



UNIVEZITET U PRIŠTINI
FAKULTET TEHNIČKIH NAUKA

Ivana Kostić Kovačević

**ANALIZA JEDNE KLASSE REKURZIVNIH STOHAŠTIČKIH
ALGORITAMA GAUSS-NEWTONOVOG TIPRA SA
PRIMENAMA U PARAMETARSKOJ IDENTIFIKACIJI FIR I
IIR DIGITALNIH SISTEMA**

Doktorska disertacija

Kosovska Mitrovica, 2008. godine

SADRŽAJ

1. Uvod.....	1
2. Modeliranje sistema.....	5
2.1. Definicija i osnovni zadaci modeliranja sistema (5)	
2.2. pristupi izgradnji modela sistema (6)	
2.2.1. Deduktivni pristup (7)	
2.2.2. Induktivni pristup (7)	
2.2.3. “Beli, sivi i crni” modeli (7)	
2.2.4. Klasifikacija modela sistema (10)	
2.2.4.1. Statički matematički modeli	
2.2.4.2. Dinamički matematički modeli	
2.2.4.2.1. Kontinualni modeli sa distribuiranim parametrima	
2.2.4.2.2. Kontinualni model sa koncentrisanim parametrima	
2.2.4.2.3. Kontinualni modeli sa događajima stanja	
2.2.4.2.4. Modeli diskretnih događaja	
2.2.4.2.5. Modeli sa semplovanim	
3. Black-box reprezentacije modela sistema.....	15
3.1. Neparametarske strukture black-box modela: Impulsni odziv i funkcija prenosa linearnog diskretnog vremenski invarijantnog sistema (16)	
3.2. Slučajni pobudni signali: Poremećaji u sistemu (22)	
3.2.1. Statističke i stohastičke karakteristične veličine za karakterizaciju poremećaja (23)	
3.2.1.1. Osnovne statističke karakteristične veličine i algoritmi za njihovu procenu	
3.2.2. Osnovne stohastičke karakteristične veličine i algoritmi za njihovu procenu (29)	
3.2.3. Osnovne osobine slučajnih pobudnih signala (33)	
3.2.4. Spektralna reprezentacija slučajnih signala (36)	
3.2.4.1. Spektar diskretnog belog šuma	
3.2.4.2. Transformacija spektra signala	
3.2.5. Opšta neparametarska reprezentacija black-box modela sistema:strukturni blok dijagram (42)	
3.3. Parametarska reprezentacija black-box modela sistema u vremenskom domenu (43)	
3.3.1. FIR model (model sa konačnim impulsnim odzivom (44)	
3.3.2. ARX model (Auto-regresioni model sa spoljašnjim ulazom (45)	
3.3.3. ARMAX model (Autoregresioni model pomičnih sredina sa spoljašnjom pobudom (47)	
3.3.4. OE model (Model greške izlaza) (48)	
3.3.5. BJ model (Box-Jenkinsov model (50)	

4. Parametarska identifikacija.....(53)

- 4.1. Modeli Predikcije: Optimalna srednje –kvadratna jednokoračna predikcija (58)
 - 4.1.1. Model predikcije za generalni parametarski model (58)
 - 4.1.2. Model predikcije za FIR reprezentaciju sistema (61)
 - 4.1.3. Model predikcije za ARX reprezentacije modela sistema (62)
 - 4.1.4. Model predikcije za ARMAX reprezentaciju modela sistema (63)
- 4.2. Parametarska identifikacija sistema: Algoritmi estimacije parametara (65)
 - 4.2.1. Linearni algoritmi estimacije parametara (65)
 - 4.2.1.1. Nerekurzivni linearni metod najmanjih kvadrata: LLS metod
 - 4.2.1.2. Nerekurzivni linearni metod težinskih (ponderisanih) najmanjih kvadrata: WLS metod
 - 4.2.1.3. Nerekurzivni pseudo-linearni metod najmanjih kvadrata: PLLS metod
 - 4.2.1.4. Rekurzivni linearni metod najmanjih kvadrata
 - 4.2.1.5. Pseudolinearni rekurzivni metod najmanjih kvadrata
 - 4.2.2. Nelinearni algoritmi estimacije parametara (86)
 - 4.2.2.1. Rekurzivni stohastički Newton-Raphsonov algoritam parametarske identifikacije.
 - 4.2.2.2. Rekurzivni nelinearni algoritam parametarske identifikacije tipa stohastičkog gradijenta
 - 4.2.2.3. Filtriranje greške predikcije

5. Analiza rekurzivnih stohastičkih algoritma parametarske identifikacije: konvergencija i konzistencija procena parametara i simulacija.....103

- 5.1. Analiza konvergencije linearnog algoritma najmanjih kvadrata (103)
- 5.2. Analiza konvergencije pseudo-linearnog metoda najmanjih kvadrata (PLLS) (111)
 - 5.2.1. Aproksimativna identifikacija (112)
 - 5.2.2. Filtriranje grešaka predikcije (116)
- 5.3. Analiza konvergencije nelinearnih rekurzivnih stohastičkih algoritama parametarske identifikacije (118)
 - 5.3.1. Analiza konvergencije nelinearnih rekurzivnih stohastičkih algoritama za estimaciju parametara u dinamičkim modelima sistema tipa linearne regresije (ARX i FIR modelima) primenom teorije martingala (118)
 - 5.3.2. Analiza konvergencije nelinearnih rekurzivnih stohastičkih algoritama za estimaciju parametara u dinamičkim modelima sistema tipa nelinearne regresije (ARMAX modelima) primenom teorije stabilnosti običnih diferencijalnih jednačina (140)

5.4. Eksperimentalna Analiza na osnovu računarske simulacije (165)

6. Optimalna parametarska

identifikacija.....171

6.1. Izbor optimalnog indeksa performanse na bazi raspoložive apriorne informacije o statističkim karakteristikama šuma (172)

6.2. Asimptotska analiza brzine konvergencije i određivanje optimalne težinske matrice kod rekurzivnih stohastičkih algoritama (182)

6.3. Statističko modeliranje poremećaja (194)

7. Adaptivna robusna estimacija parametara metodom QQ-krive.....205

7.1. QQ-metod estimacije statističkih parametara (205)

7.2. Formulacija problema parametarske estimacije (211)

7.3. Robusni rekurzivni algoritam tipa stohastičkog gradijenta (213)

7.4. Adaptacija robusnog rekurzivni algoritam tipa stohastičkog gradijenta (214)

7.5. Rezultati simulacije (219)

8. Zaključak.....221

Literatura

1. UVOD

Glavni zadatak modelovanja je da osvetli osnovne osobine realnog procesa koji se posmatra i da ih prevede na neki apstraktan jezik, pri čemu je uobičajeno da se u tu svrhu koristi jezik matematike. Na taj način, modelovanje predstavlja integralni deo nauke i tehnike, koji obuhvata sve oblasti ljudske aktivnosti, počevši od filozofije i teologije, pa preko psihologije, sociologije, ekonomije, ekologije, stići će do nauke i tehnike. Sam pojam modelovanja povezan je direktno sa pojmom procesa ili sistema, pri čemu se pod sistemom u najopštijem smislu podrazumeva izolovani deo realnosti koji je od interesa. U navedenom kontekstu sistem predstavlja subjektivni pojam i definisan je usvojenim granicama koje obuhvataju samo one osobine realnog procesa koje su od interesa za datu primenu. Ostale osobine realnog procesa koje nisu obuhvaćene izabranim ograničenjima određuju okolinu sistema. Na taj način, izbor navedenih granica predstavlja najvažniji korak u modelovanju, pošto suviše uske, kao i suviše široke granice, mogu rezultovati u praktično neprimenljive modele sa aspekta analize realnih pojava. U prvom slučaju ovakav rezultat biće posledica činjenice da se bitne osobine realnog procesa ne nalaze unutar usvojenih granica, dok će u drugom slučaju bitne osobine realnog procesa, iako su obuhvaćene izabranim granicama, biti zamaskirane nevažnim detaljima. Takođe je očigledno da postoji veliki broj modela kojima se može opisati jedna ista realna pojava, zavisno od aspekta njenog posmatranja.

Iako ne postoji generalna teorija modeliranja, u literaturi se mogu uočiti dva pristupa sintezi modela sistema.

Prvi od njih je deduktivni pristup, koji polazi od opštih saznanja o fizici realnog procesa i poznavanja idealizovanih fizičkih zakona kojima je aproksimativno opisano ponašanje posmatranog sistema. Na taj način se dolazi do kvalitativnog modela, kod koga su na osnovu poznatih fizičkih zakona definisane matematičke relacije između relevantnih fizičkih varijabli, ali u opštem slučaju nisu poznate i numeričke vrednosti određenog broja parametara koji figurišu u ovim jednačinama. Za takav model se kaže da je poznat sa tačnošću do nepoznatog vektora parametara i obično se u engleskoj literaturi naziva “*gray box*” model, a odgovarajući prevod na srpskom jeziku je model sive kutije. Ponekad je ipak na osnovu fizičkih gabarita realnog procesa moguće odrediti i numeričke vrednosti ovakvih parametara i takav model se u engleskoj literaturi naziva “*white box*” model, što se može prevesti na srpski jezik kao model bele kutije. Na primer, tipičan primer “*white box*” modela predstavljaju električna kola i sistemi.

Drugi pristup je induktivni i polazi od pretpostavke da se ne raspolaže sa adekvatnom apriornom informacijom o realnom procesu na osnovu koje bi se izabrali odgovarajući idealizovani fizički zakoni u cilju matematičkog prikaza relevantnih osobina posmatranog procesa. Stoga se ovaj korak dopunjuje apriornim usvajanjem strukture modela, koja obično predstavlja linearan vremenski kontinualan ili diskretan vremenski neporomenljiv sistem, koji je matematički opisan linearnom diferencijalnom ili diferencnom jednačinom sa fiksnim parametrima. Pošto konačan broj parametara u ovakvoj strukturi nema nikakvo fizičko tumačenje, ovakav model naziva se “*black box*” model, što se na srpski jezik može prevesti sa modelom crne kutije, pri čemu je i ovakav model kvalitativan i poznat sa tačnošću do nepoznatih parametara. Važan korak u ovakvom pristupu predstavlja estimacija parametara na osnovu raspoloživih realnih merenja na posmatranom sistemu, a ovaj zadatak se u literaturi naziva parametarska identifikacija. Ovakav pristup je vrlo popularan u

tehničkim naukama, s obzirom da linearni kontinualni i diskretni modeli navedenog tipa vrlo dobro aproksimiraju ponašanje tipičnih tehnoloških procesa oko neke izabrane radne tačke. Problem određivanja radne tačke obično se svodi na rešavanje nelinearnih algebarskih jednačina kojima je matematički opisano ponašanje tehnološkog procesa u nominalnom režimu rada.

U svetlosti rečenog, metode parametarske identifikacije za usvojene reprezentacije modela sistema različite fizičke prirode, koje su date u obliku linearnih vremenski nepromenljivih digitalnih sistema sa konačnim impulsnim odzivom (*Finite Impulse Response* ili skraćeno FIR) ili beskonačnim impulsnim odzivom (*Infinite Impulse Response* ili skraćeno IIR), predstavljaju izuzetno važan alat u nauci i tehnologiji. Ovako dobijeni modeli našli su široku primenu u raznorodnim zadacima analize tehnoloških procesa, njihove simulacije, kao i projektovanja sistema za obradu i prenos signala i upravljanje tehnološkim i industrijskim procesima. Modeli navedenog tipa popularni su i u računarskim telekomunikacijama, za prikazivanje redova čekanja poruka i podataka, kao i u ekonomiji u oblasti ekonometrije i finansijske matematike.

Kao generalni pristup za sintezu algoritama parametarske identifikacije u usvojenoj FIR ili IIR strukturi modela obično se koristi metod greške predikcije, koja predstavlja razliku između stvarne izmerene vrednosti izlaza realnog sistema i predikcije ove vrednosti na bazi samog modela. Zadatak parametarske identifikacije tada se postavlja kao problem optimizacije bez ograničavanja kriterijumske funkcije, koja predstavlja očekivanu vrednost neke nelinearne funkcije greške predikcije, kao slučajne veličine. Postavljeni optimizacioni problem, u opštem slučaju, ne može se rešiti analitički, u zatvorenoj formi, te je neophodno da se koriste alati numeričke matematike, a kao optimizaciona tehnika pogodno je da se koriste rekurzivne procedure tipa stohastičkog gradijenta. Posebno, ako je transformacija greške predikcije linearna funkcija, ovakvi rekurzivni stohastički algoritmi parametarske estimacije svode se na rekurzivni metod najmanjih kvadrata. Kao generalizacija ovakvog pristupa, u optimizacioni kriterijum se može uključiti i raspoloživa apriorna informacija o statističkim karakteristikama greške predikcije, posmatrane kao stohastička promenljiva, u formi klase raspodela kojoj pripada i realna funkcija raspodele verovatnoće razmatrane slučajne varijable. U robusnoj statistici ovakva informacija je korišćena za određivanje nelinearne transformaciji greške predikcije u odnosu na najnepovoljniju raspodelu u okviru pretpostavljene klase, koja nosi minimalnu informaciju o estimiranim parametrima, u smislu minimalne Fišerove količine informacije, odnosno maksimizacije Kramer-Raove donje međe. Ovakav pristup, u suštini, predstavlja generalizaciju Huberovog koncepta mini-maksne robusne estimacije u zadacima estimacije skalarnog parametra lokacije nepoznate funkcije raspodele verovatnoće slučajne varijable, koji se zasniva na minimizaciji asimptotske varijanse greške estimacije za najnepovoljniju raspodelu u okviru pretpostavljene klase raspodela kojoj i stvarna raspodela greške pripada.

U radu je ovakav pristup primenjen na usvojenu klasu FIR i IIR linearnih digitalnih vremenski invarijantnih modela sistema. S obzirom da je tako određen kriterijum, u opštem slučaju, nelinearna funkcija greške predikcije, kao sredstvo za više koračnu numeričku minimizaciju ovog kriterijuma korišćen je Gauss-Newtonov rekurzivni stohastički metod. Ovakvim postupkom je izbegnuta proizvoljnost i subjektivnost pri izboru algoritma parametarske identifikacije, a takođe su formalizovane i heurističke intervencije koje su uobičajene pri statističkoj analizi mernih podataka. Posebni značaj u navedenom pristupu ima teorijska analiza konvergencije, konzistentnosti i asimptotske brzine konvergencije procena parametara ka njihovim optimalnim vrednostima, koje odgovaraju minimumu usvojenog kriterijuma. Na bazi analize konvergencije dolazi se do egzaktnih teorijskih uslova

pod kojima će procene parametara, koje se generišu odabranim rekurzivnim stohastičkim algoritmom, konvergirati. Za razliku od postupaka simulacije, koji ne omogućava ekstenziju dobijenih rezultata na nove slučajeve koji nisu obuhvaćeni simulacijom, ovakav analitički rezultat ima univerzalnu vrednost i od korisnika se zahteva samo da proveri da li su izvedeni teorijski uslovi za konvergenciju algoritma parametarske identifikacije ispunjeni u razmatranom slučaju.

Za analizu konvergencije uobičajeno se u literaturi koriste sledeća dva pristupa:

1. Metod običnih diferencijalnih jednačina (*Ordinary Differential Equations* ili skraćeno ODE), koji se sastoji u pridruživanju rekurzivnom stohastičkom algoritmu odgovarajuće obične diferencijalne jednačine i primeni direktne metode Ljapunova za analizu stabilnosti rešenja ove jednačine.
2. Metod zasnovan na primeni teorije maringala, koji se zasniva na analizi stohastičke funkcije Ljapunova, koja predstavlja odgovarajuću funkciju slučajne greške estimacije parametara modela sistema, uz primenu osnovnih rezultata o konvergenciji maringala, kao specifičnih stohastičkih signala. Analiza asimptotske brzine konvergencije algoritma može se zasnivati na primeni centralne granične teoreme na relevantne stohastičke sekvence koje obrazuje maringal. Tako izvedeni analitički rezultat za asimptotsku totalnu varijansu greške estimacije omogućava optimalan izbor matrice pojačanja u rekurzivnom stohastičkom algoritmu parametarske estimacije, u smislu maksimalno moguće asimptotske brzine konvergencije procena parametara ka njihovim optimalnim vrednostima, koje odgovaraju minimumu usvojenog kriterijuma.

Sa praktičnog stanovišta, a u skladu sa centralnom graničnom teoremom matematičke statistike, posebno važnu klasu raspodela slučajnih poremećaja u sistemu predstavlja kontaminirana normalna klasa raspodela. Ovakva klasa okarakterisana je sa indeksom kontaminacije i varijansama kako nominalnog normalnog člana tako i kontaminirajućeg člana, čak i kada je ovaj član raspodeljen po normalnom zakonu raspodele verovatnoće. Pošto stepen kontaminacije i varijansa kontaminirajućeg člana kod aditivnog šuma u modelu sistema nikada nisu tačno poznate, a adekvatna procena ovih statističkih parametara je neophodna za definisanje nelinearne kriterijumske funkcije greške predikcije, u radu je predložena i modifikacija algoritma parametarske estimacije koja se zasniva na jednovremenoj estimaciji nepoznatih statističkih parametara šuma merenja sa postupkom estimacije nepoznatih parametara u usvojenoj strukturi modela razmatranog sistema, što rezultuje u novu klasu adaptivnih rekurzivnih stohastičkih algoritama parametarske identifikacije. Estimacija nepoznatih statističkih parametara slučajnih poremećaja u sistemu zasnovana je na metodologiji QQ-krive, što predstavlja novi pristup u zadacima parametarske identifikacije.

Pošto teoriska analiza osobine predloženih algoritama parametarske identifikacije ima asimptotski karakter, odnosno u važnosti je na uzorku velike dužine, pri čemu teoriska razmatranja ne ukazuju koliko velika dužina mernog uzorka treba stvarno da bude, analiza performansi algoritama na uzorku konačne dužine i u uslovima bliskim realnosti ispitivana je putem računarske simulacije. Posebna pažnja je posvećena praktičnim aspektima, kao što su planiranje eksperimenta u smislu selekcije ulaznih i izlaznih signala, izbora test pobudnih signala, izbora periode diskretnih merenja, analize mogućih tehnika prefiltriranja merenja, izbora reda usvojene FIR ili IIR strukture, procene kvaliteta estimacije parametara na bazi korelacione analize ulazno-izlaznih mernih podataka i reziduala merenja ili greške predikcije, kao i evaluacija dobijenog matematičkog modela IIR ili FIR sistema na osnovu statističke analize greške predikcije ili reziduala merenja.

Rad je formalno podeljen na uvod, šest poglavlja, zaključak i literaturu.

U **prvom poglavlju** dati su opšti aspekti matematičkog modeliranja, sa posebnim osvrtom na modeliranje linearnih diskretnih vremenski invarijantnih dinamičkih sistema različite fizičke prirode.

