkcije se moraju provesti za procijenjen VAR model (naredbe u okviru paketa tsDyn ne rade za $p-1 = 0$ u VEC modelu, već mora biti minimalno $p-1$ jednako 1). Dakle, potrebno je u tom slučaju procijeniti pripadajući VAR(1) model nad varijablama u prvim diferencijama i potom provesti naredbe koje su obrađene u prethodnom poglavlju vezane za VAR modele.

```
var<-VAR(diff(m3),p=1,type="none")
plot(irf(var,impulse="y1",response="y3",ci=0.95,n.ahead=4,ortho=T))
plot(irf(var,impulse="y3",response="y1",ci=0.95,n.ahead=4,ortho=T))
```

Slika 3.6. Naredbe za funkcije impulsnog odziva

Na slici 3.6. predočene su naredbe za grafičko predočavanje funkcija impulsnih odziva na slici 3.7. Najprije je procijenjen VAR(1) model nad diferenciranim podacima u tablici m3 i potom su procijenjeni koeficijenti u naredbama irf(...). Na slici 3.7 uočava se da šok od 1 standardne devijacije u varijabli y_1 ne rezultira sa značajnim reakcijama varijable y_3 (lijevi panel), dok šok od jedne standardne devijacije u varijabli y_3 rezultira s pozitivnom reakcijom varijable y_1 u prvome mjesecu od šoka, nakon čega šok iščezava.



Slika 3.7. Funkcije impulsnog odziva

Napomena: kao i kod VAR modela, potrebno je odrediti poredak varijabli za konstrukciju funkcije impulsnog odziva i u slučaju VEC modela.

g) Na slici 3.8 predočene su naredbe za procjenu VEC(1) modela (lag=1) i pretpostavku dvije kointegracijske jednadžbe (r=2). Sada su kointegracijske jednadžbe sljedeće:

$$\begin{aligned} y_{1t} &= -5,2 \cdot 10^{-18} y_{2t} + 2,78 y_{3t} \\ y_{2t} &= -50,36 y_{3t} \end{aligned}$$

Elementi matrica α i β su:

$$\alpha = \begin{bmatrix} 1,90 & 0,037 \\ -0,819 & -0,015 \\ 1,094 & 0,021 \end{bmatrix}, \beta = \begin{bmatrix} 1 & 0 \\ 5,2 \cdot 10^{-18} & 1 \\ -2,78 & 50,36 \end{bmatrix} \Rightarrow \Gamma = \alpha\beta' = \begin{bmatrix} 1,90 & 0,037 & -3,42 \\ -0,819 & -0,015 & 1,52 \\ 1,094 & 0,021 & -1,98 \end{bmatrix},$$

pa je VEC(1) model sljedeći:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} = \begin{bmatrix} 1,90 & 0,037 & -3,42 \\ -0,819 & -0,015 & 1,52 \\ 1,094 & 0,021 & -1,98 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} -2,24 & -0,06 & 3,89 \\ -1,50 & -0,11 & 2,52 \\ -1,26 & -0,04 & 2,18 \end{bmatrix} \begin{bmatrix} \Delta y_{1t-1} \\ \Delta y_{2t-1} \\ \Delta y_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y_{1t}} \\ \varepsilon_{y_{2t}} \\ \varepsilon_{y_{3t}} \end{bmatrix}$$

Sada parametri u prvome retku matrice α predstavljaju utjecaje pojedinih varijabli na prilagodbu prve varijable prema ravnoteži dugog roka, drugome retku su utjecaji pojedinih varijabli na prilagodbu druge varijable prema ravnoteži dugog roka, i slično tome se interpretiraju parametri u trećem retku.

```
m4<-cbind(y1,y2,y3)
vec4<-VECM(m4,lag=1,r=2,include="none",LRinclude="none",estim="ML")
summary(vec4)
```

```
## #####
## ###Model VECM
## #####
## Full sample size: 220    End sample size: 218
## Number of variables: 3    Number of estimated slope parameters 15
## AIC -1220.924    BIC -1163.388    SSR 1102.6
## Cointegrating vector (estimated by ML):
##   y1      y2      y3
## r1  1 5.20417e-18 -2.779373
## r2  0 1.00000e+00 50.364341
##
##
##          ECT1          ECT2          y1 -1
## Equation y1 1.9039(0.6537)**  0.0369(0.0126)** -2.2353(1.6519)
## Equation y2 -0.8194(1.0531)  -0.0146(0.0203)  -1.4975(2.6610)
## Equation y3 1.0944(0.3665)**  0.0212(0.0071)** -1.2550(0.9259)
##          y2 -1          y3 -1
## Equation y1 -0.0646(0.0494)  3.8852(2.9418)
## Equation y2 -0.1140(0.0796)  2.5234(4.7389)
## Equation y3 -0.0368(0.0277)  2.1778(1.6490)
```

Slika 3.3. Procjena VEC modela za slučaj 3 varijable

h) Sada slika 3.9. predočava prethodni model modificiran za uključivanje konstante u kointegracijski dio modela (include="const").

Kointegracijske jednadžbe su:

$$y_{1t} = -1,4 \cdot 10^{-17} y_{2t} + 1,14 y_{3t}$$

$$y_{2t} = 34,56 y_{3t}$$

Matrice α i β su:

$$\alpha = \begin{bmatrix} 1,14 & 0,03 \\ -0,94 & -0,02 \\ 0,69 & 0,02 \end{bmatrix}, \beta = \begin{bmatrix} 1 & 0 \\ 1,4 \cdot 10^{-17} & 1 \\ -1,14 & -34,56 \end{bmatrix} \Rightarrow \Gamma = \alpha\beta' = \begin{bmatrix} 1,14 & 0,03 & -2,17 \\ -0,94 & -0,02 & 1,74 \\ 0,69 & 0,02 & -1,31 \end{bmatrix},$$

pa je VEC(1) model:

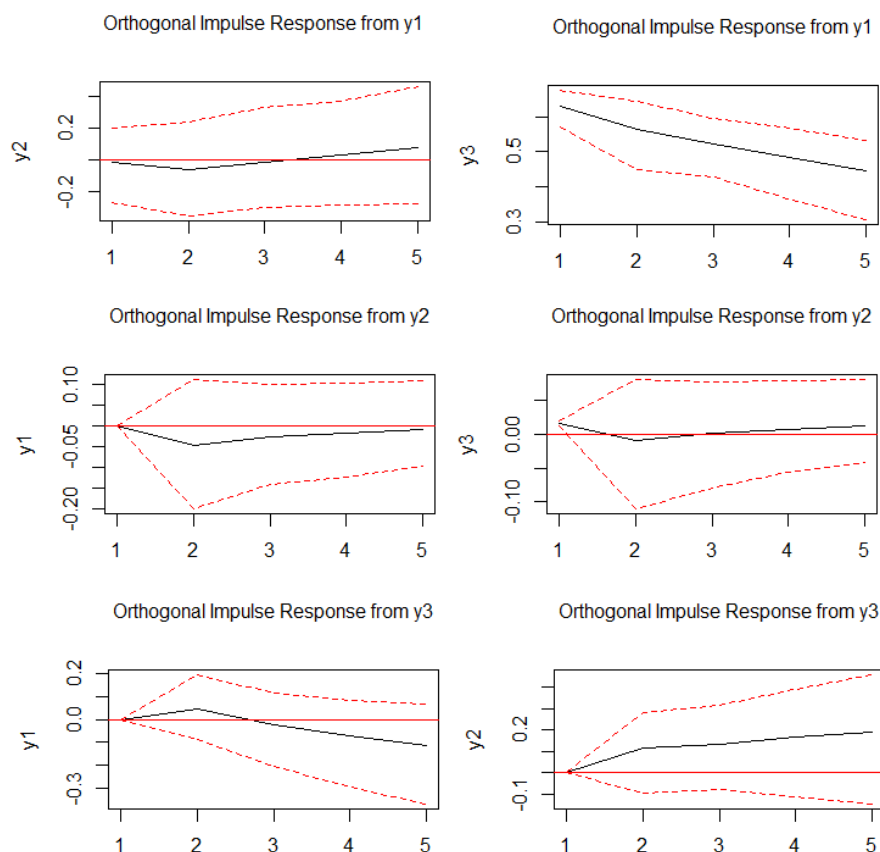
$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} = \begin{bmatrix} 0,97 \\ 0,24 \\ 0,51 \end{bmatrix} + \begin{bmatrix} 1,14 & 0,03 & -2,17 \\ -0,94 & -0,02 & 1,74 \\ 0,69 & 0,02 & -1,31 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} -2,28 & -0,07 & 4,02 \\ -1,54 & -0,11 & 2,61 \\ -1,28 & -0,04 & 2,25 \end{bmatrix} \begin{bmatrix} \Delta y_{1t-1} \\ \Delta y_{2t-1} \\ \Delta y_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y_{1t}} \\ \varepsilon_{y_{2t}} \\ \varepsilon_{y_{3t}} \end{bmatrix}$$

```
m4<-cbind(y1,y2,y3)
vec4<-VECM(m4,lag=1,r=2,include="const",LRinclude="none",estim="ML")
summary(vec4)

## #####
## ###Model VECM
## #####
## Full sample size: 220    End sample size: 218
## Number of variables: 3    Number of estimated slope parameters 18
## AIC -1230.322    BIC -1162.632    SSR 1096.422
## Cointegrating vector (estimated by ML):
##   y1          y2          y3
## r1  1 1.387779e-17 -1.141363
## r2  0 1.000000e+00 -34.561214
##
##
##          ECT1          ECT2          Intercept
## Equation y1 1.1378(0.7244)  0.0253(0.0133)  0.9702(0.4090)*
## Equation y2 -0.9386(1.1841) -0.0194(0.0218)  0.2403(0.6685)
## Equation y3 0.6890(0.4066)  0.0150(0.0075)*  0.5148(0.2296)*
##          y1 -1          y2 -1          y3 -1
## Equation y1 -2.2801(1.6348) -0.0670(0.0488)  4.0208(2.9118)
## Equation y2 -1.5436(2.6724) -0.1076(0.0797)  2.6075(4.7598)
## Equation y3 -1.2793(0.9177) -0.0380(0.0274)  2.2506(1.6346)
```

Slika 3.9. Procjena VEC modela za slučaj 3 varijable, uključene konstante u model korekcije pogreške

i)



Slika 3.10. Funkcije impulsnog odziva u VEC(1) modelu, 3 varijable

Temeljem procijenjenog modela u prethodnom postupku, na slici 3.10. predočene su funkcije impulsnih odziva. Sada u okviru paketa tsDyn možemo koristiti naredbu `irf(...)` za spremljeni VEC model (`vec4`), nije potrebno posebno procjenjivati VAR model. Uočava se da jedino šok od jedne standardne devijacije u varijabli y_1 rezultira sa značajnom negativno reakcijom varijable y_3 , koja jača tijekom razmatranog razdoblja. Ostale funkcije impulsnih odziva predočavaju neznačajne reakcije.

3.3. Deterministički regresori u VAR i VEC modelu

U prethodnom odjeljku razmatrao se VAR model i njegova VEC reprezentacija za slučaj vektorskog procesa y_t za kojeg se pretpostavljalo očekivanje jednako nul-vektoru (vektor konstanti je bio isključen iz razmatranja). Međutim, proces y_t može biti takav da je generiran i temeljem konstanti i/ili determinističkog trenda. Odabir modela u tome smislu ovisi o karakteristikama varijabli koje se razmatraju u analizi, kao i ekonomskoj teoriji i iskustvu istraživača.

Razmotrimo li slučaj u kojemu se u VAR model u (3.1) uključi vektor konstanti a_0 , tada VEC reprezentacija VAR modela izgleda ovako (Lütkepohl, 2006):

$$\Delta y_t = \Gamma(y_{t-1} - a_0) + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t, \quad (3.18)$$

$$\Delta \mathbf{y}_t = \boldsymbol{\alpha} \boldsymbol{\beta}' (\mathbf{y}_{t-1} - \mathbf{a}_0) + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad (3.19)$$

$$\Delta \mathbf{y}_t = \boldsymbol{\alpha} \boldsymbol{\beta}'_0 \begin{bmatrix} \mathbf{y}_{t-1} \\ 1 \end{bmatrix} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad (3.20)$$

$$\Delta \mathbf{y}_t = \boldsymbol{\Gamma}_0 \mathbf{y}_{t-1,0} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad (3.21)$$

gdje su $\boldsymbol{\Gamma}_0 = [\boldsymbol{\Gamma} \quad -\boldsymbol{\Gamma} \mathbf{a}_0]$, $\boldsymbol{\beta}'_0 = [\boldsymbol{\beta}' \quad -\boldsymbol{\beta}' \mathbf{a}_0]$ i $\mathbf{y}_{t-1,0} = \begin{bmatrix} \mathbf{y}_{t-1} \\ 1 \end{bmatrix}$. Dakle, ako se uključuje vektor konstanti u VAR modelu, on se „prelijeva“ u VEC specifikaciju također kao vektor konstanti, koju možemo pisati na način:

$$\Delta \mathbf{y}_t = \mathbf{v}_0 + \boldsymbol{\Gamma} \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad (3.22)$$

gdje vrijedi da je $\mathbf{v}_0 = \boldsymbol{\alpha} (-\boldsymbol{\beta}' \mathbf{a}_0)$.

Ako je vektorski proces \mathbf{y}_t definiran i pomoću konstante i linearnog trenda, tada se VEC specifikacija piše u obliku:

$$\Delta \mathbf{y}_t = \mathbf{v} + \boldsymbol{\Gamma}_+ \mathbf{y}_{t-1,+} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \quad (3.23)$$

gdje su $\mathbf{v} = -\boldsymbol{\Gamma} \mathbf{a}_0 + (I - \boldsymbol{\Gamma}_1 - \dots - \boldsymbol{\Gamma}_{p-1}) \boldsymbol{\mu}_1$, $\boldsymbol{\mu}_1$ je vektor koeficijenata uz varijablu trend t ,

$$\boldsymbol{\Gamma}_+ = \boldsymbol{\alpha} [\boldsymbol{\beta}' \quad -\boldsymbol{\beta}' \boldsymbol{\mu}_1], \mathbf{y}_{t,+} = \begin{bmatrix} \mathbf{y}_t \\ t \end{bmatrix}.$$

Prethodno analizirani modeli razmatrali su determinističke dijelove koji su bili uključeni samo u dio modela vezan uz korekciju pogreške. Ako razmatramo sljedeći zapis VEC modela:

$$\Delta \mathbf{y}_t = \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{y}_{t-1} + \underbrace{\boldsymbol{\alpha} \begin{bmatrix} \boldsymbol{\beta}' \\ a \\ b \end{bmatrix}}_{\text{kointegracijski prostor}} \begin{bmatrix} \mathbf{y}_{t-1} & 1 & t \end{bmatrix} + \underbrace{\mathbf{c} + \mathbf{d}t}_{\text{VEC dio modela}} + \boldsymbol{\varepsilon}_t, \quad (3.24)$$

uočava se da je osim dijela koji je prethodno razmatran (deterministički regresori u kointegracijskom prostoru), možemo uključiti i determinističke regresore u VEC dio modela, gdje su a i b parametri, te \mathbf{c} i \mathbf{d} vektori s konstantnim parametrima.