U **drugom polavlju** detaljno su razmatrane neparametarske i parametarske reprezentacije ovakvih tipova sistema.

U **trećem poglavlju** opisani su osnovni rekurzivni algoritmi za estimaciju nepoznatih parametara u usvojenoj strukturi modela sistema tipa stohastičkog gradijenta, kao što su algoritam dinamičke stohastičke aproksimacije i Njutn-Rafsonov ili Gaus-Njutnov algoritam.

U **četvrtom poglavlju** data je konvergencija predloženih algoritama, na bazi teorije martingala i metodologije analize stabilnosti rešenja običnih diferencijalnih jednačina. U istom poglavlju analizirane su i praktične osobine predloženih algoritama na uzorku konačne dužine i u uslovima bliskim realnosti pute Monte Karlo simulacije.

U **petom poglavlju** razmatrana je optimalna parametraska identifikacija na zadatoj klasi raspodela slučajnih poremećaja u sistemu.

U **šestom poglavlju** izvedena je nova klasa adaptivnih rekurzivnih algoritama parametarske estimacije na bazi jednovremene estimacije statističkih parametara šuma merenja, primenom metodologije QQ krive, i estmacije parametara modela sistema, na bazi rekurzivnih stohastičkih procedura razmatranih u prethodnim poglavljima. osobine predloženog algoritma ispitivane su eksperimentalno putem Monte Karlo simulacije na računaru opšte namene.

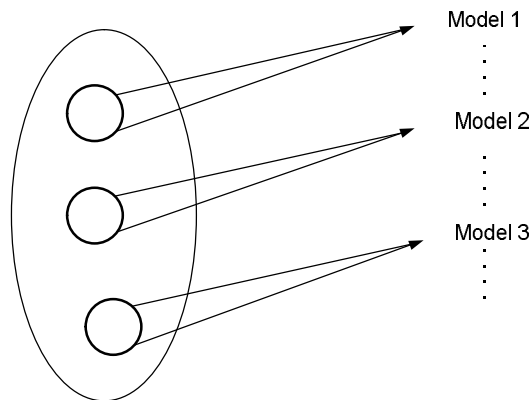
U **zaključku** su navedeni osnovni doprinosi rada i ukazano je na moguće pravce daljih istraživanja.

Literatura sadrži osnovne reference koje su korišćene prilikom istraživanja i pisanja samog teksta.

2. MODELIRANJE SISTEMA

2.1. DEFINICIJA I OSNOVNI ZADACI MODELIRANJA SISTEMA

Modeliranje je postupak dobijanja matematičkog opisa neke pojave koja se odvija u realnom svetu, kao što su na primer fizički, hemijski ili elektrotehnički procesi. Sa jedne strane, ovaj opis mora biti relativno jednostavan, a sa druge strane i dovoljno tačan, da bi odgovorio svojoj nameni koja je definisana od strane kreatora modela [7]. Primena modeliranja nije ograničena samo na tehničke procese (fizičke, hemijske, mehaničke, električke, itd.), već se ono koristi u različitim oblastima, kao što su ekonomija, biologija, sociologija, itd. Postoji praktično neograničen broj modela koji opisuju različite aspekte jedne realne pojave. Tipičan primer su električne komponente, kao što je otpornik, koji može da se proučava sa različitih stanovišta: a) kao električni element kojim je definisana linearna algebarska zavisnost između napona na njegovim krajevima i struje kroz otpornik; b) kao termički element čija se temperatura menja po zakonu koji je dat u formi linearne diferencijalne jednačine; c) kao translatorno mehaničko telo, određene mase na koga deluje sila gravitacije; d) kao rotaciono mehaničko telo, odgovarajućeg momenta inercije, kod koga je veza između obrtnog momenta i ugaonog ubrzanja data sa algebarskom relacijom [7]. Svi ti različiti pogledi na realnu pojavu rezultuju u različite modele, kao što je šematski prikazano na sledećoj slici.



Slika 2.1. Različiti pogledi na realnu pojavu rezultuju u njene različite modele

Suština postupka modeliranja je da se izaberu samo one osobine posmatranog procesa koje predstavljaju potrebne i dovoljne karakteristike da se proces opiše dovoljno tačno sa stanovišta namene modela. U svetlosti navedenog, zadatak modeliranja je da osvetli glavne osobine i fenomene realnog procesa i da ih prevede na neki apstraktan

jezik, kao što je jezik matematike [7,87]. Na taj način, modeliranje predstavlja integralan deo nauke i tehnologije, koji obuhvata skoro sve oblasti ljudskog delovanja,

počevši od filozofije i teologije, pa preko sociologije, psihologije, ekologije i ekonomije, doseže konačno i do same tehnike (građevine, hemije, fizike, mašinstva, elektrotehnike, itd.).

Pojam modeliranja je neraskidivo povezan sa pojmom procesa ili sistema. U navedenom kontekstu, sistem je skup stavki ili osobina koje predstavljaju zaokruženi deo nekog realnog fenomena koji se proučava. Pri tome, sistem je subjektivan pojam sa ograničenjima koja uključuju one osobine koje su najvažnije sa stanovišta modelara, a isključuje osobine od manjeg značaja za opis posmatranog realnog procesa. Na taj način, model predstavlja sredstvo za opisivanje najbitnijih karakteristika sistema koji se proučava. Uz navedeno, model mora posedovati prikaz objekata unutar sistema, tzv. komponenti sistema, kao i prikaz aktivnosti pod kojima će ti objekti međusobno delovati. Dakle, model reflektuje razumevanje realnog procesa, njegovih komponenti i njihove interakcije od strane samog modelara [7,29,43,87].

Važnu stavku prilikom formiranja modela predstavlja izbor ograničenja kojima je podvrgnut sistem. Ova ograničenja određuju koji će deo realnog procesa biti proglašen za sistem koji se proučava. Delovi realnog procesa koji nisu pridruženi sistemu, kao izolovanom delu realnosti, nazivaju se okolinom sistema. Ukoliko su granice sistema suviše široke može se dogoditi da je takav model sistema praktično nemoguće analizirati, pošto su mnoge važne osobine prekrivene nevažnim detaljima. Sa druge strane, ukoliko su izabrane granice sistema suviše uske, sve relevantne karakteristike realnog procesa neće biti obuhvaćene njegovim modelom, što će rezultovati u neadekvatnu analizu sistema na bazi takvog modela. Zahtevana tačnost prilikom modeliranja povezana je i sa namenom modela. Generalno, postoje dva načina primene modela sistema. Prvi od njih sastoji se u izvođenju eksperimenta u otvorenoj povratnoj sprezi. Kod ovakvog pristupa model se koristi za predviđanje budućih vrednosti relevantnih promenljivih u sistemu [23,29,30,43]. Na primer, model se može koristiti za izradu vremenske prognoze ili za predviđanje budućih vrednosti nekih ekonomskih pokazatelja, kao što su nezaposlenost, profitna stopa i sl. Zadatak predikcije zahteva adekvatan opis odgovarajućih zakona kojima su podvrgnute relevantne promenljive u modelu, tako da male greške modeliranja mogu prouzrokovati velike greške u rezultatu analize koja je izvršena na bazi takvog netačnog modela. Drugi pristup sastoji se u primeni modela u zatvorenoj povratnoj sprezi i ovakav pristup se obično koristi za projektovanje sistema upravljanja. Ukoliko je model deo sistema upravljanja u zatvorenoj povratnoj sprezi, takav sistem će redukovati dejstvo poremećaja, koji dolazi iz spoljašnje sredine, kao i same greške modeliranja. Na taj način, manje tačan model može se koristiti u ovoj drugoj primeni. Drugim rečima, zahtevana tačnost modeliranja zavisi od načina primene samog modela, odnosno da li se koristi u eksperimentu u otvorenoj povratnoj sprezi ili u okviru povratne petlje kod sistema upravljanja [16,24,80,93].

2.2. PRISTUPI IZGRADNJI MODELA SISTEMA

Postoji više načina da se izgradi model sistema, od kojih su najznačajniji sledeći pristupi [7]:

1. Deduktivni pristup (polazi od opšteg ka posebnom).
2. Induktivni pristup (polazi od posebnog ka opštem).
3. Bond graf (dijagram veza komponenti sistema).

2.2.1 Deduktivni pristup: Ovaj pristup pretpostavlja primenu opštih iskustava koja su stečena prilikom modeliranja različitih specifičnih procesa. Pristup koristi i prethodno znanje o razmatranom procesu, koje se zasniva na poznavanju fizičkih zakona koji definišu matematičke relacije između relevantnih varijabli u idealizovanom modelu procesa sa idealizovanim fizičkim komponentama. Na primer, u idealizovanim fizičkim komponentama telo odgovarajuće mase se tretira kao tačkasto, uz zanemarivanje njegovih dimenzija, protoci su laminarni, koncentracije su homogeno raspodeljene u rezervoaru, mešavine su idealne i sl. Fizički zakoni se obično izražavaju u obliku algebarskih i/ili diferencijalnih jednačina [87].

U opštem slučaju, algebarske relacije definišu ponašanje procesa u ustaljenom (ravnotežnom) stanju ili tzv. statičko ponašanje procesa. Statički modeli se koriste da se odredi radna tačka ili ravnotežno stanje procesa, na osnovu koje se dobija uvid u nominalne vrednosti relevantnih promenljivih, kao što su, na primer, pritisak, protok, temperatura ili profit. U industriji se statički modeli nazivaju “flow sheets” [7,87].

Dinamičko ponašanje procesa opisuje se diferencijalnim jednačinama [24,80,87,91]. Ove jednačine su zasnovane na osnovnim fizičkim zakonima, kao što su zakon o održanju energije, mase, momenta i sl. Apriorno (prethodno) znanje može da se koristi i da se bliže odredi struktura modela sistema, pošto pored relacije koja povezuje ulazne i izlazne varijable sistema i sama struktura modela predstavlja stavku od interesa. Na primer, ako je na osnovu prethodnog znanja ustanovljeno da sistem ima dve vremenske konstante, tada će to znanje biti iskorišćeno da se iz svih mogućih struktura modela drugog reda izabere ona kojoj odgovaraju dve realne vremenske konstante, odnosno dva realna pola, umesto da se za model usvoji opšta funkcija prenosa drugog reda koja može imati i konjugovano kompleksne polove. Navedeni pristupi obično dovode do kvalitativnog modela, u kome postoji određeni broj parametara čije vrednosti nisu poznate. Ponekada je, međutim, moguće da se procene adekvatno i vrednosti parametara u ovakvom modelu, obično koristeći prethodno znanje o fizičkim dimenzijama procesa.

2.2.2 Induktivni pristup: U opštem slučaju se ne raspolaže sa dovoljno apriornog znanja da bi se parametri u usvojenoj strukturi modela procenili adekvatno. U takvim situacijama koriste se tehnike parametarske identifikacije sistema, koje koriste merenja ulaza i izlaza sistema da bi estimirale (procenile) vrednosti parametara u modelu [23,24,30,43,63,64,68]. Postupak identifikacije zasniva se i na nekim dodatnim pretpostavkama, kao što su, na primer, klasa linearnih modela, selekcija

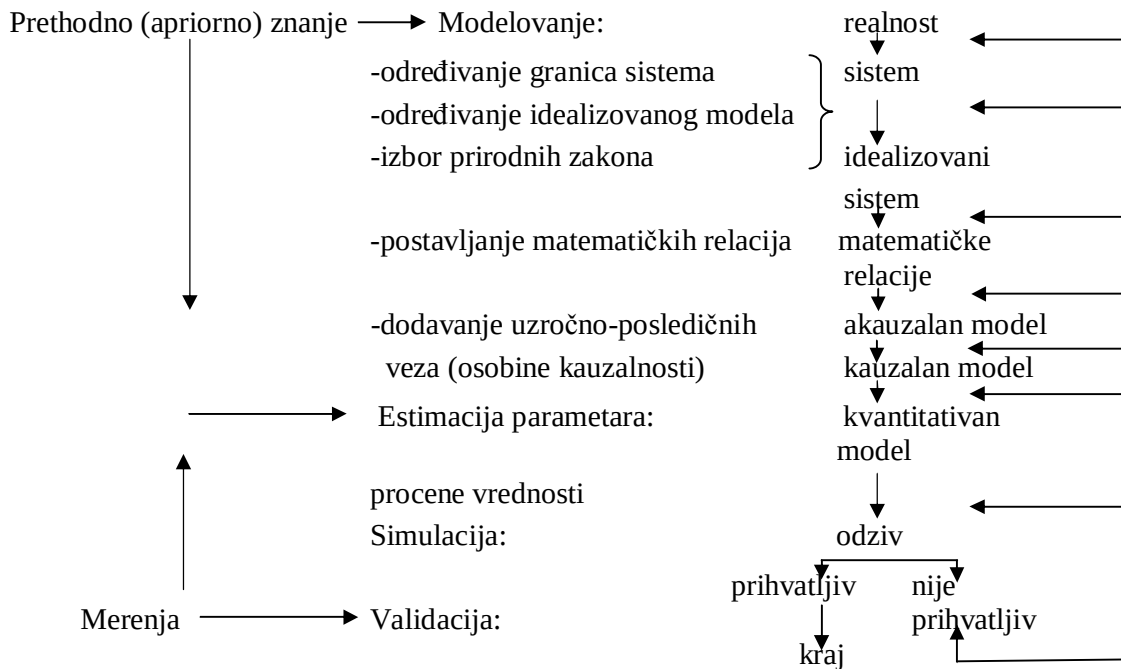
ulazno/izlaznih varijabli, red modela i sl. Sam postupak pribavljanja informacija o sistemu naziva se indukcija. U navedenom slučaju postavlja se prirodno i pitanje izbora kriterijuma za poređenje različitih modela u uslovima kada su merenja na procesu prisutna.

2.2.3 “Beli, sivi i crni” modeli: Ponekad je moguće da se model sistema izvede samo na osnovu deduktivnog pristupa, koristeći odgovarajuće idealizovane fizičke zakone i procenjene vrednosti parametara, na bazi fizičkih gabarita razmatranog procesa [7,24,80,87,91]. Takav model naziva se “beli” model, ” ili na engleskom jeziku “white-box” model. Primer takvih modela su električna i elektronska kola, a osnovni fizički zakoni koji se koriste prilikom modeliranja su Kirchofovi zakoni. U nekim slučajevima ne postoji adekvatno apriorno znanje o realnom procesu, te model mora da se postavi na osnovu raspoložive merne informacije o ulazu i izlazu sistema, ne posedujući adekvatnu informaciju o internoj strukturi i internim relacijama u sistemu [29,30,43,63,64,68,90]. Tako izveden model naziva se “crni model” ili na engleskom jeziku “black-box” model. Između ova dva granična slučaja nalazi se model u formi sive kutije ili na engleskom jeziku “gray-box” model, koji je u sebe uključio svu moguću raspoloživu apriornu informaciju o realnom procesu, što je rezultovalo u primeni odgovarajućih idealizovanih fizičkih zakona u kojima će se pojaviti određeni broj nepoznatih parametara. Pri tome, ovakva informacija zavisi uglavnom od polja primene modela i obično je veća u slučaju naučno-tehničkih procesa, nego u slučaju društvenih procesa. Tipične vrste modela u različitim oblastima date su na sledećoj slici [7].

	white box model	
elektrotehnologija	BELO	električna kola
mehnika		mehanički uređaji
hemija	NIJANSE	hemijski reaktor
ekologija	SIVOG	hidrologija (vodeni tokovi)
psihologija		biomedicina (dinamika ljudskog tela)
ekonomija	CRNO	ekonometrija (ekonomski modeli)
	black box model	

Slika 2.2. Tipovi modela u različitim oblastima: beli (white-box), sivi (gray-box) i crni (black-box) modeli

U opštem slučaju, izgradnja adekvatnog kvantitativnog modela sastoji se iz nekoliko faza: modeliranja, parametarske identifikacije, simulacije i validacije modela. Koraci (faze) u izgradnji modela prikazani su na narednoj slici.



Slika 2.3. Interaktivni postupak za dobijanje adekvatnog modela sistema

U bilo kojoj fazi opisanog postupka moguće je da se vrati jedan ili više koraka unazad, ukoliko rezultat ne zadovoljava postavljene zahteve. Poslednji korak validacije modela vrši se isključivo na osnovu realnih merenja na sistemu. Kao što je prikazano na slici 2.3, izgradnja kvantitativnog modela zahteva i uvođenje uzročno-posledične relacije (osobine kauzalnosti). Naime, same matematičke relacije definišu međusobno uticaj nekoliko varijabli u sistemu. Na primer, Omov zakon definiše vezu između napona i struje na otpornik i može se izraziti u obliku $u = Ri$, ili $i = u / R$. Obe relacije su sa matematičkog i fizičkog stanovišta iste, ali se razlikuju sa stanovišta uzroka i posledice, odnosno selekcije ulazne i izlazne varijable. Prva relacija tvrdi da će struja i (ulaz) prouzrokovati na krajevima otpornika napon u (izlaz), dok druga relacija tvrdi da će napon u (ulaz) rezultovati u struji i (izlaz) kroz otpornik. Selekcija ulazno-izlaznih varijabli naziva se postupkom uvođenja kauzalnosti (uzročno-posledičnih veza) u model. Kauzalnost nije fizički pojam i uvedena je veštačkim putem da bi se omogućila odgovarajuća istraživanja i izveo postupak simulacije [7,24,57,74,92]. Gotovo svi simulacioni paketi rade sa kauzalnim modelima. Kauzalni modeli koriste obične diferencijalne jednačine (engleski ordinary differential equations, ili skraćeno ODE) kao matematičke relacije, dok akauzalni modeli, koji iskazuju samo relacije između varijabli bez uvođenja uzročno-posledičnih veza (odnosno proglašavanje ulazno-izlaznih promenljivih) koriste diferencijalno-algebarske jednačine (engleski Differential Algebraic Equations ili skraćeno DAE). Neki noviji simulacioni programi koriste akauzalne modele [7].

Takođe, u cilju vizuelizacije veza koje postoje između varijabli u modelu koristi se grafička reprezentacija modela. Postoji više ovakvih prikaza, od kojih su neki zavisni od samog realnog procesa i njegovih komponenti, dok su drugi opšti i imaju univerzalnu primenu. Navedeni pristupi, polazeći od posebnog ka univerzalnom, prikazani su na sledećoj šemi [7].

- Crteži (autorova impresija o sistemu)
- Kola (vezana su za specifičnu primenu i zahtevaju standardni prikaz komponenti sistema)
- Dijagrami: imaju opšti karakter i mogu biti dvojaki:
 - a) Konsekutivni (sa redosledom izvođenja; primer su dijagram toka ili flow chart u programiranju ili PERT plan u organizacionim naukama)
 - b) Simultani (sa jednovremenim izvođenjem) koji mogu biti:
 - b1) kauzalni (primer su blok dijagrami i bond grafovi)
 - b2) akauzalni (primer su bond grafovi)

Dijagrami nisu vezani za specifičnu aplikaciju i predstavljaju prilično apstraktan način prikazivanja dinamike sistema. Za prikaz tehničkih sistema koriste se simultani dijagrami i to obično blok dijagrami. Kod blok dijagrama jednovremeno se uvode relacije između varijabli i definišu uzročno-posledične veze između njih (osobine kauzalnosti) [16,24,48,80]. Za razliku od blok dijagrama, bond graf dozvoljava dvokoračnu proceduru, gde se u prvom koraku definišu samo veze između varijabli, pa se tek u drugom koraku naknadno uvode uzročno-posledične veze između varijabli [7].

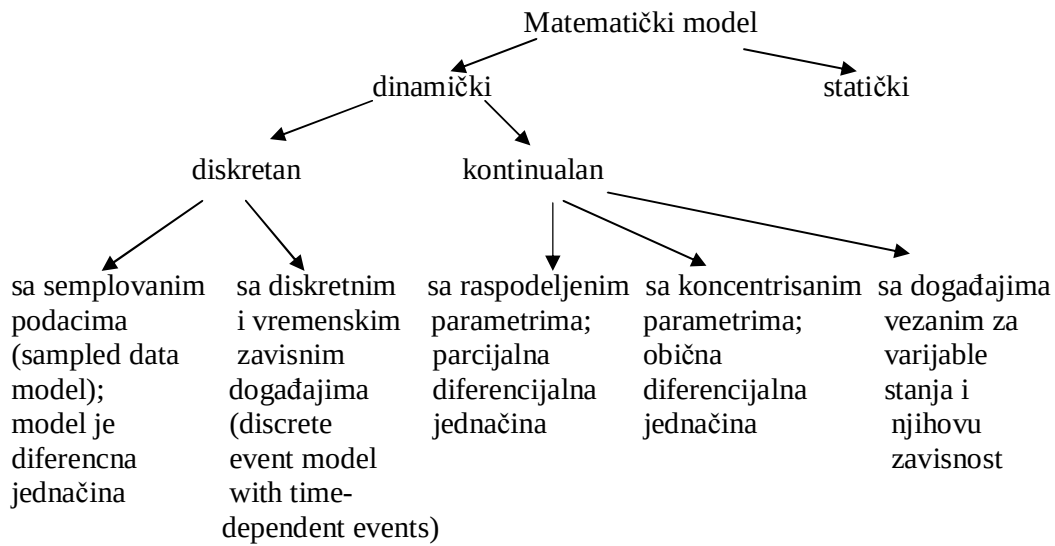
2.2.4 Klasifikacija modela sistema: Generalno se razlikuju tri različita tipa modela koji se mogu koristiti za opisivanje dinamičkog ponašanja realnog procesa : [7,87]

1) **Skalirani (proporcionalno umanjeni) model:** predstavlja proporcionalno umanjenu fizičku maketu realnog procesa, koja se koristi za ekonomično ispitivanje osobina realnog procesa (cena eksperimentisanja na ovakvom modelu je značajno niža nego na realnom procesu). Ovakav model se koristi kada ne postoji dovoljno tačan matematički model ili ukoliko izračunavanja u okviru postavljenog matematičkog modela traju neprihvatljivo dugo. Primeri ovakvog modela su aerodinamički tuneli, rezervoari za ispitivanje vodenih tokova i sl.

2) **Matematički model:** dinamičko ponašanje sistema je u opštem slučaju opisano skupom nelinearnih diferencijalnih jednačina, o čemu će kasnije biti više reči i oni su dominantni u tehnici.

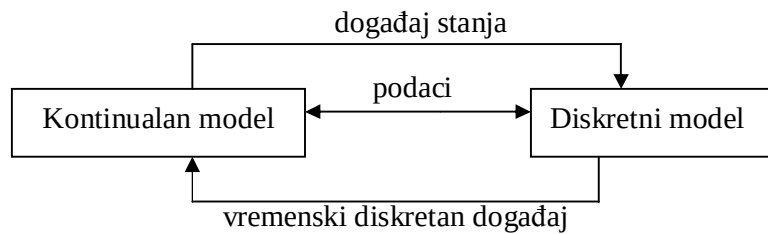
3) **Opisni (verbalni) modeli:** koriste se kada su relacije koje opisuju ponašanje sistema suviše kompleksne ili nedovoljno poznate da bi se prikazale u matematičkom obliku, i ako postoji izvesno znanje o kvalitativnim vezama između tih varijabli. Ovakvi modeli se obično koriste u sociologiji i psihologiji.

Predmet ovog rada biće samo matematički modeli. Matematički modeli mogu se dalje podeliti u sledeće kategorije: 1) statički; 2) dinamički, koji mogu biti kontinualni i diskretni. Moguća klasifikacija (tipovi) matematičkih modela prikazani su na sledećoj slici.



Sika 2.4. Tipovi matematičkih modela

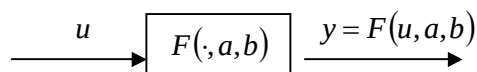
Različiti tipovi matematičkog modela mogu da se kombinuju u okviru složenog modela realnog procesa, kao što je prikazano na sledećoj slici



Sika 2.5. Primer složenog modela u kome postoji interkonekcija između kontinualnog i diskretnog modela

Pojam događaj stanja je vezan za kontinualni model i nastaje kada kontinualna varijabla stanja pređe neki prag (na primer, kada struja i prođe kroz nulu ili kada pozicija x dostigne graničnu vrednost). Detekcija događaja stanja može aktivirati neki prekidač u kontinualnom modelu ili startovati (trigerovati) neki vremenski mehanizam (brojač) u diskretnom modelu. Sa druge strane, vremenski događaj se generiše na osnovu vremenskog procesa unutar modela sa diskretnim događajima, a može aktivirati prekidač ili proizvesti neku drugu aktivnost unutar kontinualnog modela.

2.2.4.1 Statički matematički modeli opisuju nelinearnu algebarsku zavisnost između ulazne i izlazne varijable, kao što je prikazano na slici za skalarne varijable



Slika 2.6. Statički modeli

Ovakav model je linearan ukoliko postoji linearna zavisnost između ulaza u i izlaza y . Međutim, linearnost se može definisati i u odnosu na odgovarajuće parametre (a,b) koji opisuju preslikavanje $F(\cdot)$ između ulaza i izlaza. Mogući tipovi statičkih modela prikazani su sledećom tabelom

$y = F(u)$	$y = F(a,b)$	Primer
linearan	linearan	$y = (a + b)u$
nelinearan	linearan	$y = au + bu^3$
linearan	nelinearan	$y = au / b$
nelinearan	nelinearan	$y = au / (bu + 1)$

Slika 2.7. Tipovi statičkih modela

Kao što je prikazano, ovakav model može biti linearan ili nelinearan i u odnosu na ulaz i u odnosu na parametre, respektivno.

Realni procesi su generalno nelinearni, tako da adekvatan model procesa obično sadrži nelinearnosti [91]. Sa druge strane, matematički aparat je dobro razrađen za linearne modele[80]. Ovaj problem se može praktično razrešiti uvođenjem pojma radne tačke. Naime, nelinearan model se u okolini izabrane radne tačke može linearizovati, čime se dobija linearizovani model koji se dalje može koristiti za analizu i projektovanje primenom teorije linearnih matematičkih sistema. Naravno, tako dobijeni rezultati će važiti samo u okolini razmatrane radne tačke, a svaka radna tačka imaće svoj linearizovani model, tako da se sa promenom radne tačke prelazi na njoj odgovarajući linearizovani model. Dakle, linearizovani model predstavlja model sistema za male signale (predstavljaju odstupanja od radne tačke) i može se koristiti za lokalnu analizu stabilnosti sistema oko date radne tačke.

2.2.4.2 Dinamički matematički modeli: Kao što je prikazano na slici 2.4, postoji više dinamičkih modela sistema koji opisuju ponašanje sistema i u prelaznom režimu, a ne samo u ravnotežnom ili ustaljenom stanju kao što je slučaj sa statičkim modelima. Prelazni režim nastupa odmah nakon dovođenja pobude i traje sve do ulaska sistema u ravnotežno stanje [16,24,48,91,93].

2.2.4.2.1 Kontinualni modeli sa distribuiranim parametrima: Fizički sistemi su kontinualni i opisuju se sa kontinualnim modelima sa distribuiranim ili koncentrisanim parametrima. Model sa distribuiranim parametrima prikazan je u formi paracijalne diferencijalne jednačine. Kod takvih modela i prostorne koordinate (pozicija) x i vreme t predstavljaju nezavisne promenljive u modelu. Primer takvog sistema je transfer toplote kroz zid. Na primer, za $F(x,t)=0$, odgovarajuća parcijalna diferencijalna jednačina je $F_x(x,t)\dot{x} + F_t(x,t)=0$, odakle sledi

$$\frac{\partial F}{\partial t} = a \frac{\partial F}{\partial x} \quad (2.1)$$

Ako se zavisnost od pozicije x zanemari (model sa koncentrisanim parametrima, pošto su sve aktivnosti svedene samo na jednu poziciju), navedena parcijalna diferencijalna jednačina svodi se na običnu diferencijalnu jednačinu.

Ovakva aproksimacija odgovara koncentrisanju različitih materijala u zidu u jedan homogen materijal sa idealnim karakteristikama (koncentrisana masa koja nema debljinu i predstavlja homogenu površinu). Umesto jedne pozicije x , može se definisati i veći broj pozicija x_k , pri čemu se za svako x_k dobija model sa koncentrisanim parametrima (opisan običnom diferencijalnom jednačinom) koji ne zavisi od pozicije. Ovakav postupak selekcije konačnog broja ćelija ili podmodela naziva se metod konačnih elemenata (engleski Finite Element Method ili skraćeno FEM). Svaka ćelija ima svoj sopstveni skup varijabli i običnih diferencijalnih jednačina, tako da dimenzija modela raste linearno sa brojem ćelija. Veći broj ćelija rezultuje u tačnijem modelu, ali je vreme izvršavanja ovakvog modela značajno veće. Međutim, ovakvo uprošćenje je značajno, pošto je numeričko rešavanje parcijalnih diferencijalnih jednačina kompleksno i vremenski zahtevno.

2.2.4.2.2 Kontinualni model sa koncentrisanim parametrima: Ovakav model opisan je sa običnim diferencijalnim jednačinama (ODE). Obična diferencijalna jednačina je formulisana kao eksplicitna diferencijalna jednačina, kod koje je svaki izvod definisan sa desnom stranom jednačine, tj.

$$\frac{dx}{dt} = F(x(t), u(t)); \quad \text{ODE model} \quad (2.2)$$

Dakle, vrednost dx/dt može se za svako t izračunati eksplicitno. Sa druge strane, diferencijalno-algebarska jednačina (DAE) je formulisana kao implicitna diferencijalna jednačina

$$F_1(\dot{x}(t), x(t), u(t)) = 0; \quad \text{DAE model} \quad (2.3)$$

ili kao singularna diferencijalna jednačina

$$E \frac{dx}{dt} = F_2(x(t), u(t)) \quad (2.4)$$

gde je E singularna matrica (inverzija E^{-1} ne postoji) tako da se poslednja relacija ne može napisati u formi ODE. Takođe, prisustvo algebarske jednačine između varijabli za koje su definisani eksplicitno izvodi čini takvu jednačinu implicitnom ili DAE modelom. Na primer,

$$\frac{dx}{dt} = F(x(t), y(t), t) \quad (2.5)$$

$$G(x(t), y(t), t) = 0 \quad (2.6)$$

je DAE, pošto dy/dt nije definisano eksplicitno. Diferenciranjem algebarske jednačine dobija se

$$G_x(x(t), y(t), t) \frac{dx}{dt} + G_y(x(t), y(t), t) \frac{dy}{dt} + G_t(x(t), y(t), t) = 0 \quad (2.7)$$

pa ukoliko $G_y(\cdot)$ nije singularna matrica može se izvod dy/dt izraziti eksplicitno i data jednačina se svodi na ODE, a za odgovarajući DAE se kaže da ima indeks 1. Ako je $G_y(\cdot)$ singularna matrica, DAE jednačine treba da se preurede i odgovarajuća algebarska jednačina ponovo diferencira. Broj primenjenih diferenciranja definiše indeks DAE. Za razliku od ODE (predstavlja DAE sa indeksom 0), koje se relativno lako rešavaju numerički, DAE jednačine unose dodatne numeričke probleme.

2.2.4.2.3 Kontinualni modeli sa događajima stanja: su kontinualni modeli koji mogu generisati neku diskretnu akciju u trenutku kada neka kontinualna varijabla stanja pređe neki definisani prag. Na primer, ako se u sistemu za grejanje kao grejno telo (izvršni organ) koristi gasni ventil koji ima samo dva položaja (otvoreno i zatvoreno). Takva regulacija se naziva on-off regulacija i ako se ventil uključuje kada sobna temperatura padne ispod nekog nivoa, a isključuje kada sobna temperatura pređe neki drugi prag, takva on-off regulacija je u potpunosti određena sa vrednošću sobne temperature koja se meri termostatom. Trenutak uključivanja gasnog ventila (njegovog prelaska iz jednog u drugo stanje) nije sinhronisan ni sa jednim drugim procesom. Drugi primer su energetske konvertori, koji konvertuju veliku količinu energije iz nekog izvora, sa minimalnim gubicima, na električni motor. Mali gubici se postižu korišćenjem prekidačkih kola (elektronskih prekidača), kao što su tiristori i tranzistori sange. Ove poluprovodničke komponente se mogu kontrolisati i računar (mikroprocesor) može izračunati u kojim vremenskim trenucima tiristor treba da se aktivira (upali). Ove tačke u vremenu zavise od vrednosti odgovarajućih napona i struja i tiristor se gasi (zakoči) kada struja ili napon padnu ispod određene vrednosti. Položaj ovih tačaka na vremenskoj osi zavisi od stvarnih vrednosti signala u modelu. Sa druge strane, diferencijalne jednačine sa konstantnim parametrima, koje opisuju kontinualni model sa koncentrisanim parametrima, obično se rešavaju numerički i rešenje se dobija u određenom broju ekvidistantnih diskretnih vremenskih trenutaka, čije se rastojanje naziva integracioni interval. Problem može nastati ukoliko se prekidački trenuci u modelu ne poklapaju sa trenucima u kojima se vrši numerička integracija. Ukoliko je period prekidanja u modelu znatno veći od integracionog intervala, učiniće se male greške ukoliko se trenuci aktiviranja prekidačke logike izjednače sa odgovarajućim trenucima integracije, odnosno ukoliko se trenuci prekidanja podese prema trenucima numeričke integracije. Ukoliko to nije slučaj, neophodno je da se trenuci integracije podese prema trenucima aktiviranja prekidačke logike. Prvi tip simulacije vrši se automatski, ali postojeći simulacioni paketi ne vode računa o drugom tipu sinhronizacije i potrebno je učiniti dodatan napor kako bi se ove sinhronizacije obezbedile.

2.2.4.2.4 Modeli diskretnih događaja: su diskretni modeli koji nemaju fiksni period ponavljanja, tj. imaju aperiodično odabiranje, a sve varijable zadržavaju svoje vrednosti između dva događaja. U trenutku pojave diskretnog događaja varijable mogu menjati svoje vrednosti. Događaj može biti prouzrokovan nekim spoljašnjim događajem, na primer događajem stanja, ili izazvan na osnovu neke vremenske tabele, koja se javlja kao rezultat internih aktivnosti. Primeri ovakvih modela su redovi čekanja (poruke na serveru, vozila na rampi, ljudi na šalterima, itd.). U ovakvim modelima varijable predstavljaju vremenske procese, kao što je na primer trenutak dospeća u red, a stanje varijable označava vremenski trenutak u kome varijabla postaje aktivna, a ne predstavlja fizičku veličinu, što je slučaj kod događaja stanja.

2.2.4.2.5 Modeli sa semplovanim podacima: su najrasprostranjeniji diskretni modeli koji se primenjuju kada se računar koristi kao sredstvo za kontrolu procesa [18,85]. Računar odabira (semplove) podatke upotrebom A/D konvertora, a kontroliše proces na bazi upotrebe D/A konvertora. Semplovanje podataka se vrši sa fiksnom periodom odabiranja (semplovanja). Ovakav tip modela opisan je diferencnom jednačinom, koja se lako rešava upotrebom odgovarajućeg programskog jezika i računara. Tipičan oblik diferencne jednačine koja povezuje izlazni diskretni niz $\{y(i)\}$ i ulazni diskretni niz $\{u(i)\}$ je

$$\sum_{i=0}^n a_i y(i) = \sum_{i=1}^m b_i u(i) \quad (2.8)$$

gde su $a_i, i = 0, \dots, n$; $b_i, i = 1, \dots, m$ koeficijenti ili parametri, koji obično nemaju fizičko tumačenje.

Ovakvi modeli su široko rasprostranjeni u prirodno-tehničkim naukama ali i u društvenim naukama, pre svega ekonometriji. Takođe, sami kontinualni modeli kada se primenjuju na računarskoj mašini, prevode se na odgovarajuću diskretnu formu i rešavaju kao vremenski diskretni modeli, odnosno odgovarajuće diferencne jednačine. Modeli navedenog tipa biće detaljno razmatrani u drugom poglavlju. Ovi modeli predstavljaju i osnovnu strukturu koja se koristi u zadacima parametarske identifikacije i sintezi kvantitativnih modela procesa tipa crne kutije, čemu je u celini i posvećen ovaj rad. Zadatak parametarske identifikacije sastoji se u određivanju nepoznatih vrednosti parametara u (2.8) na optimalan način, u određenom smislu, a na bazi raspoloživih merenja ulazno-izlaznih diskretnih signala. Ovaj problem biće detaljno razmatran u poglavljima 3-6.

3. BLACK-BOX REPREZENTACIJE MODELA SISTEMA

Kao što je navedeno u prethodnom poglavlju, jedan mogući pristup modelovanju sistema zasniva se na poznavanju fizike samog procesa (sistema) koji se modeluje. Naime, polazeći od fizičkih zakona (Newton zakon, Kirchhoff-ovi zakoni, zakoni o održanju energije, mase, momenta itd.) i pretpostake da su komponente sistema idealne dolazi se do modela sistema, kojim je opisano ponašanje sistema. Ovakav pristup zahteva potpuno poznavanje fizičkog procesa i stoga se ovako izvedeni model naziva u anglosaksonskoj literaturi “white-box” model (ili reprezentacija modela u vidu “bele kutije”).