Restriktivni oblik modela je slučaj kada su $a = b = 0$, te \mathbf{c} i \mathbf{d} nul-vektori ostaje samo model:

$$\Delta \mathbf{y}_t = \sum_{i=1}^{p-1} \Gamma_i \Delta \mathbf{y}_{t-i} + \alpha \beta' \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (3.25)$$

koji je u prethodnom odjeljku obrađivan. Međutim, rijetko se primjenjuje u praksi. Drugi ekstremi slučaj je da su $a \neq 0$, $b \neq 0$ te c i d ne-nul vektori. To znači da postojanje linearnog trenda u VEC modelu opisuje kvadratno ponašanje trenda u podacima – što se također rijetko događa u praksi, i teško je izvršiti smislene ekonomske interpretacije.

Tri iduća slučaja su češća od prethodno spomenuta dva:

- Samo je $a \neq 0$ u (3.24), dok su $b = 0$ te c i d nul-vektori: varijable su kointegrirane oko konstantnog vektora.
- Samo su $a \neq 0$ i $c \neq \mathbf{0}$ u (3.25), dok su $b = 0$ i $d = \mathbf{0}$: varijable u prvim diferencijama nemaju jednaku očekivanu razinu (konstantu).
- Samo je $d = \mathbf{0}$ u (3.24), dok su $a \neq 0$, $b \neq 0$ i $c \neq \mathbf{0}$: uključen je deterministički trend u dio kointegracijskog prostora, jer se pretpostavlja da varijable u razinama imaju deterministički trend.

Primjer 3.2.

Temeljem podataka u primjeru 3.1., između varijabli y_2 i y_4 procijenimo VEC(0) model u kojemu je $a \neq 0$.

U okviru naredbe VECM(...) se ograničenja mogu mijenjati iz „none“ (slučaj gdje su $a = b = 0$ te $c = d = \mathbf{0}$) u „const“ za konstantu, „trend“ za trend i „both“ za uključivanje tih determinističkih regresora u pojedini dio modela. Dio koji se odnosi na „include=“ je kointegracijski prostor u (3.24), dok se „LRinclude=“ odnosi na VEC dio u (3.24), vidjeti sliku 3.11. U RStudiju je u okviru stupca Intercept dan umnožak $a \cdot \alpha$.

Kointegracijska jednadžba je: $y_{2t} = 40,12y_{4t}$. Elementi matrica α i β su:

$$\alpha = \begin{bmatrix} -0,002 \\ 0,002 \end{bmatrix}, \beta = \begin{bmatrix} 1 \\ -40,12 \end{bmatrix} \Rightarrow \alpha\beta' = \begin{bmatrix} -0,002 & 0,08 \\ 0,002 & -0,08 \end{bmatrix},$$

pa je VEC(0) model sljedeći:

$$\begin{bmatrix} \Delta y_{2t} \\ \Delta y_{4t} \end{bmatrix} = \begin{bmatrix} 0,01 \\ 0,60 \end{bmatrix} + \begin{bmatrix} -0,002 & 0,08 \\ 0,002 & -0,08 \end{bmatrix} \begin{bmatrix} y_{2t-1} \\ y_{4t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y_{2t}} \\ \varepsilon_{y_{4t}} \end{bmatrix}.$$

```
vec6<-VECM(m2, lag=0, include="const", LRinclude="none", estim="ML")
summary(vec6)
```

```
## #####
## ###Model VECM
## #####
## Full sample size: 220      End sample size: 219
## Number of variables: 2    Number of estimated slope parameters 4
## AIC 143.8678      BIC 157.4241      SSR 864.8431
## Cointegrating vector (estimated by ML):
##   y2      y4
## r1  1 -40.11634
##
##
##                               ECT                               Intercept
```

```
## Equation y2 -0.0018(0.0016)    0.0130(0.4772)
## Equation y4 0.0019(0.0006)** 0.6017(0.1914)**
```

Slika 3.11. VEC model za slučaj $a \neq 0$

Napomena: Kao i u slučaju modeliranja VAR modela, i kod procjene VEC modela potrebno je provesti adekvatne testove nad vektorom procijenjenih reziduala, vezanim uz karakteristike vektora slučajnih varijabli VAR modela.

3.4. Primjer

Učitajmo datoteku „vec_primjer.txt“ u RStudio. Datoteka sadrži mjesečne podatke o tri varijable: x , y_1 i y_2 , s početnim datumom siječanj 2003. Definirajmo naredbom `ts(...)` da se radi o vremenskim nizovima.

- a) Provedimo ADF testove jediničnog korijena za sve tri varijable, s uključenom konstantom kao determinističkim regresorom u DF jednadžbi i donesimo zaključak o ishodu testa uz razinu značajnosti od 5%.

```
x<-ts(vec$x,start=c(2003,1),frequency = 12)
y1<-ts(vec$y1,start=c(2003,1),frequency = 12)
y2<-ts(vec$y2,start=c(2003,1),frequency = 12)

library(urca)
summary(ur.df(x,type="drift",selectlags = "BIC"))
summary(ur.df(y1,type="drift",selectlags = "BIC"))
summary(ur.df(y2,type="drift",selectlags = "BIC"))
```

Slika 3.12. Učitavanje datoteke i provođenje ADF testova

DF test veličine redom iznose: $-1,88$, $-2,51$ i $-2,47$ za razmatrane varijable, dok je kritična jednaka $-2,88$, pa su sve tri veličine veće od kritične granice. Zato uz razinu značajnosti od 5% ne odbacujemo pojedinačne hipoteze da je pojedina varijabla nestacionarna. S obzirom da su sve varijable nestacionarne, postoji mogućnost kointegracije u dugome roku, između odabranog para varijabli ili pak između sve tri.

- b) Provjerimo koji je optimalni pomak p u $VAR(p)$ modelu za varijable x i y_1 . Isključimo sve determinističke regresore iz VAR modela. Temeljem odabranog pomaka p , provedimo Johansenov pristup testiranja kointegracije između odabranog para varijabli pomoću testa ranga matrice i testa najveće svojstvene vrijednosti. Razina značajnosti je 5%.

```
m1<-cbind(x,y1)
VARselect(m1,12,type="none")$selection
## AIC(n)  HQ(n)  SC(n)  FPE(n)
##      1      1      1      1
vec1<-VECM(m1,lag=0,include="none",LRinclude="none",estim="ML")
summary(rank.test(vec1))
##   r      trace trace_pval trace_pval_T      eigen eigen_pval
## 1 0 3.5299544    0.7674    0.7766 3.1534852    0.7593
## 2 1 0.3764692    0.6107    0.6021 0.3764692    0.6024
```

Slika 3.13. Testiranje kointegracije pomoću testa ranga matrice i testa najveće svojstvene vrijednosti, varijable x i y_1

Na slici 3.13. uočava se da je za odabrani par varijabli optimalni pomak u VAR modelu jednak 1, stoga je VEC(0) odgovarajuća reprezentacija VAR(1) modela. Procijenjen je VEC(0) model, te Johanesov pristup testiranja kointegracije na dnu ispisa slike 3.13. upućuje da se nulta hipoteza testa da ne postoji kointegracija između razmatranog para varijabli ne može odbaciti. Test veličine za oba testa redom iznose 3,53 i 3,15, s odgovarajućim p -vrijednostima 0,77 i 0,76, što je veće od razine značajnosti od 5%. U ovome slučaju ne postoji kointegracija između razmatranog para varijabli, no za varijable u prvim diferencijama se može razmotriti VAR model i analizirati postoji li povezanost u kratkome roku.