Iako je ovakav pristup prirodan, postoje izvesne situacije gde njegova primena nije moguća:

1. Proces je suviše kompleksan da bi mogao da se opiše idealizovanim fizičkim zakonima ili se ne raspolaže sa adekvatnim znanjem o fizici procesa. U takvoj situaciji može se usvojiti opšta struktura modela u kojoj su nepoznate numeričke vrednosti odgovarajućih parametara (kaže se da je model poznat sa tačnošću do nepoznatih vrednosti parametara u modelu), te se ove vrednosti moraju naknadno proceniti (estimirati) na osnovu raspoloživih merenja ulazno-izlaznih varijabli procesa, uz korišćenje pogodne estimacione procedure (algoritma). Ovakvo dobijeni model naziva se na engleskom jeziku “black-box” model, s obzirom da apriorno znanje o fizici procesa nije uključeno u postupak modeliranja.

2. Moguće je koristiti izvesne idealizovane fizičke zakone da bi se opisalo ponašanje procesa, ali su pri tome numeričke vrednosti odgovarajućih fizičkih parametara i dalje nepoznate (na primer, nisu poznate inercija ili koeficijent trenja u mehaničkom sistemu ili kapacitivnost u električnom sistemu). U ovakvoj situaciji poznat je fizički model procesa ali sa tačnošću do nepoznatih numeričkih vrednosti odgovarajućih fizičkih parametara, koje se moraju proračunati na neki drugi način. Ovakvo dobijeni model predstavlja kombinaciju “white-box” modeliranja sa postupkom estimacije parametara, te se naziva u anglosaksonskoj literaturi “gray-box” model. U opštem slučaju, bilo kakva kombinacija “black-box” modeliranja sa apriornim znanjem o fizici procesa naziva se “gray-box” modeliranje.

U ovom poglavlju biće razmatrano “black-box” modeliranje. Kod ovakvog pristupa ne postoji raspoloživa apriorna informacija o fizici procesa (ne postoje ili nisu poznati idealizovani fizički zakoni koji se mogu uključiti u postupak modeliranja sistema), te se izgradnja modela mora bazirati na raspoloživim merenjima ulaza i izlaza procesa (sistema)[9,23,28-30,43,63,64,88-90].

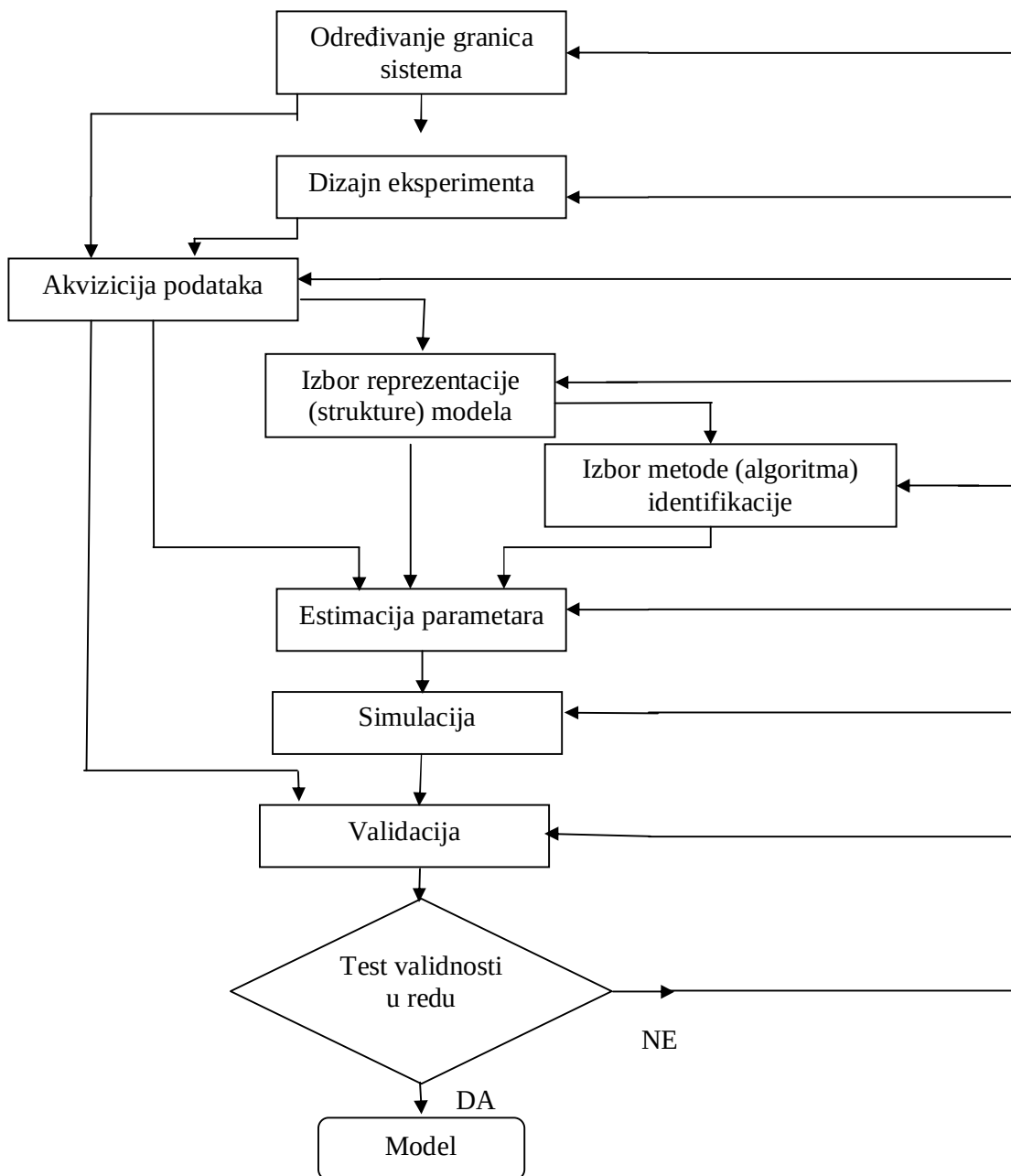
Postupak “black-box” modeliranja sastoji se iz više koraka (aktivnosti), koji su prikazani na sledećoj slici.

Kao što je navedeno u prethodnom poglavlju, prvi korak u postupku “black-box” modeliranja predstavlja određivanje granica kojima je definisan sistem, kao izolovani i idealizovani deo nekog realnog (fizičkog) procesa. Granice sistema određuju koje će se osobine realnog procesa uzeti u obzir prilikom definisanja sistema, kao i koje će osobine realnog procesa biti zanemarene i pridružene okolini sistema.

Drugi korak je planiranje i dizajn eksperimenta, u kome se donosi odluka koji će signali da se mere i sa kojom učestanošću odabiranja, ukoliko se merenja vrše u diskretnim trenucima vremena, kao i u kome se vrši selekcija ulaznih (pobudnih) signala.

Treći korak je akvizicija podataka, koji se sastoji u prikupljanju mernih podataka o ulazu i izlazu sistema, koji su neophodni za proceduru identifikacije (estimacije nepoznatih parametara u usvojenoj strukturi modela).

Četvrti korak je izbor reprezentacije ili struktura modela sistema. Mogu se koristiti različite reprezentacije (strukture) modela sistema, a izbor odgovarajuće strukture zavisi od namene modela. U ovom poglavlju će biti prikazane neke od mogućih struktura “black-box” modela sistema.



Slika3.1: Koraci (aktivnosti) u “black-box” modeliranju

Peti korak je simulacija, koja ima za cilj da se izvrši analiza ponašanja sistema u različitim uslovima rada (predstavlja neki vid validacije modela u odnosu na test podatke).

Šesti korak je validacija, koja ima za cilj da odgovori da li model dovoljno tačno opisuje ponašanje realnog procesa, sa aspekta namene modela.

Najčešće korišćena reprezentacija (struktura) “black-box” modela je tzv. linearan vremenski-invarijantan diskretan model (sistem) [16,24,48,49,85]. Ovakvi modeli su tipični za ekonomiju i računarske nauke, ali se primenjuju i u elektrotehnici, a nastaju kao rezultat diskretizacije (semplovanja) odgovarajućih linearnih vremenski-invarijantnih kontinualnih modela. Pošto se u postupku modeliranja ne koristi apriorno znanje o fizičkim osobinama procesa, kod ovakvih “black-box” modela ne postoji direktna veza između parametara u modelu i fizičkih osobina samog procesa koji se modeluje. Inače, reprezentacija sistema u vidu diskretnog sistema predstavlja parametarski model u vremenskom području. U opštem slučaju reprezentacija modela sistema može biti parametarska (sa konačnim brojem parametara) ili neparametarska (sa beskonačnim brojem parametara), dok sam model može biti prikazan u vremenskom, frekvencijskom i kompleksnom domenu.

3.1. NEPARAMETARSKE STRUKTURE BLACK-BOX MODELA: IMPULSNI ODZIV I FUNKCIJA PRENOSA LINEARNOG DISKRETNOG VREMENSKI INVARIJANTNOG SISTEMA

Pod sistemom se podrazumeva skup relacija između promenljivih i signala, koji predstavlja pogodnu idealizaciju realnog fenomena (realna pojava se još naziva i proces ili postrojenje, a odgovarajući termin na engleskom jeziku je “process” ili “plant”). Prilikom izrade modela procesa neophodno je da se uvedu dodatne pretpostavke o procesu, kao što su granice procesa (odnose se na karakteristike procesa koje su uključene u sistem) i izbor ulazno-izlaznih varijabli i poremećaja.

Sistem je **diskretan** ukoliko su ulazno-izlazne promenljive sistema diskretni signali. Diskretan signal je skup realnih brojeva, tzv. odbiraka, na primer $\{u(k), k = 1, 2, \dots, N\}$, koji mogu nastati semplovanjem vremenski kontinualnog signala $u(t)$, u ekvidistantnim trenucima odabiranja $t_k = kT_s$, gde je k celobrojna promenljiva (indeks diskretnog trenutka), a T_s je perioda odabiranja (semplovanja). Ovakav signal je deterministički ukoliko je vrednost odbirka u potpunosti poznata, na osnovu nekog matematičkog izraza ili tabelarnog prikaza. Dakle, buduće vrednosti determinističkog signala mogu se predvideti na osnovu poznavanja njegovih karakteristika i prethodnih vrednosti. Sa druge strane, egzaktna vrednost stohastičkog signala se ne može predvideti, već se ta vrednost može bliže opisati na osnovu statističkih pokazatelja, kao što su na primer srednja (očekivana) vrednost i varijansa. Pored navedenog, signali mogu biti merljivi i nemerljivi, a diskretan signal se meri samo u ekvidistantnim diskretnim trenucima vremena $t_k = kT_s$, gde je T_s perioda semplovanja (diskretizacije, odabiranja ili uzorkovanja). Na taj način, kontinualni

signal $u(t)$, gde je nezavisno promenljiva t realna promenljiva ($t \in R$) koja predstavlja vreme, a realna matematička funkcija $u(t)$ predstavlja signal, reprezentovan je skupom (sekvencom) odbiraka $\{u(kT_s), k = 0, \pm 1, \pm 2, \dots\}$, odnosno $k \in Z$ gde je Z skup celih brojeva, koji se naziva diskretnim signalom. Uobičajeno je da se zbog jednostavnosti pisanja usvoji $T_s = 1$, pa tada diskretan niz $\{u(k), k \in Z\}$ označava diskretan signal. Diskretan sistem transformiše (preslikava) ulazni diskretan niz brojeva u izlazni diskretan niz brojeva. Osnovne osobine ovakvih sistema su linearnost, vremenska-invarijantnost (stacionarnost), kauzalnost i stabilnost.

Sistem je **linearan** ukoliko linearna kombinacija diskretnih ulaznih signala generiše odziv (izlaz) koji predstavlja linearnu kombinaciju izlaza koji se dobijaju za pojedinačne ulaze. Dakle, ako diskretan ulaz $\{u_1(k)\}$ generiše na izlazu sistema diskretan odziv $\{y_1(k)\}$, a diskretan ulaz $\{u_2(k)\}$ ima za posledicu izlaz sistema $\{y_2(k)\}$, tada ulaz $u(k) = c_1 u_1(k) + c_2 u_2(k)$, gde su c_1 i c_2 realne konstante, generiše kao izlazni signal diskretnog sistema $y(k) = c_1 y_1(k) + c_2 y_2(k)$.

Sistem je **vremenski-invarijantan** ukoliko se njegove osobine ne menjaju sa vremenom. Dakle, ako diskretan ulaz $\{u_1(k)\}$ generiše kao izlaz sistema diskretan signala $\{y_1(k)\}$, tada će pomereni ulaz $\{u(k)\} = \{u_1(k - k_0)\}$, $\forall k, k_0 \in Z$, dati na izlazu sistema odziv $\{y(k)\} = \{y_1(k - k_0)\}$. Ova osobina naziva se i invarijantnost u odnosu na vremenski šift (pomeraj).

Važna osobina sistema, pored linearnosti i vremenske-invarijantnosti, je i kauzalnost. Sistem je **kauzalan** (engleski “causal” ili “proper”) ukoliko izlaz u nekom vremenskom trenutku zavisi samo od ulaza do tog trenutka vremena, uključujući i sam posmatrani vremenski trenutak, a ne i od vrednosti ulaza u budućim trenucima vremena. Ukoliko se posmatrani vremenski trenutak isključi iz ove formulacije, kaže se da je sistem striktno kauzalan (engleski “strictly causal”).

Takođe, sistemi se mogu klasifikovati kao monovarijabilni, odnosno sistemi sa jednim ulazom i jednim izlazom (označavaju se kao SISO sistemi, a skraćenica potiče od engleskog termina “single input – single output”) i multivarijabilni, odnosno sistemi sa više ulaza i izlaza (označavaju se kao MIMO, od engleskog termina “multi-input multi-output”).

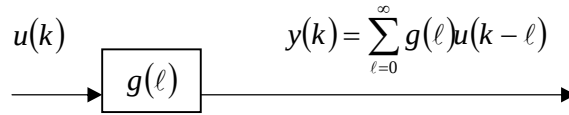
Najpoznatije reprezentacije “black box” modela predstavljaju impulsni odziv i funkcija prenosa linearnog diskretnog vremenski invarijantnog sistema.

Impulsni odziv predstavlja neparametarsku reprezentaciju modela u vremenskom domenu. Impulsni odziv linearnog vremenski-invarijantnog diskretnog sistema (eng. “Linear Time-Invariant Discrete-Time System” ili skraćeno LTI vremenski diskretan sistem) je beskonačna sekvenca realnih brojeva $g(k), k \in Z$, pri čemu je $g(k) = 0$ za $k < 0$ za kauzalne sisteme, koja se dobija kao odziv (izlaz)

sistema na jedinični diskretan impuls $\delta(k)$, koji je definisan sa $\delta(k)=1$ za $k=0$ i $\delta(k)=0$ za $k \neq 0, k \in \mathbb{Z}$. Tada je izlaz (odziv) $y(k)$ ovakvog modela sistema na ulazni kauzalan diskretan signal $u(k), k \in \mathbb{Z}$ i $k > 0 (u(k)=0$ za $k < 0$ na osnovu osobine kauzalnosti) definisan diskretnom konvolucijom između ulaza i impulsnog odziva, tj.

$$y(k) = g(k) * u(k) = \sum_{\ell=0}^{\infty} g(\ell) u(k-\ell) = \sum_{\ell=0}^k g(\ell) u(k-\ell) \quad (3.1)$$

Uočimo da je u konvolucionoj sumi donja granica $\ell = 0$ (osobina kauzalnosti) kao i da izlaz $y(k)$ zavisi samo od ulaza do trenutka k , uključujući i sam trenutak k ($u(k-\ell)=0$ za $k-\ell < 0$, tj. $\ell > k$, pošto je kauzalan signal), tako da je sistem kauzalan. Za striktno kauzalan sistem donja granica u konvolucionoj sumi je $\ell = 1$, tako da će izlaz u trenutku k , $y(k)$, zavisiti samo od ulazne sekvence $\{u(k-1), u(k-2), \dots, u(0)\}$, a ne i od samog $u(k)$. Drugim rečima, kod striktno kauzalnog sistema je $g(0)=0$. Šematski prikaz “black-box” reprezentacije modela u vidu impulsnog odziva dat je na sledećoj slici



Slika 3.1 Karakterizacija linearnog, diskretnog, vremenski-invarijantnog, kauzalnog modela sistema na osnovu impulsnog odziva.

Primetimo da se navedena konvoluciona suma koja definiše odziv LTI vremenski diskretnog sistema može dokazati relativno jednostavno na sledeći način. Ako sa $L(\cdot)$ označimo linearnu transformaciju koju sistem vrši nad ulaznim signalom da bi generisao odziv (izlazni) signal, tada je odziv $y(k)$ na pobudu $u(k)$ definisan sa

$$y(k) = L(u(k)), \quad \forall k \in \mathbb{Z} \quad (3.2)$$

Posebno, za impulsnu pobudu $u(k) = \delta(k)$ dobija se impulsni odziv

$$g(k) = L(\delta(k)), \quad \forall k \in \mathbb{Z} \quad (3.3)$$

Sa druge strane, proizvoljna diskretna i kauzalna pobuda $u(k)$ se predstavlja u vidu konvolucije jedinične impulsne sekvence $\delta(k)$ i same pobudne sekvence $u(k)$, tj.

$$u(k) = \sum_{i=0}^{\infty} u(i) \delta(k-i) = \sum_{i=0}^{\infty} u(k-i) \delta(i) \quad (3.4)$$

pošto je $\delta(k-i)=1$ za $k-i=0$ i $\delta(k-i)=0$ za $k-i \neq 0$. Tada je odziv sistema na zadatu pobudu

$$y(k) = L(u(k)) = L\left(\sum_{i=0}^{\infty} u(k) \delta(i)\right) = \sum_{i=0}^{\infty} u(k-i)g(i) \quad (3.5)$$

što predstavlja traženu relaciju. Prilikom dobijanja poslednje relacije odbirci pobude $u(k-i)$ tretirani su kao poznate konstante koje se mogu izvući van linearnog operatora sistema, dok je sama linearna operacija nad impulsnom pobudom rezultovala u impulsni odziv sistema.

Ako se definiše operator prednjačenja q (engleski “forward shift” operator)

$$qu(k) = u(k+1) \quad (3.6)$$

i operator kašnjenja q^{-1} (engleski. “backward shift” operator)

$$q^{-1}u(k) = u(k-1) \quad (3.7)$$

konvoluciona suma se može napisati kao

$$y(k) = \sum_{\ell=0}^{\infty} g(\ell) [q^{-\ell}u(k)] = \left[\sum_{\ell=0}^{\infty} g(\ell) q^{-\ell} \right] u(k) \quad (3.8)$$

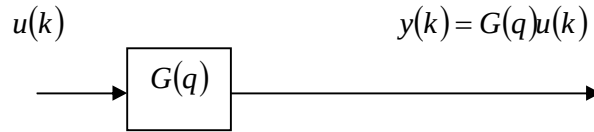
Definišući dalje funkciju prenosa sa

$$G(q) = \sum_{\ell=0}^{\infty} g(\ell) q^{-\ell} \quad (3.9)$$

konvoluciona suma se može zapisati u polinomijalnoj formi

$$y(k) = G(q)u(k) \quad (3.10)$$

Poslednja relacija se može šematski prikazati na sledeći način



Slika3.3: Karakterizacija linearnog, diskretnog, vremenski-invarijantnog kauzalnog modela sistema na osnovu funkcije prenosa

Primetimo da se u teoriji sistema funkcija prenosa definiše kao z -transformacija impulsnog odziva $g(\ell)$, $\ell = 0, 1, 2, \dots$ [16, 24, 85].

$$G(z) = \sum_{\ell=0}^{\infty} g(\ell) z^{-\ell} \quad (3.11)$$

gde je z kompleksna promenljiva. Kada se radi u vremenskom domenu sa diferencnim jednačinama (diskretna konvolucija predstavlja vid diferencne jednačine)

uobičajeno je da koristi operator pomeraja u vremenu q umesto komplekse varijable z . Iako ovakva notacija može da izazove malu zabunu, ona je pogodnija za korišćenje u vremenskom domenu i znatno olakšava analizu.

Pored već navedenih osobina (linearnost, vremenska-invarijantnost i kauzalnost) važno je uvesti i pojam **stabilnosti** [16,24,85,93]. Iako postoji više tipova stabilnosti, u ovom materijalu ćemo koristiti samo pojam stabilnosti **ograničen-ulaz ograničen-izlaz** (engleski “bounded-input bounded-output” ili skraćeno BIBO), koja označava da će ograničen ulazni signal generisati na izlazu sistema takođe ograničen signal. Prilikom analize stabilnosti sistema uobičajeno je da se koristi definicija funkcije prenosa u kompleksnom z domenu i da se uvede pojam nula i polova funkcije prenosa $G(z)$. Nule sistema su vrednosti β_i kompleksne varijable z za koje je $G(\beta_i) = 0$, dok su polovi sistema vrednosti α_i kompleksne varijable z za koje je $G(\alpha_i) = \infty$ (nule i polovi nazivaju se zajedničkim imenom kritične učestanosti). Posebno, ukoliko je $G(z)$ realna racionalna funkcija, tj. količnik dva polinoma kompleksne varijable z sa fiksnim realnim parametrima, na primer $G(z) = B(z)/A(z)$, tada su koreni (rešenja) jednačine $B(z) = 0$ nule sistema, dok su koreni (rešenja) jednačine $A(z) = 0$ polovi sistema (nule polinoma u brojiocu funkcije $G(z)$ su nule sistema, dok su nule polinomi u imeniocu funkcije $G(z)$ polovi sistema).

Kada se analizira BIBO stabilnost sledeći iskazi su ekvivalentni:[16,24,48]

1. Sistem je BIBO stabilan.
2. $\sum_{\ell=0}^{\infty} |g(\ell)| < \infty$
3. $\lim_{l \rightarrow \infty} g(l) = 0$ i $|g(l)| < \infty$
4. $G(z)$ ima polove α_i koji svi leže unutar oblasti ograničene jediničnim krugom $|z| = 1$, odnosno $|\alpha_i| < 1$.

U engleskoj literaturi se koristi termin “monic” da označi funkciju prenosa $G(q)$ čiji je slobodan član $g(0) = 1$, odnosno monic-funkcija prenosa

$$G(q) = \sum_{\ell=0}^{\infty} g(\ell)q^{-\ell}; \quad g(0) = 1 \quad (3.12)$$

Navedeni iskazi lako se dokazuju nalaženjem apsolutne vrednosti leve i desne strane relacije koja definiše odziv sistema kao konvolucionu sumu pobude i impulsnog odziva, tj.

$$|y(k)| = \left| \sum_{\ell=0}^{\infty} g(\ell)u(k-\ell) \right| \leq \sum_{\ell=0}^{\infty} |g(\ell)| |u(k-\ell)| \quad (3.13)$$

pa ukoliko je pobuda ograničena to je $|u(k-\ell)| \leq C_u < \infty$ gde je C_u nenegativna konačna konstanta, dalje sledi

$$|y(k)| \leq C_u \sum_{\ell=0}^{\infty} |g(\ell)| \quad (3.14)$$

pa da bi izlaz (odziv) bio konačan, odnosno

$$|y(k)| \leq C_y < \infty \quad (3.15)$$

mora biti ispunjeno

$$\sum_{\ell=0}^{\infty} |g(\ell)| < \infty \quad (3.16)$$

što predstavlja iskaz broj 2. Da bi impulsni odziv predstavljao apsolutno konvergentan niz dalje mora da važi $g(\ell) \rightarrow 0$ kada $\ell \rightarrow \infty$ i $|g(\ell)| < \infty$ za $\forall \ell \in \mathbb{Z}$, što odgovara iskazu broj 3. Sa druge strane, funkcija prenosa sistema u z -kompleksnom domenu data je sa

$$G(z) = \sum_{\ell=0}^{\infty} g(\ell) z^{-\ell} \quad (3.17)$$

i ovaj izraz će konvergirati u prstenu $R_d |z| \leq \infty$, gde je R_d poluprečnik kruga sa centrom u koordinatnom početku i koji prolazi kroz najudaljeniji pol funkcije prenosa od koordinatnog početka. Pošto je za $z = 1$, $G(\ell) = \sum_{\ell=0}^{\infty} g(\ell)$, a na osnovu uslova BIBO stabilnosti $G(\ell) < \infty$, to se najudaljeniji pol mora nalaziti unutar jediničnog kruga $|z| = 1$ da bi sistem bio BIBO stabilan.

U zaključku se može navesti da impulsni odziv i funkcija prenosa predstavljaju dve neparametarske reprezentacije (strukture) modela sistema u vremenskom području. Ovakva reprezentacija ima beskonačno parametara, pošto impulsni odziv $\{g(\ell)\}$ ima beskonačno odbiraka, odnosno njegovo trajanje je beskonačno, pa se ovakav odziv na engleskom jeziku naziva “Infinite Impulse Response” i označava sa IIR [16,24,48,65,85]. Obe ove reprezentacije definišu relaciju između ulaznog i izlaznog signala u vremenskom domenu. Ukoliko je ulazni signal deterministički i izlazni signal biće deterministički. Međutim, navedena reprezentacija nije mnogo realna i zahteva dodatnu nadogradnju na osnovu uvođenja signala poremećaja. Naime, kod prethodne reprezentacije ulaz je deterministički signal, koji će na izlazu sistema generisati takođe deterministički signal. Međutim, na realne sisteme po pravilu deluju i poremećaji, koji su obično stohastičke prirode i nemerljivi, tako da su van kontrole dizajnera, iako deluju na kompletan odziv (izlaz) sistema.

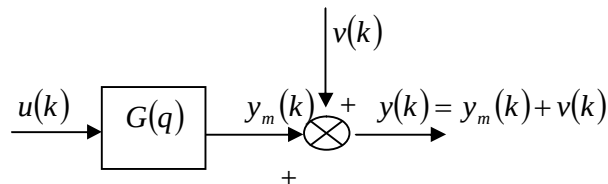
3.2. SLUČAJNI POBUDNI SIGNALI: POREMEĆAJI U SISTEMU

Signali poremećaja su po pravilu prisutni u realnom sistemu, pri čemu tačka u kojoj deluju na sistem obično nije poznata, a karakteristike im se takođe menjaju sa vremenom i uslovima rada. Uz navedeno, ovakvi signali su obično i nemerljivi, te se tretiraju kao stohastički signali [41,49,69,71,81], koji se mogu svrstati u dve klase:

1. **Šum merenja:** potiče od senzora (mernog uređaja) kojim se meri neki signal i koji nije idealan. Naime, svaki elektronski uređaj karakterišu temperaturno zavisni šum i drift, tako da će se u postupku akvizicije podataka uvek pojaviti slučajne greške, koje su posledica ovih fenomena.

2. **Neupravljivi (nekontrolabilni) ulazi:** na realan sistem pored kontrolnih (kontrolabilnih ili upravljivih) ulaza deluju i ulazi koji nisu kontrolabilni i predstavljaju poremećaje procesa. Na pimer, pri navigaciji brodom kontrolabilni ulaz je zadati pravac krme, ali će na pravac broda uticati i vodene struje, koje su generalno nepredvidive i predstavljaju slučajne poremećaje.

Kada se poremećaji uključe u reprezentaciju modela sistema, uobičajeno je da se razmatra njihovo ekvivalentno dejstvo na sam izlaz. Tako se dejstvo svih poremećaja koji deluju na sistem svodi na jedan ekvivalentan stohastički signal $v(k)$, koji se superponira na deterministički izlaz $y_m(k)$ koji je generisan kontrolnim ulaznim signalom $u(k)$ i takav poremećaj $v(k)$ naziva se aditivni šum izlaza. Blok šema sistema sa aditivnim šumom izlaza predstavljena je na sledećoj slici



Slika3.4: Sistem sa poremećajem svedenim na izlaz

Ovakav model je u mnogim situacijama realan, s obzirom da se i poremećaji na ulazu mogu svesti na izlaz, pošto će takvi poremećaji propagirati kroz sistem i na izlazu izazvati odgovarajući odziv. Za opis poremećaja se koriste statistički i stohastički pokazatelji.

3.2.1 Statističke i stohastičke karakteristične veličine za karakterizaciju poremećaja

Poremećaji su slučajni (stohastički) signali, tako da se njihove vrednosti ne mogu predvideti, odnosno vrednosti signala su slučajne i ne mogu se egzaktano odrediti čak i kada su poznate prethodne vrednosti signala. Stoga se ovakvi signali opisuju na osnovu odgovarajućih statističkih i stohastičkih pokazatelja (karakterističnih veličina) [41,49,69,71,81].

Statističke karakteristike (pokazatelj) signala vezani su za distribuciju (raspodelu) amplitude signala, kao slučajne veličine, odnosno vode računa o tome koliko se često neka vrednost amplitude pojavljuje u sekvenci (vremenskoj seriji), ne vodeći pri tome računa u kome redosledu (vremenu) se ove vrednosti signala pojavljuju. Dakle, statističke karakteristike vezane su za prostornu raspodelu amplitude signala (prostornu korelaciju), ne vodeći računa o korelaciji (zavisnosti) amplituda u vremenu.

Stohastičke karakteristike (pokazatelji) signala vezane su za dinamiku signala, odnosno promenu slučajne amplitude u vremenu i daju odgovor na pitanje da li postoji korelacija (zavisnost) između trenutne amplitude signala i amplitude signala u nekom drugom trenutku. Dakle, stohastičke karakteristike (pokazatelji) vezani su za vremensku korelaciju, odnosno opisuju dinamiku signala u vremenu ili raspodelu frekvencija u frekvencijskom liku (spektru) signala.

Kao što je poznato iz teorije slučajnih procesa, stohastički signal (slučajni proces) sastoji se iz ansambla (familije) funkcija vremena, koje se nazivaju realizacije. Međutim, u praksi se obično raspolaze sa jednom realizacijom slučajnog signala, dobijenom na osnovu merenja (na primer, ako je signal izmeren, šum merenja predstavlja samo jednu realizaciju slučajnog signala). Stoga je važno da se u praksi koriste slučajni procesi koji su u potpunosti opisani sa samo jednom svojom realizacijom, koja se naziva i vremenska serija (niz). Sa praktičnog aspekta je takođe važno da karakteristike procesa ne zavise od samog vremenskog trenutka u kome je izvršeno merenje amplitude signala. Stohastički procesi kod kojih osobine ne zavise od apsolutne vrednosti vremena nazivaju se široko ili slabo stacionarni procesi (engleski “weakly stationary”), dok stohastički procesi čije se osobine mogu sagledati na osnovu samo jedne realizacije nazivaju se široko ili slabo ergodični (engleski “weakly ergodic”) [41,49,69,71]. Prilikom modeliranja sistema obično se pretpostavlja da su poremećaji široko stacionarni ili široko ergodični slučajni procesi, pri čemu široko ergodični procesi jesu istovremeno i širokostacionarni, dok obrnuto ne važi.

3.2.1.1 Osnovne statističke karakteristične veličine i algoritmi za njihovu procenu

Dva najvažnija statistička pokazatelja su srednja vrednost (očekivana vrednost ili matematičko očekivanje) i varijansa.

Definicija srednje vrednosti: Srednja vrednost (engleski “mean value”) ili necentrirani statistički momenat prvog reda diskretnog slučajnog signala $\{v(k)\}$ definisana je sa

$$m_v = \bar{v} = E\{v(k)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N v(k) \quad (3.18)$$

gde je $E\{\cdot\}$ operator matematičkog očekivanja, odnosno

$$E\{x\} = \int_{-\infty}^{\infty} x f(x) dx \quad (3.19)$$

U navedenom izrazu $f(x)$ predstavlja funkciju gustine verovatnoće (f.g.v.) slučaja varijable $x = v(k)$. Primetimo da kod širokostacionarnog slučajnog niza funkcije gustine verovatnoće $f(x, k) = f(x)$ za $\forall k \in Z$, odnosno funkcija gustine verovatnoće ne zavisi od vremena, tako da srednja vrednost takođe ne zavisi od vremena, tj. predstavlja konstantu. Matematičko očekivanje $E\{\cdot\}$ je linearan operator, što znači da za bilo koje dve slučajne promenljive $v(k)$ i $\omega(k)$ (predstavljaju odбирke diskretnih slučajnih signala) i proizvoljne realne konstante c_1 i c_2 važi

$$E\{c_1 v(k) + c_2 \omega(k)\} = c_1 E\{v(k)\} + c_2 E\{\omega(k)\} = c_1 \bar{v} + c_2 \bar{\omega} \quad (3.20)$$

Definicija varijanse: Varijansa ili centrirani statistički momenat drugog reda diskretnog slučajnog procesa $\{v(k)\}$ definisana je sa

$$\sigma_v^2 = E\{[v(k) - \bar{v}]^2\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N [v(k) - \bar{v}]^2 \quad (3.21)$$

gde je

$$E\{[v(k) - \bar{v}]^2\} = \int_{-\infty}^{\infty} [v - \bar{v}]^2 f(v) dv; \quad f(v, k) = f(v) \quad \text{za} \quad \forall k \in Z \quad (3.22)$$

U praksi se odgovarajuća matematička očekivanja, odnosno statistički pokazatelji moraju izračunati na osnovu uzorka konačne dužine N (vremenske serije). Procena srednje vrednosti na osnovu uzorka (jedne realizacije) signala $\{v(k)\}$ konačne dužine N data je sa aritmetičkom sredinom

$$\bar{v}(N) = \hat{m}_v(N) = \frac{1}{N} \sum_{k=1}^N v(k) \quad (3.23)$$

Ako se na dati izraz primeni operator matematičkog očekivanja, koji je linearan operator, dobija se

$$E\{\hat{m}_v(k)\} = E\left\{\frac{1}{N} \sum_{k=1}^N v(k)\right\} = \frac{1}{N} \sum_{k=1}^N E\{v(k)\} = \frac{1}{N} \sum_{k=1}^N m_v = m_v \quad (3.24)$$

Dobijeni rezultat pokazuje da je srednja vrednost procene, kao slučajne veličine, jednaka samoj determinističkoj veličini koja se procenjuje, te se takva procena naziva nepomerenom (engleski unbiased estimate). Sa druge strane, varijansa ove procene definisana je izrazom

$$\begin{aligned} \hat{\sigma}_{m_v}^2(N) &= E\{[\hat{m}_v(N) - m_v]^2\} \\ &= E\left\{\left[\frac{1}{N} \sum_{k=1}^N v(k) - m_v\right]^2\right\} = E\left\{\left[\frac{1}{N} \sum_{k=1}^N [v(k) - m_v]\right]^2\right\} \end{aligned}$$

$$\begin{aligned}
 &= E \left\{ \frac{1}{N^2} \sum_{k=1}^N [v(k) - m_v]^2 \right\} + E \left\{ \frac{2}{N^2} \sum_{k=1}^N \sum_{\substack{j=1 \\ j \neq k}}^N [v(k) - m_v][v(j) - m_v] \right\} \\
 &= \frac{1}{N^2} \sum_{k=1}^N E \{ [v(k) - m_v]^2 \} + \frac{2}{N^2} \sum_{k=1}^N \sum_{\substack{j=1 \\ j \neq k}}^N E \{ [v(k) - m_v][v(j) - m_v] \} \quad (3.25)
 \end{aligned}$$

Pošto je na osnovu definicionog izraza za varijansu varijable $v(k)$

$$\sigma_v^2 = E \{ [v(k) - m_v]^2 \} \quad (3.26)$$

i ako se pretpostavi da su centrirane slučajne varijable $v(k) - m_v$ i $v(j) - m_v$ međusobno nekorelisane, tada je na osnovu definicije nekorelisanosti

$$E \{ [v(k) - m_v][v(j) - m_v] \} = 0, \quad k \neq j \quad (3.27)$$

odakle sledi

$$\hat{\sigma}_{m_v}^2(N) = \frac{1}{N^2} \sum_{k=1}^N \sigma_v^2 = \frac{\sigma_v^2}{N} \quad (3.28)$$

Dakle, kada dužina uzorka N neograničeno raste ($N \rightarrow \infty$) varijanse nepomerene procene $\hat{m}_v(N)$ težiće nuli, tj.

$$\lim_{N \rightarrow \infty} \hat{\sigma}_{m_v}^2(N) = \lim_{N \rightarrow \infty} \frac{\sigma_v^2}{N} = 0 \quad (3.29)$$

i takva nepomerena procena je konzistentna.