- c) Provjerimo koji je optimalni pomak p u VAR(p) modelu za varijable x i y_2 . Isključimo sve determinističke regresore iz VAR modela. Temeljem odabranog pomaka p , provedimo Johansenov pristup testiranja kointegracije između odabranog para varijabli pomoću testa ranga matrice i testa najveće svojstvene vrijednosti. Razina značajnosti je 5%.

```
m2<-cbind(x,y2)
VARselect(m2,12,type="none")$selection
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      1      1      1      1

vec2<-VECM(m2,lag=0,include="none",LRinclude="none",estim="ML")
summary(rank.test(vec2))

##      r      trace trace_pval trace_pval_T      eigen eigen_pval
## 1 0 147.590311 <0.001 <0.001 146.476284 <0.001
## 2 1  1.114027  0.3398  0.3369  1.114027  0.337
```

Slika 3.14. Testiranje kointegracije pomoću testa ranga matrice i testa najveće svojstvene vrijednosti, varijable x i y_2

Za drugi par varijabli je ponovno optimalan broj pomaka u VAR modelu 1, stoga je procijenjen VEC(0) model i u ovome slučaju (slika 3.14.). Sada se nulta hipoteza o nepostojanju kointegracije između ovog para varijabli odbacuje jer su pripadajuće p -vrijednosti u oba testa manje od 5%. Nadalje, nulta hipoteza da postoji 1 kointegracijska relacija između odabranog para varijabli se sada ne može odbaciti, jer su pripadajuće p -vrijednosti oba testa veće od 5% (iznose 0,34 svaka). Stoga u idućem koraku procjenjujemo VEC(0) model i interpretiramo njegove rezultate.

- d) Temeljem zaključka u c), procijenimo VEC($p-1$) model. Zapišimo procijenjene koeficijente matrica α (vektor brzine prilagodbe) i β (kointegracijski vektor), te izračunajmo matricu Γ . Zapišimo VEC model u matricnoj formi, te kointegracijsku jednadžbu.

```
summary(vec2)

## #####
## ###Model VECM
## #####
## Full sample size: 200      End sample size: 199
## Number of variables: 2      Number of estimated slope parameters 2
## AIC 4.860005      BIC 14.73992      SSR 525.1199
## Cointegrating vector (estimated by ML):
##      x      y2
## r1 1 -1.256386
##
##      ECT
```

```
## Equation x -0.1304(0.0570)*
## Equation y2 0.7132(0.0725)***

coefA(vec2); coefB(vec2);coefPI(vec2)

##          ECT
## Equation x -0.1304307
## Equation y2  0.7132276

##          r1
## x  1.000000
## y2 -1.256386

##          x          y2
## Equation x -0.1304307  0.1638712
## Equation y2  0.7132276 -0.8960889
```

Slika 3.15. VEC(0) model za varijable x i y_2 , bez determinističkih regresora

Kointegracijska jednadžba je: $x_t = 1,26y_{2t}$. Matrice α i β su:

$$\alpha = \begin{bmatrix} -0,13 \\ 0,71 \end{bmatrix}, \beta = \begin{bmatrix} 1 \\ -1,26 \end{bmatrix} \Rightarrow \alpha\beta' = \begin{bmatrix} -0,13 & 0,16 \\ 0,71 & -0,90 \end{bmatrix},$$

pa je VEC(0) model sljedeći:

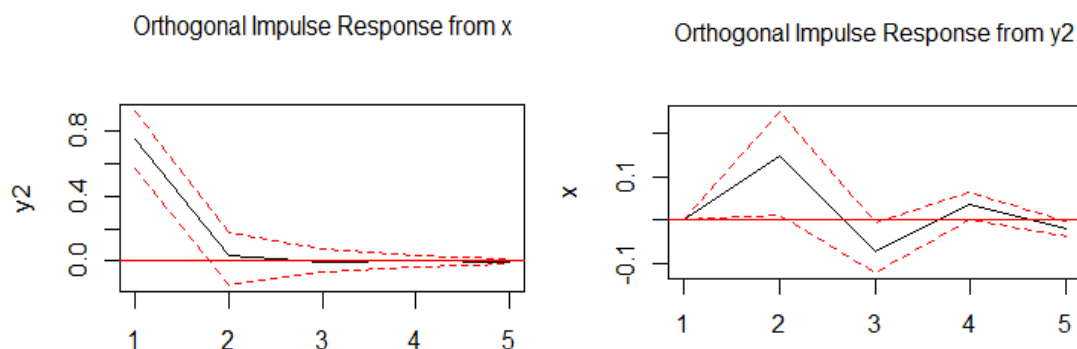
$$\begin{bmatrix} \Delta x_t \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} -0,13 & 0,16 \\ 0,71 & -0,90 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{y2t} \end{bmatrix}.$$

e) Interpretirajmo brzine prilagodbe iz matrice α VEC modela.

Kako je vektor α s elementima brzine prilagodbe prema ravnotežnom stanju u dugome roku (slično kao parametar γ u EC modelu u odjeljku 5.14.), vrijednosti se interpretiraju na sljedeći način: Svakog mjeseca se 13% neravnoteže kratkog roka varijable x korigira prema ravnotežnom stanju dugog roka zbog utjecaja te iste varijable, dok se 71% neravnoteže kratkog roka varijable x korigira prema ravnotežnom stanju dugog roka zbog utjecaja varijable y_2 .

f) S obzirom na procijenjen VEC model procijenimo funkcije impulsnog odziva pomoću dekompozicije Choleskyog te interpretirajmo rezultate.

```
var2<-VAR(diff(m2), type="none", p=1)
plot(irf(var2, impulse = "x", response = "y2", ci=.95, n.ahead=4, ortho = T))
plot(irf(var2, impulse = "y2", response = "x", ci=.95, n.ahead=4, ortho = T))
```



Slika 3.16. IRF funkcije u VEC(0) modelu

Kako je ovdje procijenjen VEC(0) model, u paketu tsDyn nije moguće procijeniti IRF funkcije za taj pomak, već za minimalno prvi pomak u VEC modelu, najprije je procijenjen VAR model s varijablama u prvim diferencijama i potom se razmatraju odgovarajuće IRF funkcije (slika 3.16.). Uočava se da šok od jedne standardne devijacije u varijabli x rezultira s pozitivnom reakcijom varijable y_2 samo u prvome mjesecu, te već u drugome učinak nestaje. S druge strane, šok od jedne standardne devijacije u varijabli y_2 u prvome mjesecu ne rezultira sa značajnom reakcijom varijable x , nakon toga u drugome mjesecu je reakcija pozitivna i značajna i nakon toga nestaje.