Na sličan način procena varijanse na uzorku konačne dužine

$$\hat{\sigma}_v^2(N) = \frac{1}{N} \sum_{k=1}^N [v(k) - \hat{m}_v(N)]^2 = \frac{1}{N} \sum_{k=1}^N \left[v(k) - \frac{1}{N} \sum_{i=1}^N v(i) \right]^2 \quad (3.30)$$

je pomena, što se može dokazati ako se na poslednji izraz primeni linearni operator matematičkog očekivanja $E\{\cdot\}$, tj.

$$\begin{aligned}
 E \{ \hat{\sigma}_v^2(N) \} &= E \left\{ \frac{1}{N} \sum_{k=1}^N \left[v(k) - \frac{1}{N} \sum_{i=1}^N v(i) \right]^2 \right\} \\
 &= \frac{1}{N} \sum_{k=1}^N E \left\{ \left[v(k) - \frac{1}{N} \sum_{i=1}^N v(i) \right]^2 \right\} = \frac{1}{N^3} \sum_{k=1}^N E \left\{ \left[Nv(k) - \sum_{i=1}^N v(i) \right]^2 \right\}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{N^3} \sum_{k=1}^N E \left\{ \left[(N-1)v(k) - \sum_{\substack{i=1 \\ i \neq k}}^N v(i) \right]^2 \right\} \\
 &= \frac{1}{N^3} \sum_{k=1}^N E \left\{ \left[(N-1)v(k) - (N-1)m_v - \sum_{\substack{i=1 \\ i \neq k}}^N v(i) + (N-1)m_v \right]^2 \right\} \\
 &= \frac{1}{N^3} \sum_{k=1}^N E \left\{ \left[(N-1)(v(k) - m_v) - \sum_{\substack{i=1 \\ i \neq k}}^N (v(i) - m_v) \right]^2 \right\} \\
 &= \frac{1}{N^3} \sum_{k=1}^N E \left\{ \left[(N-1)^2 (v(k) - m_v)^2 \right] + \frac{1}{N^3} \sum_{k=1}^N E \left\{ \left[\sum_{\substack{i=1 \\ i \neq k}}^N (v(i) - m_v) \right]^2 \right\} \right\} \\
 &\quad - \frac{2}{N^3} E \left\{ (N-1)(v(k) - m_v) \sum_{\substack{i=1 \\ i \neq k}}^N (v(i) - m_v) \right\} \\
 &= \frac{1}{N^3} \sum_{k=1}^N (N-1)^2 E \left\{ [v(k) - m_v]^2 \right\} + \frac{1}{N^3} \sum_{k=1}^N E \left\{ \sum_{\substack{i=1 \\ i \neq k}}^N (v(i) - m_v)^2 \right\} \\
 &\quad + \frac{2}{N^3} \sum_{k=1}^N E \left\{ \sum_{\substack{i=1 \\ i \neq k}}^N \sum_{\substack{j=1 \\ j \neq i}}^N (v(i) - m_v)(v(j) - m_v) - \frac{2(N-1)}{N^3} E \left\{ \sum_{\substack{i=1 \\ i \neq k}}^N (v(k) - m_v)(v(i) - m_v) \right\} \right\} \quad (3.31)
 \end{aligned}$$

Pošto je na osnovu pretpostavke o nekorelisanosti slučajnog procesa $\{v(k)\}$

$$E\{(v(i) - m_v)(v(j) - m_v)\} = 0 \quad \text{za} \quad \forall i \neq j \in Z \quad (3.32)$$

a na osnovu definicionog izraza za varijansu ovog slučajnog procesa

$$\sigma_v^2 = E\{(v(k) - m_v)^2\} \quad (3.33)$$

dalje sledi

$$E\{\hat{\sigma}_v^2(N)\} = \frac{1}{N^3} \sum_{k=1}^N (N-1)^2 \sigma_v^2 + \frac{1}{N^3} \sum_{k=1}^N \sum_{\substack{i=1 \\ i \neq k}}^N E\{(v(i) - m_v)^2\}$$

$$= \frac{N(N-1)^2}{N^3} \sigma_v^2 + \frac{N(N-1)}{N^3} \sigma_v^2 = \frac{1}{N^2} \left[(N-1)^2 + (N-1) \right] \sigma_v^2 \quad (3.34)$$

odnosno

$$E\{\hat{\sigma}_v^2(N)\} = \frac{N-1}{N} \sigma_v^2 \quad (3.35)$$

Dakle, u opštem slučaju je $E\{\hat{\sigma}_v^2(N)\} \neq \sigma_v^2$, te je ovakva procena pomenjena (procena je nepomerena jedino za $N \rightarrow \infty$, tj. u asimptotskom slučaju). Dakle, da bi se dobila nepomerena procena varijanse, usvojenu procenu je potrebno pomnožiti sa faktorom $N/(N-1)$, odnosno nepomerena procena varijanse na uzorku konačne dužine N ima oblik

$$\hat{\sigma}_v^2(N) = \frac{1}{N-1} \sum_{k=1}^N \left[v(k) - \frac{1}{N} \sum_{i=1}^N v(i) \right]^2 \quad (3.36)$$

Ovakva procena zadovoljava uslov $E\{\hat{\sigma}_v^2(N)\} = \sigma_v^2$, a može se pokazati da je ova procena i konzistentna, odnosno da varijansa procene teži nuli kada dužina uzorka N neograničeno raste ($N \rightarrow \infty$). Zaista, varijansa procene definisana je izrazom

$$\text{var } \hat{\sigma}_v^2(N) = E\left\{ \left[\hat{\sigma}_v^2(N) - \sigma_v^2 \right]^2 \right\} \quad (3.37)$$

gde je σ_v^2 tačna vrednost varijanse, koja se estimira, pa ako se pretpostavi radi jednostavnosti pisanja da je srednja vrednost slučajnog procesa $v(k)$ jednaka nuli, tj. $m_v = 0$, tada je sama nepomerena procena varijanse

$$\hat{\sigma}_v^2(N) = \frac{1}{N-1} \sum_{k=1}^N v^2(k) \quad (3.38)$$

pa varijansa procene postaje

$$\begin{aligned} \text{var } \hat{\sigma}_v^2(N) &= E\left\{ \left(\hat{\sigma}_v^2(N) - \sigma_v^2 \right)^2 \right\} = E\left\{ \left[\frac{1}{N-1} \sum_{k=1}^N v^2(k) - \sigma_v^2 \right]^2 \right\} \\ &= \frac{1}{(N-1)^2} E\left\{ \left[\sum_{k=1}^N v^2(k) \right]^2 \right\} - \frac{2\sigma_v^2}{N-1} E\left\{ \sum_{k=1}^N v^2(k) \right\} + \sigma_v^4 \\ &= \frac{1}{(N-1)^2} \sum_{k=1}^N \sum_{j=1}^N E\{v^2(k)v^2(j)\} - \frac{2\sigma_v^2}{N-1} \sum_{k=1}^N E\{v^2(k)\} + \sigma_v^4 \end{aligned}$$

$$= \frac{1}{(N-1)^2} \sum_{k=1}^N \sum_{\substack{j=1 \\ k \neq j}}^N E\{v^2(k)v^2(j)\} + \frac{1}{(N-1)^2} \sum_{k=1}^N E\{v^4(k)\} - \frac{2\sigma_v^2}{N-1} \sum_{k=1}^N E\{v^2(k)\} + \sigma_v^4 \quad (3.39)$$

Imajući u vidu dobro poznat rezultat iz teorije gausovskih slučajnih varijabli koja tvrdi da je za nekorelisane slučajne varijable $v(i)$ i $v(j)$ ispunjeno [49]

$$E\{v^2(i)v^2(j)\} = \begin{cases} E\{v^2(i)\}E\{v^2(j)\} = \sigma_v^4, & i \neq j \\ E\{v^4(i)\} = 3\sigma_v^4, & i = j \end{cases} \quad (3.40)$$

dalje se dobija (kod gausovski raspodeljene slučajne varijable parni momenat reda $2n$ može se izračunati na osnovu varijanse σ^2 po formuli $1 \cdot 3 \cdot \dots \cdot (2n-1)\sigma^2$, a u našem slučaju $n = 2$)

$$\begin{aligned} \hat{\sigma}_v^2(N) &= \frac{1}{(N-1)^2} N(N-1)\sigma_v^4 + \frac{1}{(N-1)^2} N3\sigma_v^4 - \frac{2\sigma_v^2}{N-1} N\sigma_v^2 + \sigma_v^4 \\ &= \frac{1}{(N-1)^2} [N(N-1)\sigma_v^4 + 3N\sigma_v^4] - \frac{2N}{N-1} \sigma_v^4 + \sigma_v^4 \\ &= \frac{1}{(N-1)^2} [N^2\sigma_v^4 + 2N\sigma_v^4] - \frac{2N\sigma_v^2 - (N-1)\sigma_v^4}{N-1} \\ &= \frac{(N+2)N\sigma_v^4}{(N-1)^2} - \frac{\sigma_v^4(N+1)}{(N-1)} = \frac{(N+2)N\sigma_v^4 - \sigma_v^4(N^2-1)}{(N-1)^2} \\ &= \frac{\sigma_v^4(2N+1)}{(N-1)^2} \xrightarrow{N \rightarrow \infty} 0 \end{aligned} \quad (3.41)$$

Dakle, pošto varijansa nepomerene procene teži nuli sa porastom dužine mernog uzorka, ovakava nepomerena procena je i konzistentna.

3.2.2 Osnovne stohastičke karakteristične veličine i algoritmi za njihovu procenu

Dva važna stohastička pokazatelja slučajnog procesa su autokovarijaciona i autokorelaciona funkcija.

Definicija autokovarijacione funkcije: Autokovarijaciona funkcija diskretnog široko stacionarnog slučajnog procesa $\{v(k)\}$ definisana je izrazom

$$R_v(\tau) = E\{[v(k) - m_v][v(k-\tau) - m_v]\}, \quad \tau \in Z \quad (3.42)$$

gde je parametar τ celobrojna promenljiva, koji se naziva i kašnjenje.

Autokovarijaciona funkcija je simetrična funkcija parametra τ

$$R_v(-\tau) = R_v(\tau) \quad (3.43)$$

pri čemu je

$$R_v(0) = \sigma_v^2 \quad (3.44)$$

Ova funkcija takođe zadovoljava uslov

$$|R_v(\tau)| \leq R_v(0) \quad \text{za} \quad \forall \tau \in Z \quad (3.45)$$

Definicija autokorelacione funkcije: Autokorelaciona funkcija diskretnog široko-stacionarnog slučajnog procesa definisana je izrazom

$$X_v(\tau) = E\{v(k)v^T(k-\tau)\}, \quad \tau \in Z \quad (3.46)$$

Slično kao autokovarijaciona funkcija i ova funkcija je simetrična, tj.

$$X(-\tau) = X(\tau) \quad (3.47)$$

Na osnovu kovarijacione funkcije često se u literaturi definiše i koeficijent korelacije, koji u osnovi predstavlja normalizovanu kovarijacionu funkciju čije vrednosti pripadaju intervalu $(-1, +1)$ [49]

$$\rho_v(\tau) = \frac{E\{v(k)v(k-\tau)\}}{E\{v(k)v(k)\}} \quad (3.48)$$

Primetimo da se kovarijaciona funkcija svodi na koeficijent korelacije ukoliko se koriste normalizovani (skalirani) i centrirani signali pri njenoj definiciji, odnosno

$$\tilde{v}(k) = \frac{v(k) - m_v}{\sqrt{\sigma_v^2}} = \frac{v(k) - m_v}{\sigma_v} \quad (3.49)$$

pri čemu se veličina σ_v naziva standardnom devijacijom.

Pošto je $E\{\tilde{v}(k)\} = m_{\tilde{v}} = 0$, tada je

$$R_{\tilde{v}}(\tau) = X_{\tilde{v}}(\tau) = \rho_{\tilde{v}}(\tau) = E\{\tilde{v}(k)\tilde{v}(k-\tau)\} = E\{\tilde{v}(k)\tilde{v}(k-\tau)\} \quad (3.50)$$

pošto je

$$\delta_{\tilde{v}}^2 = E\{\tilde{v}^2(k)\} = E\{[v(k) - m_v]^2\} = 1 \quad (3.51)$$

Na osnovu definicionih izraza, veza između autokovarijacione i autokorelacione funkcije data je izrazom

$$\begin{aligned}
 R_v(\tau) &= E\{(v(k) - m_v)(v(k - \tau) - m_v)\} \\
 &= E\{v(k)v(k - \tau)\} - E\{v(k)\}m_v - m_v E\{v(k - \tau)\} + m_v^2 \\
 &= X_v(\tau) - m_v^2 - m_v^2 + m_v^2 = X_v(\tau) - m_v^2
 \end{aligned} \tag{3.52}$$

Dakle, ukoliko je srednja vrednost signala $m_v = 0$, tada se ove dve funkcije poklapaju, tj. $R_v(\tau) = X_v(\tau)$.

Slično se mogu definisati i kros-kovariaciona i kros-korelaciona funkcija dva diskretna široko-stacionarna slučajna procesa $\{v(k)\}$ i $\{\omega(k)\}$ na osnovu izraza

$$R_{v\omega}(\tau) = E\{[v(k) - m_v][\omega(k - \tau) - m_\omega]\}, \quad \forall \tau \in Z \tag{3.53}$$

$$X_{v\omega}(\tau) = E\{v(k)\omega(k - \tau)\}, \quad \tau \in Z \tag{3.54}$$

Na sličan način kao u prethodno razmatranom slučaju, lako se pokazuje da je veza između ovih funkcija definisana izrazom

$$R_{v\omega}(\tau) = X_{v\omega}(\tau) - m_v m_\omega \tag{3.55}$$

Uočava se da ukoliko jedan od posmatranih signala ima nultu srednju vrednost, ove dve funkcije predstavljaju istu funkciju. Takođe važe relacije

$$R_{v\omega}(-\tau) \neq R_{v\omega}(\tau); \quad R_{v\omega}(-\tau) = R_{v\omega}(\tau) \tag{3.56}$$

$$X_{v\omega}(-\tau) \neq X_{v\omega}(\tau); \quad X_{v\omega}(-\tau) = X_{v\omega}(\tau) \tag{3.57}$$

pošto je za široko stacionarne slučajne procese ispunjeno

$$E\{v(k)\omega(k + i)\} = E\{\omega(k + \tau)v(k)\} = E\{\omega(k)v(k)\} = E\{\omega(k)v(k - \tau)\} \tag{3.58}$$

Slično kao statistički pokazatelji i stohastički pokazatelji se mogu proceniti (estimirati) na osnovu uzorka signala konačne dužine. Na primer, procene autokovarijacione funkcije može se generisati na osnovu izraza

$$\hat{R}_N(\tau) = \frac{1}{N} \sum_{k=1}^N [v(k) - \hat{m}_v][v(k - \tau) - \hat{m}_v]; \quad \hat{m}_v = \frac{1}{N} \sum_{k=1}^N v(k); \quad \forall \tau \in Z \tag{3.59}$$

dok je odgovarajuća procena auto-korelacione funkcije

$$\hat{X}_N(\tau) = \frac{1}{N} \sum_{k=1}^N v(k)v(k - \tau); \quad \forall \tau \in Z \tag{3.60}$$

Umesto ovakve procene može se koristiti i alternativni izraz koji usrednjava sve moguće proizvode parova odbiraka koji se nalazi na rastojanju $\tau \in Z$ na sledeći način

$$\begin{aligned}\hat{X}_N(\tau) &= \frac{v(0)v(|\tau|) + v(1)v(|\tau|+1) + \dots + v(N-1-|\tau|)v(N-1)}{N} \\ &= \frac{1}{N} \sum_{i=0}^{N-|\tau|-1} v(i)v(i+|\tau|); \quad \tau = 0, \pm 1, \pm 2, \dots, \pm N-1\end{aligned}\quad (3.61)$$

Primetimo da će za veliko τ postojati svega nekoliko proizvoda parova, dok za $\tau \geq N$ neće na raspolaganju ostati ni jedan par uzoraka čije je rastojanje jednako datom τ , tako da je

$$\hat{X}_N(\tau) = 0, \quad \tau \geq N \quad (3.62)$$

Ovako estimirana auto-korelaciona funkcija ispunjava uslov parnosti, koji važi i za originalnu funkciju $X(\tau)$

$$\hat{X}_N(-\tau) = \hat{X}_N(\tau) \quad (3.63)$$

Može se pokazati da je srednja vrednost ove procene [40,71]

$$E\{X_N(\tau)\} = \frac{1}{N} \sum_{j=0}^{N-1-|\tau|} E\{v(j)v(j+|\tau|)\} = \left(1 - \frac{|\tau|}{N}\right) X(\tau) \quad (3.64)$$

gde je $X(\tau)$ stvarna autokorelaciona funkcija koja se estimira. Dakle, pošto je $E\{X_N(\tau)\} \neq X(\tau)$ definisana procena je u opštem slučaju pomerena. Međutim, ova procena je asimptotski nepomerena, pošto $|\tau|/N \rightarrow 0$ kada $N \rightarrow \infty$. Slično kao kod procene varijanse, da bi dobili nepomerenu procenu data procena se mora pomnožiti sa faktorom $N/(N-|\tau|)$, odnosno odgovarajuću sumu treba podeliti sa $(N-|\tau|)$ umesto sa N . U primeni je ipak uobičajeno da se koristi izraz za pomerenu procenu, o čemu će više biti reči u poglavlju posvećenom estimaciji spektra signala (spektralnoj analizi).

Varijansa nepomerene procene definisana je izrazom

$$\text{var } \hat{X}_N(\tau) = E\left\{\left[\hat{X}_N(\tau) - E\{\hat{X}_N(\tau)\}\right]^2\right\} = E\left\{\left[\hat{X}_N(\tau - X(\tau))\right]^2\right\} \quad (3.65)$$

gde je

$$\hat{X}_N(\tau) - X(\tau) = \frac{1}{N} \sum_{j=0}^{N-1-|\tau|} [v(j)v(j+|\tau|) - X(\tau)] \quad (3.66)$$

odakle sledi

$$\text{var } \hat{X}_N(\tau) = \frac{1}{N^2} \sum_{j=0}^{N-1-|\tau|} \sum_{i=0}^{N-1-|\tau|} E \{ [v(j)v(j+|\tau|) - X(\tau)] [v(i)v(i+|\tau|) - X(\tau)] \} \quad (3.67)$$

Izračunavanje ovog izraza zahteva poznavanje četvrtog momenta

$$E \{ v(j)v(j+|\tau|)v(i)v(i+|\tau|) \} = m_{ij}(\tau) \quad (3.68)$$

Dakle, dalje izračunavanje zahteva dodatnu pretpostavku o četvrtom momentu razmatranog slučajnog procesa. Ako se uvede dodatna pretpostavka da je razmatrani slučajni proces normalan (gausovski), može se pokazati da je [49,71]

$$\text{var } \hat{X}_N^v(\tau) \approx \frac{2}{N} \sum_{r=0}^{N-1-|\tau|} \left[1 - \frac{\tau+r}{N} \right] [X^2(r) + X(r+\tau)X(r-\tau)] \quad (3.69)$$

Pošto u većini slučajeva $X(r)$ iščezava sa porastom sumiranja r , izraz u velikim (vitičastim zgradama) težiće u ovakvim slučajevima nuli sa porastom indeksa sumiranja. Sa druge strane, za veliko N gornja granica sume biće približno ∞ , a izraz u srednjim zgradama biće približno jednak 1 za one vrednosti r za koje izraz u velikim zgradama nije blizak nuli. Na taj način, suma će težiti ka konačnoj vrednosti, koja je jednaka zbiru nenultih članova u velikoj zgradi. Dakle, za veliko N jedina zavisnost procesa $\hat{X}_N(\tau)$ od dužine uzorka N biće preko faktora ispred sume $1/N$, tako da je izvedeni izraz za varijansu procene srazmeran sa $1/N$, te će težiti nuli kada N neograničeno raste, tj. ovakva procena je konzistentna. Drugim rečima, uslov $N \gg |\tau|$ obezbeđuje odgovarajuću tačnost ovakve procene, kada je približno ispunjen i uslov nepomerivosti.

Slično se može definisati i procena kros-korelacione funkcije

$$\hat{X}_N^{v\omega}(\tau) = \begin{cases} \frac{1}{N} \sum_{i=0}^{N-1-|\tau|} v(i)\omega(i+\tau) & \tau = 0, 1, \dots, N-1 \\ \frac{1}{N} \sum_{i=|\tau|}^{N-1} v(i)\omega(i+\tau) & \tau = -1, -2, \dots, -N+1 \\ 0 & |\tau| \geq N \end{cases} \quad (3.70)$$

U ovakvom definicionom izrazu postoji samo po N odbiraka za svaki od razmatranih široko-stacionarnih diskretnih slučajnih procesa $\{v(k)\}$ i $\{\omega(k)\}$. Analogno sa procenom auto-korelacione funkcije, ova procena će biti utoliko verodostojnija ukoliko je dužina mernog uzorka N znatno veće od usvojenog kašnjenja τ . U opštem slučaju važi $\hat{X}_{v\omega}(-\tau) \neq \hat{X}_{v\omega}(\tau)$ a srednja vrednost i varijansa procene mogu se analizirati na identičan način kao u slučaju autokorelacione funkcije.

3.2.3 Osnovne osobine slučajnih pobudnih signala

Do sada uvedene definicije i izvršene analize bazirale su se na pretpostavci o osobini široke (slabe) stacionarnosti ili široke (slabe) ergodičnosti razmatranog slučajnog procesa. Na ovom mestu definisaćemo nešto strožije ove dve osobine slučajnog procesa [49,71].

Definicija široke (slabe) stacionarnosti: Diskretni slučajni proces $\{v(k)\}$ je **široko (slabo) stacionaran** ukoliko ispunjava sledeća tri zahteva:

1. Srednja snaga signala ograničena je sa gornje strane konačnom veličinom $E\{v^2(k)\} < \infty$.

2. Srednja vrednost signala je vremenski-invarijanta $m_v(k) = E\{v(k)\} = m_v = c$ gde je c realna konstanta.

3. Autokorelaciona funkcija slučajnog procesa $X_v(\tau) = E\{v(k)v(k-\tau)\}$ zavisi samo od kašnjenja τ , a ne i od apsolutne vrednosti vremenskih trenutaka k , $k-\tau$ u kojima se traži srednja vrednost proizvoda odbiraka signala.

Primetimo da pretpostavke 1 i 2 označavaju i da je varijansa slučajnog procesa $\sigma_v^2 = E\{v^2(k)\} - m_v^2$ konačna konstantna, tj. ne zavisi od vremenskog trenutka k i ima konačnu vrednost.

Slučajni proces je **široko ergodičan** ukoliko ispunjava tri sledeća uslova:

1. Diskretni signal $\{v(k)\}$ je široko stacionaran.

2. Srednja vrednost, odnosno usrednjavanje po ansamblu realizacija na osnovu matematičkog očekivanja $E\{\cdot\}$ jednaka je usrednjavanju odbiraka po vremenu (aritmetičkoj sredini vrednosti odbiraka)

$$m_v = E\{v(k)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N v(k) \quad (3.71)$$

3. Autokorelaciona funkcija dobijena usrednjavanjem po ansamblu realizacija (primenom matematičkog očekivanja) jednaka je proceni autokorelacione funkcije dobijene na osnovu usrednjavanja proizvoda odbiraka po vremenu

$$X_v(\tau) = E\{v(k)v(k-\tau)\} = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{i=-N}^N v(i)v(i-|\tau|); \quad \tau \in Z \quad (3.72)$$

Važan stohastički pokazatelj predstavlja i **kovarijaciona matrica**, koja u osnovi predstavlja generalizaciju pojma varijanse na slučaj vektorskih (multiidmenzionalnih) slučajnih varijabli [63,64].

Definicija kovarijacione matrice: Kovarijaciona matrica $R_v^N(k)$ diskretnog široko stacionarnog slučajnog signala $\{v(k)\}$ definisano je kao matematičko očekivanje proizvoda vektora kolone, koji se sastoji od N prethodnih odbiraka signala, i odgovarajućeg vektora vrste, tj.

$$R_v^N(k) = E \left\{ \begin{bmatrix} v(k) - m_v \\ v(k-1) - m_v \\ \vdots \\ v(k-N+1) - m_v \end{bmatrix} \begin{bmatrix} v(k) - m_v \\ v(k-1) - m_v \\ \vdots \\ v(k-N+1) - m_v \end{bmatrix}^T \right\} = E \{ v_N(k) v_N^T(k) \} \quad (3.73)$$

gde je $m_v = E\{v(k)\}$, a $v_N(k)$ N -dimenzioni vektor kolona koji sadrži tekući odbirak signala i prethodne odbitke u okviru prozoru podataka.

Ova matrica je simetrična, pošto je $(R_v^N(k))^T = R_v^N(k)$ kao posledica činjenice da je $E\{[v(k) - m_v][v(k-\tau) - m_v]\} = E\{[v(k-\tau) - m_v][v(k) - m_v]\}$.

Kovarijaciona matrica se može izraziti preko kovarijacione funkcije $R_v(\tau)$ na sledeći način

$$R_v^N(k) = \begin{bmatrix} R_v(0) & R_v(1) & \dots & R_v(N-1) \\ R_v(-1) & R_v(0) & \dots & R_v(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_v(1-N) & R_v(2-N) & \dots & R_v(0) \end{bmatrix} \quad (3.74)$$

Diskretni beli šum: Slučajni proces sa proizvoljnim korelacionim svojstvima može se generisati propuštanjem specifičnog slučajnog procesa, tzv. „belog šuma“, kroz linearan vremenski-invarijantan kauzalan diskretni sistem. Beli šum je široko stacionaran slučajni proces koji je opisan samo na osnovu prvog i drugog statističkog momenta. Specijalan slučaj belog šuma je beli slučajni proces sa nultom srednjom vrednošću, koji je definisan sa [41,63,64,71,81]

$$m_e = E\{e(k)\} = 0 \quad \text{za} \quad \forall k \in Z \quad (3.75)$$

$$X_e(k, \ell) = X_e(k - \ell) = E\{e(k)e(\ell)\} = \begin{cases} \sigma_e^2 & \text{za} \quad k = \ell \in Z \\ 0 & \text{za} \quad k \neq \ell; k, \ell \in Z \end{cases} \quad (3.76)$$

Ako se posmatra distribucija frekvencija u frekvencijskom liku (spektru) ovakvog signala sve frekvencije biće ravnopravno zastupljene, odnosno spektralna gustina snage takvog signala biće ista na svim frekvencijama, slično kao kod bele svetlosti u optici.

Uvedena definicija ništa ne govori o zakonu raspodele verovatnoće slučajne amplitude $e(k)$ u diskretnom trenutku k . U praksi se obično uvodi dodatna pretpostavka da je $e(k)$ normalno (gausovski) raspodeljen slučajna varijabla i takav

diskretan signal se zove beli gausovski (normalan) slučajni proces (beli gausovski šum). Normalna raspodela nulte srednje vrednosti $m_e = 0$ i varijanse σ_e^2 , označava se kao $N(0, \sigma_e^2)$. Kao što je istaknuto u prethodnom poglavlju, autokorelaciona funkcija belog šuma može se proceniti na osnovu jedne realizacije slučajnog procesa, tj. vremenske serije $\{e(k), k = 0, 1, \dots, N\}$ kao

$$\hat{R}_N^e(\tau) = \frac{1}{N} \sum_{i=0}^{N-|\tau|-1} e(i)e(i+|\tau|); \quad \tau = 0, \pm 1, \dots, \pm N \quad (3.77)$$

pri čemu je srednja vrednost ove procene

$$E\{\hat{R}_N^e(\tau)\} = \left(1 - \frac{|\tau|}{N}\right) R_e(\tau) \approx R_e(\tau) \quad \text{za} \quad \tau \ll N \quad (3.78)$$

gde je $R_e(\tau)$ teorijska autokorelaciona funkcija belog šuma, dok je varijansa procene

$$\begin{aligned} \text{var } \hat{R}_N^e(\tau) &\approx \frac{2}{N} \sum_{r=0}^{N-1-|\tau|} \left(1 - \frac{\tau+r}{N}\right) [R_e^2(r) + R_e(\tau+r)R_e(r-\tau)] \\ &\approx \frac{2}{N} \left(1 - \frac{\tau}{N}\right) R_e^2(0) \approx \frac{2}{N} R_e^2(0) \quad \text{za} \quad \tau \ll N \end{aligned} \quad (3.79)$$

Da bi se procenilo da li procene $\hat{R}_N^e(1), \hat{R}_N^e(2), \dots$ zaista odgovaraju autokorelacionoj funkciji belog šuma potrebno je da ove vrednosti budu istog reda veličine ili manje od empirijske standardne devijacije $\sqrt{\text{var } \hat{R}_N^e(\tau)} \approx R_e(0)/\sqrt{N/2} = \sigma_e^2/\sqrt{N/2}$. U protivnom, zaključuje se da vrednosti $\hat{R}_N^e(1), \hat{R}_N^e(2), \dots$, znatno odstupaju od teorijske nulte vrednosti autokorelacione funkcije $R_e(\tau) = 0$ za $\tau \neq 0$, te takva procena autokorelacione funkcije ne potiče od belog šuma.

Korelisani (obojeni) šum: Kao što je prethodno istaknuto, poremećaj $\{v(k)\}$ se obično generiše propuštanjem belog šuma $\{e(k)\}$ kroz linearan sistem, [49, 71, 81] tj.

$$v(k) = \sum_{\ell=0}^{\infty} h(\ell)e(k-\ell) \quad (3.80)$$

ili ekvivalentno koristeći koncept funkcije prenosa

$$v(k) = H(q)e(k) \quad (3.81)$$

gde je $h(\ell)$ impulsni odziv linearnog diskretnog vremenski-invarijantnog kauzalnog sistema na čiji se ulaz dovodi beli šum a čiji je izlaz slučajni poremećaj $v(k)$, a $H(q)$ je funkcija prenosa ovog sistema. Pošto se poremećaj $\{v(k)\}$ dobija filtriranjem belog

šuma $\{e(k)\}$, on se naziva i obojeni (korelisani) šum, po analogiji sa obojenom svetlošću.

3.2.4 Spektralna reprezentacija slučajnih signala

Široko stacionaran slučajni proces kompletno je opisan u vremenskom domenu statističkim i stohastičkim pokazateljima, odnosno srednjom vrednošću (statistički pokazatelj) i kovarijacionom funkcijom (stohastički pokazatelj). Ekvivalentan opis sistema u frekvencijskom domenu je spektar snage (eng. "power spectrum") slučajnog (stohastičkog) signala. Spektar snage široko stacionarnog slučajnog signala (procesa) se definiše kao Fourierova transformacija njegove autokovarijacione funkcije i daje informacije o raspodeli energije signala na različitim frekvencijama [49,64,71,81]. Pre same formalne definicije spektra, razmotrimo neke dodatne činjenice.

Kao što je prethodno istaknuto, kovarijaciona funkcija je definisana samo za slučajne (stohastičke) signale. Da bi se u razmatranje uvela i kategorija determinističkih signala, definiše se operator generalisanog matematičkog očekivanja [63,64]

$$\overline{E\{\cdot\}} = \frac{1}{N} \sum_{k=1}^N E\{\cdot\} \quad (3.82)$$

U slučaju stohastičkih signala generalisano matematičko očekivanje $\overline{E\{\cdot\}}$ se svodi na klasično matematičko očekivanje $E\{\cdot\}$, pošto sumiranje nema efekta, odnosno veličine $E\{\cdot\}$ se može izvući van sume, nakon čega se dobija

$$\overline{E\{s(n)\}} = \frac{1}{N} \sum_{k=1}^N E\{s(k)\} = E\{s(k)\} \frac{1}{N} \sum_{k=1}^N 1 = E\{s(k)\} = m_s = \text{const.} \quad (3.83)$$

Sa druge strane, u slučaju determinističkih signala klasičan operator matematičkog očekivanja nema efekta, pošto je matematičko očekivanje determinističkog signala jednako samom signalu, te se generalisano matematičko očekivanje svodi na aritmetički zbir odbiraka ovog signala, odnosno na vremensko usrednjavanje

$$\overline{E\{s(n)\}} = \frac{1}{N} \sum_{k=1}^N E\{s(k)\} = \frac{1}{N} \sum_{k=1}^N s(k) = m_s = \text{const.} \quad (3.84)$$

Uvođenjem operatora generalisanog matematičkog očekivanja i stohastički i deterministički signali mogu se tretirati na isti način, odnosno iste definicije i osobine mogu se primeniti na obe vrste signala. Na primer, pretpostavka o širokoj (slaboj) stacionarnosti i ergodičnosti slučajnog signala može se na sličan način definisati i za slučaj determinističkih signala. Imajući ovo u vidu, dalje ćemo pretpostaviti da su svi razmatrani signali, bilo da su stohastički ili deterministički, kvazi-stacionarni u smislu sledeće definicije.

Definicija osobine kvazi-stacionarnosti signala: Signal, bilo da je deterministički ili stohastički, naziva se kvazistacionarnim ukoliko su ispunjene sledeće tvrdnje:

1. Generalisano matematičko očekivanje $\overline{E}\{s(t)\} = \bar{s}(t)$ je ograničeno za svako celobrojno t , tj. $\forall t \in Z$.
2. Kovarijaciona funkcija $\overline{E}\{[s(t) - \bar{s}(t)][s(r) - \bar{s}(r)]\} = R_s(t, r)$ je ograničena za sve celobrojne vrednosti argumenata t i r , tj. $\forall t, r \in Z$, pri čemu je

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N R_s(t, t - \tau) = R_s(\tau) \quad \text{za} \quad \forall t, \tau \in Z \quad (3.85)$$

Navedena definicija, u osnovi, podrazumeva da srednja vrednost i energija signala imaju konačne vrednosti. Polazeći od jedinstvene definicije autokovarijacione funkcije, za stohastičke i determinističke signale, sada se na jedinstven način za obe vrste signala može definisati i njihov spektar snage, koji se često naziva i spektralna gustina snage.

Definicija spektra snage signala: Spektar snage $S_v(\omega)$ vremenski diskretnog signala $\{v(k)\}$ je Fourierova transformacija njegove auto-kovarijacione funkcije $R_v(\tau)$, odnosno

$$S_v(\omega) = \sum_{\tau=1}^N R_v(\tau) e^{-i\tau\omega} \quad (3.86)$$

gde je auto-kovarijaciona funkcija definisana na osnovu generalisanog matematičkog očekivanja kao

$$R_v(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N R_v(t, t - \tau); \quad R_v(t, r) = \overline{E}\{[v(t) - \bar{v}(t)][v(r) - \bar{v}(r)]\} \quad (3.87)$$

pri čemu je $\bar{v}(t) = \overline{E}\{v(t)\}$, a $\overline{E}\{\cdot\}$ označava operator generalisanog matematičkog očekivanja, podrazumevajući da navedena beskonačna suma postoji.

Kros (među) – spektar snage diskretnih signala $\{v(k)\}$ i $\{\omega(k)\}$, $S_{vw}(\omega)$, predstavlja Fourierovu transformaciju njihove kros-kovarijacione funkcije, $R_{vw}(\tau)$, odnosno

$$S_{vw}(\omega) = \sum_{\tau=-\infty}^{\infty} R_{vw}(\tau) e^{-i\tau\omega} \quad (3.88)$$

gde je kros-kovarijaciona funkcija definisana sa

$$R_{vw}(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N R_{vw}(t, t - \tau) \quad (3.89)$$

pri čemu je

$$R_{vw}(t, r) = \overline{E} \{ [v(t) - \bar{v}(t)] [w(r) - \bar{w}(r)] \}; \quad \bar{s}(t) = \overline{E} \{ s(t) \} \quad (3.90)$$

pretpostavljajući egzistenciji navedene beskonačne sume.

Auto-spektar snage

$$S_v(\omega) = \sum_{\tau=-\infty}^{\infty} R_v(\tau) \cos(\omega\tau) - i \sum_{\tau=-\infty}^{\infty} R_v(\tau) \sin(\omega\tau) \quad (3.91)$$

je realna i parna funkcija ($S_v(-\omega) = S_v(\omega)$) realnog argumenta ω , pošto je $R_v(\tau)$ parna funkcija pa je imaginarni deo navedenog izraza jednak nuli, dok je kros-spektar snage $S_{vw}(\omega)$ u opštem slučaju kompleksna funkcija realnog vremena ω , pa se njegov realni deo $R_e \{ S_{vw}(\omega) \}$ naziva kospektar (engleski “cospectrum”) a imaginarni deo $I_m \{ S_{vw}(\omega) \}$ kvadrturni spektar (eng. “quadrature spectrum”).

Pošto je na osnovu definicije inverzne Fourierove transformacije [24,48]

$$R_v(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_v(\omega) e^{i\pi\omega} d\omega \quad (3.92)$$

zaključuje se da je

$$R_v(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_v(\omega) d\omega \quad (3.93)$$

pri čemu je $R_v(0) = \sigma_v^2$.

3.2.4.1 Spektar diskretnog belog šuma

Posmatrajmo diskretan beli slučajni proces $\{e(k)\}$, nulte srednje vrednosti $\bar{e}(k) = 0$ i varijanse $\overline{E} \{ e^2(k) \} = \sigma_e^2$. Kovarijaciona funkcija ovog signala data je sa

$$R_e(\tau) = \overline{E} \{ [e(t) e(t - \tau)] \} = \sigma_e^2 \Delta(\tau) \quad (3.94)$$

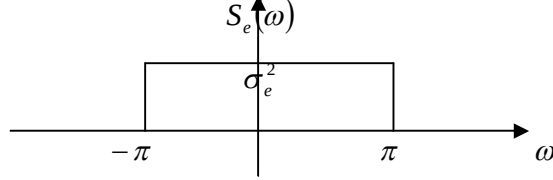
gde je jedinični Kroneckerov delta impuls $\Delta(\tau)$ definisan sa:

$$\Delta(\tau) = \begin{cases} 1, & \tau = 0 \\ 0, & \tau \neq 0 \text{ i } \tau \in Z \end{cases} \quad (3.95)$$

Tada je spektar snage ovog signala definisan izrazom

$$S_e(\omega) = \sum_{\tau=-\infty}^{\infty} R_e(\tau) e^{-i\tau\omega} = \sigma_e^2, \quad -\pi \leq \omega \leq \pi \quad (3.96)$$

Spektar snage diskretnog belog šuma prikazan je na sledećoj slici



Slika3.5: Sepktar snage belog šuma

Primetimo da zbog periodičnosti kompleksne eksponencijalne funkcije

$e^{-i\omega_0\tau} = \cos(\omega\tau) - i\sin(\omega\tau)$ sa osnovnom periodom 2π i spektralna karakteristika snage $S_e(\omega)$ predstavlja periodičnu matematičku funkciju argumenta ω sa periodom 2π , ali osnovni fizički spektar snage diskretnog signala $\{e(k)\}$ je ograničen samo na osnovni opseg učestanosti $(-\pi, +\pi)$, koji se naziva i Niquistovo područje.