- g) Procijenimo VEC(0) model između x i y_2 uz uključenu konstantu u kointegracijski prostor. Zapišimo procijenjene koeficijente matrica α i β , te izračunajmo matricu Γ . Zapišimo VEC model u matricnoj formi, te kointegracijsku jednadžbu. Interpretirajmo brzine prilagodbe.

Kointegracijska jednadžba je: $x_t = 1.25y_{2t}$. Matrice α i β su:

$$\alpha = \begin{bmatrix} -0,13 \\ 0,72 \end{bmatrix}, \beta = \begin{bmatrix} 1 \\ -1,25 \end{bmatrix} \Rightarrow \alpha\beta' = \begin{bmatrix} -0,13 & 0,17 \\ 0,72 & -0,89 \end{bmatrix},$$

pa je VEC(0) model sljedeći:

$$\begin{bmatrix} \Delta x_t \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} -0,02 \\ 0,02 \end{bmatrix} + \begin{bmatrix} -0,13 & 0,17 \\ 0,72 & -0,89 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{y2t} \end{bmatrix}.$$

Svakog mjeseca se 13% neravnoteže kratkog roka varijable x korigira prema ravnotežnom stanju dugog roka zbog utjecaja te iste varijable, dok se 72% neravnoteže kratkog roka varijable x korigira prema ravnotežnom stanju dugog roka zbog utjecaja varijable y_2 .

Uočimo da konstante nisu značajne u modelu.

```
vec3<-VECM(m2,lag=0,include="const",LRinclude="none",estim="ML")
summary(vec3)
## #####
## ###Model VECM
## #####
## Full sample size: 200    End sample size: 199
## Number of variables: 2    Number of estimated slope parameters 4
## AIC 6.550309    BIC 19.72353    SSR 525.7622
## Cointegrating vector (estimated by ML):
##   x          y2
## r1 1 -1.247197
##
##
##           ECT           Intercept
## Equation x -0.1341(0.0575)* -0.0268(0.0715)
## Equation y2 0.7161(0.0733)*** 0.0246(0.0912)
coefA(vec3); coefB(vec3); coefPI(vec3)
##           ECT
## Equation x -0.1341477
## Equation y2 0.7161387
```



```
##          r1
## x    1.000000
## y2 -1.247197

##          x          y2
## Equation x -0.1341477  0.1673086
## Equation y2  0.7161387 -0.8931657
```

Slika 3.17. VEC(0) model za varijable x i y_2 , uključena konstanta u kointegracijski prostor

- h) Procijenimo VEC(1) model između x i y_2 uz uključenu konstantu u kointegracijski prostor. Zapišimo procijenjene koeficijente matrica α i β , te izračunajmo matricu Γ . Zapišimo VEC model u matricnoj formi, te kointegracijsku jednadžbu.

```
vec4<-VECM(m2, lag=1, include="const", LRinclude="none", estim="ML")
summary(vec4)

## #####
## ###Model VECM
## #####
## Full sample size: 200    End sample size: 198
## Number of variables: 2    Number of estimated slope parameters 8
## AIC 11.52204    BIC 41.11644    SSR 511.9875
## Cointegrating vector (estimated by ML):
##    x          y2
## r1 1 -1.256319
##
##
##          ECT                Intercept          x -1
## Equation x -0.0793(0.0827)    -0.0284(0.0718)    -0.0443(0.0848)
## Equation y2 0.7549(0.1043)*** -0.0084(0.0906)    0.0338(0.1070)
##          y2 -1
## Equation x  0.0706(0.0728)
## Equation y2 0.0662(0.0918)

coefA(vec4); coefB(vec4); coefPI(vec4)

##          ECT
## Equation x -0.07933001
## Equation y2  0.75489259

##          r1
## x    1.000000
## y2 -1.256319

##          x          y2
## Equation x -0.07933001  0.09966379
## Equation y2  0.75489259 -0.94838577
```

Slika 3.18. VEC(1) model za varijable x i y_2 , uključena konstanta u kointegracijski prostor

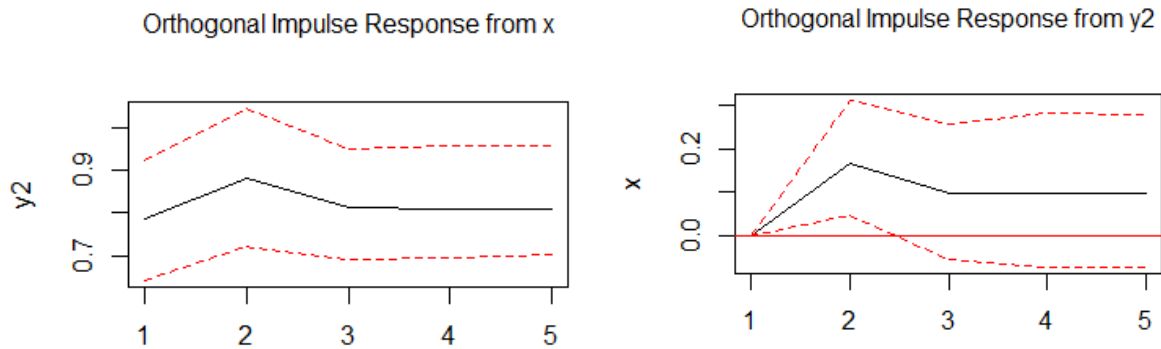
Kointegracijska jednadžba (slika 3.13.) je: $x_t = 1.26y_{2t}$. Matrice α i β su:

$$\alpha = \begin{bmatrix} -0,08 \\ 0,75 \end{bmatrix}, \beta = \begin{bmatrix} 1 \\ -1,26 \end{bmatrix} \Rightarrow \alpha\beta' = \begin{bmatrix} -0,08 & 0,10 \\ 0,75 & -0,95 \end{bmatrix},$$

pa je VEC(1) model:

$$\begin{bmatrix} \Delta x_t \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} -0,03 \\ 0,01 \end{bmatrix} + \begin{bmatrix} -0,08 & 0,10 \\ 0,75 & -0,95 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} -0,04 & 0,07 \\ 0,03 & 0,07 \end{bmatrix} \begin{bmatrix} \Delta x_{t-1} \\ \Delta y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x_t} \\ \varepsilon_{y_{2t}} \end{bmatrix},$$

- i) S obzirom na procijenjen VEC model u postupku h) procijenimo funkcije impulsnog odziva pomoću dekompozicije Choleskyog te interpretirajmo rezultate.



Slika 3.19. IRF funkcije u VEC(1) modelu

Uočavamo (slika 3.19.) da šok od jedne standardne devijacije u varijabli x rezultira sa značajnom pozitivnom reakcijom varijable y_2 u svim razmatranim razdobljima, dok šok od jedne standardne devijacije u varijabli y_2 rezultira samo s pozitivnom značajnom reakcijom varijable x u drugome mjesecu.

- j) Pretpostavimo da postoje 2 kointegracijske relacije između varijabli x , y_1 i y_2 . Procijenimo VEC(0) model, uz uključivanje konstante u VEC dio modela. Zapišimo procijenjene koeficijente matrica α i β , te izračunajmo matricu Γ . Zapišimo VEC model u matricnoj formi, te kointegracijsku jednadžbu. Interpretirajmo brzine prilagodbe.

Kointegracijske jednadžbe (slika 3.20.) su: $x_t = 3,5 \cdot 10^{-18} y_{1t} + 1,24 y_{2t} + 0,09$ i $y_{1t} = 0,34 y_{2t} + 7,81$. Matrice α i β su:

$$\alpha = \begin{bmatrix} -0,14 & -0,03 \\ 0,069 & -0,05 \\ 0,714 & -0,04 \end{bmatrix}, \beta = \begin{bmatrix} 1 & 0 \\ -3,5 \cdot 10^{-18} & 1 \\ -1,24 & 0,34 \\ 0,09 & 7,81 \end{bmatrix} \Rightarrow \alpha\beta' = \begin{bmatrix} -0,14 & -0,03 & 0,16 & -0,24 \\ 0,069 & -0,05 & -0,10 & -0,39 \\ 0,714 & -0,04 & -0,90 & -0,24 \end{bmatrix},$$

pa je VEC(0) model:

$$\begin{bmatrix} \Delta x_t \\ \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} -0,14 & -0,03 & 0,16 & -0,24 \\ 0,0069 & -0,05 & -0,10 & -0,39 \\ 0,714 & -0,04 & -0,90 & -0,24 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{1t-1} \\ y_{2t-1} \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{x_t} \\ \varepsilon_{y_{1t}} \\ \varepsilon_{y_{2t}} \end{bmatrix}.$$

```
m3<-cbind(x,y1,y2)
vec5<-VECM(m3,lag=0,r=2,include="none",LRinclude="const",estim="ML")
summary(vec5)

## #####
## ###Model VECM
## #####
## Full sample size: 200    End sample size: 199
## Number of variables: 3    Number of estimated slope parameters 6
## AIC 8.848515    BIC 35.19495    SSR 720.2231
## Cointegrating vector (estimated by ML):
##   x      y1      y2    const
## r1 1 -3.469447e-18 -1.2398906 0.09203653
## r2 0 1.000000e+00 0.3386793 7.81256533
##
##
```

```

##          ECT1          ECT2
## Equation x -0.1395(0.0576)* -0.0294(0.0215)
## Equation y1 0.0692(0.0576) -0.0514(0.0215)*
## Equation y2 0.7142(0.0735)** -0.0397(0.0274)

coefA(vec5);coefB(vec5);coefPI(vec5)

##          ECT1          ECT2
## Equation x -0.13951360 -0.02937913
## Equation y1 0.06924198 -0.05137222
## Equation y2 0.71416599 -0.03969189

##          r1          r2
## x      1.000000e+00 0.0000000
## y1     -3.469447e-18 1.0000000
## y2     -1.239891e+00 0.3386793
## const  9.203653e-02 7.8125653

##          x          y1          y2          const
## Equation x -0.13951360 -0.02937913 0.1630315 -0.2423667
## Equation y1 0.06924198 -0.05137222 -0.1032512 -0.3949760
## Equation y2 0.71416599 -0.03969189 -0.8989305 -0.2443661

```

Slika 3.20. VEC(0) model za varijable x , y_1 i y_2 , uključena konstanta u VEC prostor

3.5. Pitanja za ponavljanje

1. Zapišite VEC($p-1$) model u općenitom zapisu i definirajte oznake koje koristite.
2. Što je VEC model? Kada se koristi?
3. Kod razmatranja matrice Γ u VEC modelu, postoje tri slučaja u postupku testiranja kointegriranosti u modelu. Ukratko objasnite sva tri.
4. Kako se interpretiraju parametri u matrici α ?
5. Koja su dva testa koji se koriste u testiranju kointegriranosti varijabli u VEC modelu? Koja je nulta hipoteza svakoga testa?
6. Učitajte datoteku „vec_vjezba.txt“ u RStudio i ponovite sve postupke kao u sveobuhvatnom primjeru 3.4.

Rješenja

Zadatak 6)

```
summary(ur.df(x,type="drift",selectlags = "BIC"))

##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.033399 -0.007162  0.000106  0.008231  0.023410
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.15103    0.08032   1.880  0.0615 .
## z.lag.1      -0.03317    0.01761  -1.884  0.0611 .
## z.diff.lag   0.01591    0.07196   0.221  0.8253
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01064 on 195 degrees of freedom
## Multiple R-squared:  0.01788,    Adjusted R-squared:  0.007802
## F-statistic: 1.775 on 2 and 195 DF,  p-value: 0.1723
##
##
## Value of test-statistic is: -1.8839 1.8374
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

summary(ur.df(y1,type="drift",selectlags = "BIC"))

##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
```

```

## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.0263037 -0.0074288 -0.0000104  0.0069892  0.0269193
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.26555    0.10586   2.508  0.0129 *
## z.lag.1      -0.05854    0.02331  -2.512  0.0128 *
## z.diff.lag   0.02884    0.07152   0.403  0.6872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01077 on 195 degrees of freedom
## Multiple R-squared:  0.03134, Adjusted R-squared:  0.0214
## F-statistic: 3.154 on 2 and 195 DF, p-value: 0.04485
##
##
## Value of test-statistic is: -2.5118 3.2538
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

summary(ur.df(y2,type="drift",selectlags = "BIC"))

##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.050505 -0.010707  0.001342  0.011714  0.034205
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.34668    0.14030   2.471  0.0143 *
## z.lag.1      -0.07594    0.03070  -2.474  0.0142 *
## z.diff.lag   -0.28417    0.06755  -4.206 3.95e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01501 on 195 degrees of freedom
## Multiple R-squared:  0.1339, Adjusted R-squared:  0.125
## F-statistic: 15.08 on 2 and 195 DF, p-value: 8.172e-07
##
##
## Value of test-statistic is: -2.4737 3.1226
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

```

```

library(vars)
library(tsDyn)
m1<-cbind(x,y1)
VARselect(m1,12,type="none")$selection

## AIC(n)  HQ(n)  SC(n) FPE(n)
##      1      1      1      1

vec1<-VECM(m1,lag=0,include="none",LRinclude="none",estim="ML")

summary(rank.test(vec1))

##   r      trace trace_pval trace_pval_T      eigen eigen_pval
## 1 0 3.6812859    0.7470      0.7564 3.5098505    0.7057
## 2 1 0.1714354    0.7497      0.7400 0.1714354    0.7403

m2<-cbind(x,y2)
VARselect(m2,12,type="none")$selection

## AIC(n)  HQ(n)  SC(n) FPE(n)
##      8      3      1      8

vec2<-VECM(m2,lag=0,include="none",LRinclude="none",estim="ML")

summary(rank.test(vec2))

##   r      trace trace_pval trace_pval_T      eigen eigen_pval
## 1 0 74.44295028 <0.001      <0.001 74.37110487 <0.001
## 2 1 0.07184542  0.8502      0.8412 0.07184542  0.8416

summary(vec2)

## #####
## ###Model VECM
## #####
## Full sample size: 200   End sample size: 199
## Number of variables: 2   Number of estimated slope parameters 2
## AIC -3553.94   BIC -3544.06   SSR 0.06614567
## Cointegrating vector (estimated by ML):
##   x      y2
## r1 1 -0.9979015
##
##
##           ECT
## Equation x -0.1616(0.0536)**
## Equation y2 0.4497(0.0768)***

coefA(vec2); coefB(vec2);coefPI(vec2)

##           ECT
## Equation x -0.1615954
## Equation y2 0.4497499

##           r1
## x 1.0000000
## y2 -0.9979015

##           x      y2
## Equation x -0.1615954 0.1612563
## Equation y2 0.4497499 -0.4488061

vec2

##           ECT
## Equation x -0.1615954
## Equation y2 0.4497499

```

```

var2<-VAR(diff(m2),type="none",p=1)
plot(irf(var2,impulse = "x",response = "y2",ci=.95,n.ahead=4,ortho = T))

plot(irf(var2,impulse = "y2",response = "x",ci=.95,n.ahead=4,ortho = T))

vec3<-VECM(m2,lag=0,include="const",LRinclude="none",estim="ML")

summary(rank.test(vec3))

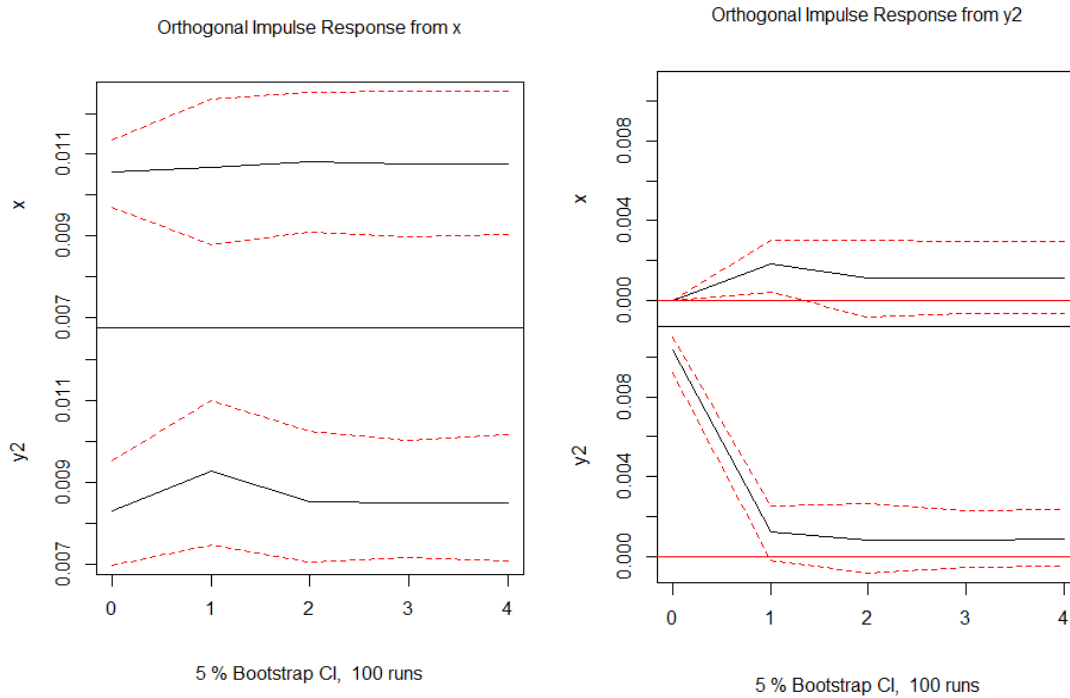
##      r      trace trace_pval trace_pval_T      eigen eigen_pval
## 1 0 151.192991 < 0.001 < 0.001 147.940980 < 0.001
## 2 1  3.252011  0.07134  0.07296  3.252011  0.07133

vec4<-VECM(m2,lag=1,include="const",LRinclude="none",estim="ML")
summary(rank.test(vec4))

##      r      trace trace_pval trace_pval_T      eigen eigen_pval
## 1 0  88.394001 < 0.001 < 0.001 84.709425 < 0.001
## 2 1  3.684576  0.05492  0.05632  3.684576  0.05492

plot(irf(vec4,impulse="x",response="y2",ci=0.95,n.ahead=4,ortho=T))

```



AWN

LITERATURA

1. Anderson, T. W., Darling, D. A. (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23, str. 193–212.
2. Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown changepoint. *Econometrica*, 61, str. 821–856.
3. Andrews, D. W. K., Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62, str. 1383–1414.
4. Balakrishnan, N., Koutras, M. V., Politis, K. G. (2020). *Introduction to Probability – Models and Applications*, New Jersey: Wiley, USA.
5. Breusch, T. S. (1978). Testing for Autocorrelation in Dynamic Linear Models. *Australian Economic Papers*, 17, str. 334–355.
6. Brooks, C. (2014). *Introductory Econometrics for Finance*, 3rd Edition, University Printing House, Cambridge, United Kingdom.
7. Canova, F. (2007). *Methods for Applied Macroeconomic Research*, Princeton University Press.
8. Chow, Gregory C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*, 28(3), str. 591–605.
9. Cochran, D., Orcutt, G. H. (1949). Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms. *Journal of the American Statistical Association*, 44(245), str. 32–61.
10. Cramér, H. (1928). On the Composition of Elementary Errors. *Scandinavian Actuarial Journal*, 1928 (1), str. 13–74.
11. EC (2018), European Commission, *Handbook on Seasonal Adjustment*. eurostat, Luxembourg. doi: 10.2785/941452.
12. Engle, R., Granger, C. W. J. (1991). Cointegration and Error Correction: Representation, Estimation, and Testing. In Engle and Granger (eds.), *Long Run Economic Readings in Cointegration*, Oxford University Press, New York, 1991, 81-113.
13. Engle, R., Granger, C. W. J., (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251-276.
14. Eurostat (2020). <https://ec.europa.eu/eurostat/data/database>
15. Farrar, D. E., Glauber, R. R. (1967). Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics*, 49, str. 92–107.
16. Farrar, D. E., Glauber, R. R. (1967). Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics*, 49(1), str. 92–107.
17. Glejser, H. (1969). A new test for heteroskedasticity. *Journal of the American Statistical Association*, 64(325), str. 316-323.
18. Goldberger, A. S. (1964). *Econometric Theory*. New York: John Wiley and Sons.
19. Granger, C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–38.
20. Greene, W. H. (2002). *Econometric analysis*, 5th edition, New Jersey: Prentice Hall, USA.
21. Greene, W. H. (2018). *Econometric analysis*, 8th edition, New York: Pearson.
22. Guidolin, M., Pedio, M. (2019). *Essentials of Time Series for Financial Applications*. UK: Academic Press.
23. Gujarati, D. N., Porter, D. C. (2010). *ESSEntials of Econometrics*, Fourth edition. New York: McGraw-Hill, USA.
24. Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press, Princeton.
25. Hayashi, F. (2002). *Econometrics*. Princeton University Press.

26. Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *Journal of the American Statistical Association*, 62(318), str. 399–402.
27. Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin.
28. Lütkepohl, H. (1993). "Introduction to Multiple Time Series Analysis", 2nd. ed., Berlin: Springer-Verlag.
29. Lütkepohl, H. (2000). "Bootstrapping impulse responses in VAR analyses". SFB 373 Discussion Paper, No. 2000, 22, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, Berlin.
30. Lütkepohl, H. (2006). "New Introduction to Multiple Time Series Analysis". Berlin: Springer.
31. Lütkepohl, H. (2010). "Vector Autoregressive Models". Economics Working Paper ECO 2011/30, European University Institute
32. Lütkepohl, H. (2011). "Vector Autoregressive Models". EUI Working Papers, ECO 2011/30, Department of Economics.
33. Jorgenson, D. W. (1966). Rational distributed lag functions. *Econometrica*, 34(1), 135-149.
- Kotzé, K. (2022). Autoregressive-distributed lag models, dostupno na: <https://kevin-kotze.gitlab.io/tsm/ts-9-note/>.
34. Wall, K. D. (1974). The rational distributed lag structural form – a general econometric model. NBER working paper No. 65, National Bureau of Economic Research.
35. Maddala, G. S. (1988). *Introduction to Econometrics*. Macmillan, New York.
36. Maddala, G. S. (1992). *Introduction to Econometrics*, 2nd edition, Toronto: Macmillan Pub.
37. Maddala, G. S., Lahiri, K. (2009). *Introduction to Econometrics*, 4th Edition, John Wiley and Sons.
38. Pesaran, M. H. (1987). *Econometrics*. The New Palgrave Dictionary of Economics, 1–25. doi:10.1057/978-1-349-95121-5_188-1.
39. Svjetska banka (2020). <https://databank.worldbank.org/>
40. Šego, B., Aljinović, Z., Marasović, B. (2011). *Financijsko modeliranje*. 2. izmijenjeno i dopunjeno izdanje. Split: Ekonomski fakultet – Split.
41. Šego, B., Škrinjarić, T. (2012). Modeliranje dnevne sezonalnosti prinosa na Zagrebačkoj burzi. U: *Matematički modeli u analizi razvoja hrvatskog financijskog tržišta*, Aljinović, Zdravka ; Marasović, Branka (ur.). Split: Ekonomski fakultet u Splitu, str. 159-172.
42. Šošić, I. (2006). *Statistika*, 2. Izmijenjeno izdanje. Zagreb: Školska knjiga.
43. von Mises, R. E. (1928). *Wahrscheinlichkeit, Statistik und Wahrheit*. Julius Springer.
44. Wei, W. S. (2005). *Time Series Analysis - Univariate and Multivariate Methods*, Second edition. USA: Addison Wesley.
45. Wooldridge, J. F. (2016). *Introductory Econometrics, A modern approach*, 6th edition. Boston: Cengage Learning, USA.
46. Abadie, A. (2005). "Semiparametric difference-in-differences estimators". *Review of Economic Studies*. 72 (1): 1–19
47. Austin, Peter C. (2011). "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies". *Multivariate Behavioral Research*. 46 (3): 399–424.
48. Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press.

POPIS POJMOVA

- ACF ..44, 57, 58, 59, 61, 64, 65, 67, 69, 70, 71, 72, 73, 74, 75, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 99, 105, 106, 107, 109, 110, 111, 114, 115, 116, 118, 119, 120, 121, 122, 123, 124, 125, 139, 140
- ARDL77, 144, 145, 146, 147, 148, 149, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 170, 174, 175, 176, 177, 178, 185, 186, 209
- autokorelacijska funkcija44, 45, 46, 47, 53, 57, 63, 67, 68, 69, 70, 73, 80, 98
- Autoregresivni model distribuiranih pomaka144
- autoregresivni proces.....54, 66, 79
- Cholesky dekompozicija214
- čisti slučajni proces51
- dekompozicija varijance.....222
- desezoniranje.....3, 4, 33, 34, 29, 30, 32, 34
- deterministički trend....75, 105, 106, 116, 137, 139, 143, 235, 260, 267, 268
- detrendiranje.....108, 112, 121
- Dickey-Fullerov test.....125, 126, 127, 129, 143
- diferenciranje.....121, 143
- egzogene varijable.....160, 200
- endogene varijable.....160, 197, 209
- funkcija impulsnog odziva77, 149, 151, 153, 157, 158, 175, 176, 177, 213
- funkcija impulsnog odziva ARMA procesa77
- Grangerova uzročnost ..173, 179, 180, 209, 210, 211, 212, 239, 243
- Hipoteza efikasnog tržišta117
- informacijski kriteriji.....227
- integriran proces110, 114, 115, 117, 135, 141
- inverzna karakteristična jednadžba55
- jedinični korijen.....127, 129
- Johansenov pristup258, 259, 269, 270
- kauzalnost.....160
- klasično desezoniranje.....34
- kointegracija164
- kointegracija u vektorskim modelima .. 255
- kointegracijski parametar..... 164
- korelogram..... 45, 46, 48, 52, 98, 105, 114, 124, 125
- logaritmiranje..... 3
- model korekcije pogreške 161, 170, 171, 172, 173, 180, 181, 186, 255, 256, 260, 265
- model slučajnog pomaka 114, 115
- multivarijatni test heteroskedastičnosti 232, 241
- multivarijatni test normalnosti..... 233, 242
- multivarijatni testovi autokorelacije 229, 241
- nestacionarnost 105, 124, 126, 128, 257
- PACF47, 48, 57, 58, 59, 61, 64, 65, 67, 69, 70, 71, 72, 73, 74, 75, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 99, 119, 139, 140
- pomični prosjek 33
- proces pomičnih prosjeka 67, 68
- reducirani model ... 198, 199, 205, 208, 243
- sezonski učinci..... 9
- slaba stacionarnost..... 43
- Stabilnost i stacionarnost modela 201
- stacionarni niz 3, 43
- stohastički trend3, 54, 75, 76, 77, 98, 105, 106, 109, 112, 143, 166, 169, 174
- stopa rasta 3, 4, 46, 47, 155, 156, 157, 158, 161, 162, 174, 175, 176, 177, 180, 186, 205, 211, 218, 219, 220, 221, 225
- striktna stacionarnost 43
- strukturni model... 198, 199, 208, 214, 216, 243
- test jediničnog korijena..... 125, 127, 165
- test omjera vjerodostojnosti..... 226
- trend . 1, 3, 9, 10, 14, 15, 29, 34, 29, 31, 32, 105, 106, 107, 108, 109, 112, 113, 120, 121, 123, 128, 129, 130, 131, 132, 135, 136, 137, 139, 142, 143, 164, 169, 174, 205, 228, 259, 267, 268
- uzročnost..... 160, 162, 210
- VAR77, 197, 198, 200, 201, 202, 203, 204, 205, 206, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221,

223, 224, 225, 226, 227, 228, 229, 231, 232, 233, 234, 236, 237, 239, 240, 241, 242, 243, 244, 255, 256, 257, 258, 259, 260, 261, 263, 264, 266, 267, 269, 270, 272	vektorski autoregresijski model.... 173, 197
vektorski	VMA 201
model korekcije pogreške.160, 173, 197, 229, 234, 243, 255, 256, 267	vremenski
	nizovi . 3, 6, 7, 12, 29, 33, 34, 31, 34, 43, 44, 54, 98, 105, 106, 143, 213, 239
	Waldov test..... 10, 19, 20

Izdavač:
Hrvatska narodna banka

Urednici:
Vedran Šošić
Davor Kunovac

Izvršni urednik:
Katja Gattin Turkalj

Recenzent:
izv. prof. dr. sc. Petar Sorić

Grafički urednik:
Slavko Križnjak

Dizajn naslovnice:
Vjekoslav Gjergja

Lektura predgovora:
Sanda Uzun-Ikić

Tehnički urednik:
Nevena Jerak Muravec

Fotografija:
Božidar Bengez

Pri korištenju ove publikacije obavezno navesti izvor.



HRVATSKA NARODNA BANKA

ODABRANE TEME PRIMIJENJENE
EKONOMETRIJE
**UVOD U ANALIZU VREMENSKIH
NIZOVA**

ISBN 978-953-8013-11-9