3.2.4.2 Transformacija spektra signala

Osnovni zadatak analize sistema sastoji se u određivanju izlaznog signala (odziva) na zadati ulazni signal (pobudu ili eksitaciju). Ovakva analiza se vrši u vremenskom domenu a njen ekvivalent u frekvencijskom domenu predstavlja zadatak određivanja spektra snage izlaznog signala za poznati spektar snage signala na ulazu sistema.

Izlazni signal generalnog „black box“ modela sistema na slici 3.4 gde je $v(k)$ dato sa (3.80) definisan je relacijama

$$y(k) = G(q)u(k) + H(q)e(k) = \sum_{\ell=1}^{\infty} g(\ell)u(k-\ell) + \sum_{\ell=0}^{\infty} h(\ell)e(k-\ell) \quad (3.97)$$

gde su $y(k)$, $u(k)$ i $e(k)$ signali izlaza, ulaza i poremećaja, respektivno. Dakle, sistem u suštini ima dve pobude, koristan ulazni signal $u(k)$ i parazitnu eksitaciju $e(k)$. Pretpostavimo, dalje, da ovi pobudni signali imaju nulte srednje vrednosti i da su nekorelisani međusobno, odnosno da je $\bar{u} = \bar{e} = 0$, $R_{ue}(\tau) = 0$ za $\forall \tau \in Z$.

Tada je auto-kovarijaciona funkcija izlaza

$$\begin{aligned} R_y(\tau) &= \bar{E}\{y(k)y(k-\tau)\} \\ &= \bar{E}\left\{\left[\sum_{\ell=1}^{\infty} g(\ell)u(k-\ell) + \sum_{\ell=0}^{\infty} h(\ell)e(k-\ell)\right]\left[\sum_{m=1}^{\infty} g(m)u(k-\tau-m) + \sum_{m=0}^{\infty} h(m)e(k-\tau-m)\right]\right\} \end{aligned}$$

$$\begin{aligned}
 &= \bar{E} \left\{ \sum_{\ell=1}^{\infty} g(\ell) u(k-\ell) \sum_{m=1}^{\infty} g(m) u(k-\tau-m) \right\} + \bar{E} \left\{ \sum_{\ell=0}^{\infty} h(\ell) e(k-\ell) \sum_{m=0}^{\infty} h(m) e(k-\tau-m) \right\} \\
 &+ \bar{E} \left\{ \sum_{\ell=1}^{\infty} g(\ell) u(k-\ell) \sum_{m=0}^{\infty} h(m) e(k-\tau-m) \right\} + \bar{E} \left\{ \sum_{\ell=0}^{\infty} h(\ell) e(k-\ell) \sum_{m=1}^{\infty} g(m) u(k-\tau-m) \right\} \\
 &= \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} g(\ell) g(m) \bar{E} \{ u(k-\ell) u(k-\tau-m) \} + \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} h(\ell) h(m) \bar{E} \{ e(k-\ell) e(k-\tau-m) \} \\
 &+ \sum_{\ell=1}^{\infty} \sum_{m=0}^{\infty} g(\ell) h(m) \bar{E} \{ u(k-\ell) e(k-\tau-m) \} + \sum_{\ell=0}^{\infty} \sum_{m=1}^{\infty} h(\ell) g(m) \bar{E} \{ e(k-\ell) u(k-\tau-m) \} \\
 &= \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} g(\ell) g(m) R_u \{ (k-\ell) - (k-\tau-m) \} + \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} h(\ell) h(m) R_e \{ (k-\ell) - (k-\tau-m) \} \\
 &+ \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} g(\ell) h(m) R_{ue} \{ (k-\ell) - (k-\tau-m) \} + \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} h(\ell) g(m) R_{eu} \{ (k-\ell) - (k-\tau-m) \}
 \end{aligned} \tag{3.98}$$

Prilikom izvođenja poslednjeg izraza iskorišćena je činjenica da je $\bar{E}\{\cdot\}$ linearan operator, koji može da zameni redosled sa operacijom sumiranja, kao i da je model procesa (sistema) $G(q)$ striktno kauzalan, odnosno da je odbirak impulsnog odziva $g(0)=0$. Imajući dalje u vidu da je na osnovu pretpostavke o nekorelisanosti korisne i parazitne pobude kros-korelacione funkcije ovih signala $R_{eu}(\tau)=0, R_{ue}(\tau)=0$ za $\forall \tau \in Z$, konačno se dobija

$$R_y(\tau) = \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} g(\ell) g(m) R_u(\tau+m-\ell) + \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} h(\ell) h(m) R_e(\tau+m-\ell) \tag{3.99}$$

Spektar snage izlaznog signala definisan je izrazom

$$\begin{aligned}
 S_y(\omega) &= \sum_{\tau=-\infty}^{\infty} R_y(\tau) e^{-i\tau\omega} \\
 &= \sum_{\tau=-\infty}^{\infty} \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} g(\ell) g(m) R_u(\tau+m-\ell) e^{-i\tau\omega} + \sum_{\tau=-\infty}^{\infty} \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} h(\ell) h(m) R_e(\tau+m-\ell) e^{-i\tau\omega} \\
 &= \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} \sum_{\tau=0}^{\infty} g(\ell) g(m) R_u(\tau+m-\ell) e^{-i\tau\omega} e^{-i\ell\omega} e^{-im\omega} e^{-im\omega} \\
 &+ \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} h(\ell) h(m) R_e(\tau+m-\ell) e^{-i\tau\omega} e^{-i\ell\omega} e^{-im\omega} e^{-im\omega} \\
 &= \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} \sum_{\tau=-\infty}^{\infty} g(\ell) e^{-\ell\omega} g(m) e^{im\omega} R_u(\tau+m-\ell) e^{-i(\tau+m-\ell)\omega}
 \end{aligned}$$

$$+ \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} \sum_{\tau=-\infty}^{\infty} h(\ell) e^{-i\ell\omega} h(m) e^{im\omega} R_e(\tau + m - \ell) e^{-i(\tau+m-\ell)\omega} \quad (3.100)$$

Uvodeći smenu $j = \tau + m - \ell$, pri čemu je j celobrojna promenljiva iz intervala $(-\infty, \infty)$, konačno se može pisati

$$S_y(m) = \sum_{\ell=1}^{\infty} g(\ell) e^{-i\ell\omega} \sum_{m=1}^{\infty} g(m) e^{im\omega} \sum_{j=-\infty}^{\infty} R_u(j) e^{-ij\omega} + \sum_{\ell=0}^{\infty} h(\ell) e^{-i\ell\omega} \sum_{m=0}^{\infty} h(m) e^{im\omega} \sum_{j=-\infty}^{\infty} R_e(j) e^{-ij\omega} \quad (3.101)$$

odnosno

$$S_y(m) = G(e^{i\omega}) G(e^{-i\omega}) S_u(\omega) + H(e^{i\omega}) H(e^{-i\omega}) S_e(\omega) \quad (3.102)$$

Pošto je

$$G(e^{i\omega}) G(e^{-i\omega}) = G(e^{i\omega}) G^*(e^{i\omega}) = |G(e^{i\omega})|^2 \quad (3.103)$$

gde simbol “*” označava konjugovano-kompleksnu vrednost odgovarajućeg kompleksnog broja, a “|·|” označava apsolutnu vrednost odgovarajućeg kompleksnog broja, takođe se može pisati

$$S_y(\omega) = |G(e^{i\omega})|^2 S_u(\omega) + |H(e^{i\omega})|^2 S_e(\omega) \quad (3.104)$$

Dakle, pod uslovom da su ulazni signal i signal poremećaja (parazitna pobuda) međusobno nekorelisani, spektar snage izlaznog signala jednak je linearnoj kombinaciji (težinskoj sumi) spektara snaga korisne i parazitne pobude, pri čemu koeficijenti linearne kombinacije (težinski koeficijenti) definišu pojačanja sistema po snazi u odnosu na odgovarajući pobudni signal, bilo da predstavlja korisnu ili parazitnu pobudu.

Na sličan način se može izvesti kros-spektralna karakteristika snage između signala izlaza i signala ulaza. Naime, pošto je kros-korelaciona funkcija ovih signala

$$R_{yu}(\tau) = \bar{E}\{y(k)u(k-\tau)\} \quad (3.105)$$

gde je $\bar{E}\{\cdot\}$ operator generalisanog matematičkog očekivanja, a izlazni signal je definisan sa

$$y(k) = G(q)u(k) + H(q)e(k) = \sum_{\ell=1}^{\infty} g(\ell)u(k-\ell) + \sum_{\ell=0}^{\infty} h(\ell)e(k-\ell) \quad (3.106)$$

dalje se može pisati

$$\begin{aligned}
 R_{yu}(\tau) &= \bar{E} \left\{ \sum_{\ell=1}^{\infty} g(\ell) u(k-\ell) u(k-\tau) \right\} + \sum_{\ell=0}^{\infty} h(\ell) e(k-\ell) u(k-\tau) \\
 &= \sum_{\ell=1}^{\infty} g(\ell) \bar{E} \{ u(k-\ell) u(k-\tau) \} + \sum_{\ell=0}^{\infty} h(\ell) \bar{E} \{ e(k-\ell) u(k-\tau) \} \\
 &= \sum_{\ell=1}^{\infty} g(\ell) R_u(\tau-\ell) + \sum_{\ell=0}^{\infty} h(\ell) R_{eu}(\tau-\ell)
 \end{aligned} \tag{3.107}$$

Imajući u vidu da su $\{e(k)\}$ i $\{u(k)\}$ međusobno nekorelisani vremenski diskretni signali, to je $R_{eu}(\tau) = 0$ za $\forall \tau \in Z$, pa je

$$R_{yu}(\tau) = \sum_{\ell=1}^{\infty} g(\ell) R_u(\tau-\ell) = \sum_{\ell=0}^{\infty} g(\ell) R_u(\tau-\ell) \tag{3.108}$$

pošto je $g(0) = 0$.

Dalje je kros-spektar ovih signala definisan sa

$$\begin{aligned}
 S_{yu}(\omega) &= \sum_{\tau=-\infty}^{\infty} R_{yu}(\tau) e^{-i\tau\omega} \\
 &= \sum_{\tau=-\infty}^{\infty} \sum_{\ell=0}^{\infty} g(\ell) R_u(\tau-\ell) e^{-i\tau\omega} e^{i\ell\omega} e^{-i\ell\omega} = \sum_{\ell=0}^{\infty} \sum_{\tau=-\infty}^{\infty} g(\ell) e^{-i\ell\omega} R_u(\tau-\ell) e^{-i\omega(\tau-\ell)}
 \end{aligned} \tag{3.109}$$

Uvodeći smenu $j = \tau - \ell$, celobrojna promenljiva j pripada skupu $(-\infty, \infty)$, te se može pisati

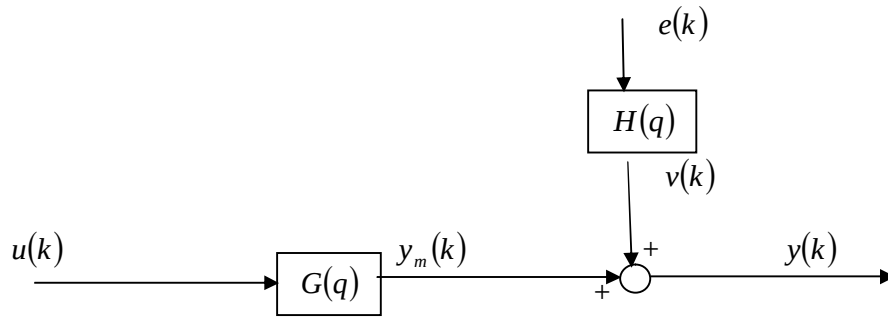
$$S_{yu}(\omega) = \sum_{\ell=0}^{\infty} g(\ell) e^{-i\ell\omega} \sum_{j=-\infty}^{\infty} R_u(j) e^{-ij\omega} \tag{3.110}$$

odnosno

$$S_{yu}(\omega) = G(e^{i\omega}) S_u(\omega) \tag{3.111}$$

3.2.5 Opšta neparametarska reprezentacija black-box modela sistema: strukturni blok dijagram

Na osnovu predhodne analize model sistema, koji obuhvata i kontrolni signal $u(k)$ i poremećaj $v(k)$, prikazan je sledećim blok dijagramom



Slika3.6: Model sistema sa poremećajem čije je dejstvo svedeno na izlaz sistema.

Prikazani model je opšti “black-box” model i predstavlja neparametarsku reprezentaciju u vremenskom domenu. Izlaz modela dat je relacijom

$$y(k) = y_m(k) + v(k) = G(q)u(k) + H(q)e(k) \quad (3.112)$$

U ovakvom modelu obično se podrazumevaju sledeće tvrdnje:

1. Funkcija prenosa $G(q)$ je striktno kauzalna, tj. $g(\ell) = 0$ za $\ell \leq 0$.
2. Funkcija prenosa $H(q)$ je kauzalni monic-funkcija, tj. $h(\ell) = 0$ za $\ell < 0$ i $h(0) = 1$.
3. Razmatrani sistemi koji generišu izlaz procesa $y_m(k)$ i ekvivalentni poremećaj na izlazu $v(k)$ su linearni i vremenski-invarijantni, a modelovani su funkcijama prenosa $G(q)$ i $H(q)$, respektivno.
4. Svi slučajni procesu su generisani kao široko (slabo) ergodični slučajni procesi.

3.3. PARAMETARSKA REPREZENTACIJA BLACK-BOX MODELA SISTEMA U VREMENSKOM DOMENU

Korišćenje modela sistema, bilo da se radi o simulaciji sistema, u cilju analize njegovih osobina, ili predikciji, u cilju predviđanja osobina sistema u budućnosti, ili sintezi, odnosno projektovanju sistema koji ima unapred zadata svojstva, znatno je jednostavnije ukoliko je broj parametara u modelu konačan, pri čemu je poželjno da se broj parametara što je moguće smanji. Sa tog aspekta parametarski modeli su poželjniji za rad u poređenju sa neparametarskim modelima.

Najčešće korišćene parametarske reprezentacije modela sistema su sledeće [11,28,49,63]:

1. Model sa konačnim impulsnim odzivom (eng. “Finite Impulse Response Model”), koji se označava kao FIR model.

3. Autoregresioni model sa spoljašnjom pobudom (eng. "Auto-Regressive with eXogenous input model"), koji se skraćeno označava sa ARX.

3. Autoregresioni model sa pomičnim sredinama i spoljašnjom pobudom (eng. "Auto-regressive Moving Average with eXogenous input model"), ili skraćeno ARMAX model.

4. Model greške jednačine (eng. "Output Error Model") ili skraćeno OE model.

5. Box-Jenkins-ov model, koji se po inicijalima autora označava kao BJ.

Svi navedeni modeli predstavljaju specijalan slučaj sledeće opšte forme reprezentacije (strukture) modela u vremenskom području:

$$A(q)y(k) = \frac{B(q)}{E(q)}u(k) + \frac{C(q)}{D(q)}e(k) \quad (3.113)$$

gde su $A(q), B(q), C(q), D(q)$ i $E(q)$ polinomi po promenljivoj q , koja predstavlja operator jediničnog prednjačenja, dok je q^{-1} operator jediničnog kašnjenja, tj.

$$A(q) = 1 + a_1q^{-1} + a_2q^{-2} + \dots + a_{n_a}q^{-n_a} = 1 + \sum_{i=1}^{n_a} a_i q^{-i}$$

$$B(q) = b_1q^{-1} + b_2q^{-2} + \dots + b_{n_b}q^{-n_b} = \sum_{i=1}^{n_b} b_i q^{-i}$$

$$C(q) = 1 + c_1q^{-1} + c_2q^{-2} + \dots + c_{n_c}q^{-n_c} = 1 + \sum_{i=1}^{n_c} c_i q^{-i}$$

$$D(q) = 1 + d_1q^{-1} + d_2q^{-2} + \dots + d_{n_d}q^{-n_d} = 1 + \sum_{i=1}^{n_d} d_i q^{-i}$$

$$E(q) = 1 + f_1q^{-1} + f_2q^{-2} + \dots + f_{n_f}q^{-n_f} = 1 + \sum_{i=1}^{n_f} f_i q^{-i} \quad (3.114)$$

U navedenoj opštoj strukturi modela diskretni signali $\{y(k)\}, \{u(k)\}$ i $\{e(k)\}$ predstavljaju izlaz, ulaz i poremećaj, respektivno. Dakle, struktura (reprezentacija) modela je u potpunosti definisana izborom navedenih polinoma, odnosno njihovog reda i brojnih vrednosti odgovarajućih parametara. Drugim rečima, prikazana opšta struktura (reprezentacija) modela je poznata sa tačnošću do nepoznatih parametara u navedenim polinomima koji definišu ovu strukturu. Zavisno od namene modela, ponekad je uobičajeno da se izvesni od polinoma u opštoj strukturi modela fiksiraju na jedinicu (predstavlja polinom nultog reda). Na primer, FIR i ARX strukture rezultuju u linearne jednačine u odgovarajućim algoritmima za identifikaciju sistema,

tj. estimaciji parametara modela na bazi merenja ulazno-izlaznih signala, što je vrlo povoljno sa računskog stanovišta.

3.3.1 FIR model (model sa konačnim impulsnim odzivom)

Kao što je prethodno navedeno, model sa beskonačnim impulsnim odzivom (eng. Infinite Impulse Response model) ili skraćeno IIR model definisan je sa diferencnom (diferencnom) jednačinom

$$y(k) = \sum_{\ell=1}^{\infty} g(\ell)u(k-\ell) + e(k) \quad (3.115)$$

gde su $g(\ell)$ odbirci impulsnog odziva, a nizovi $y(k)$, $u(k)$ i $e(k)$ predstavljaju izlaz, ulaz i poremećaj, respektivno. Odgovarajuća funkcija prenosa sistema je

$$G(q) = \sum_{\ell=1}^{\infty} g(\ell)q^{-\ell} \quad (3.116)$$

gde je q^{-1} operator jediničnog kašnjenja. Ukoliko je $G(q)$ stabilna funkcija prenosa tada impulsni odziv iščezava sa vremenom, odnosno nakon završetka prelaznog režima izlaz sistema (impulsni odziv) u ravnotežnom stanju biće jednak nuli (u ravnotežnom stanju izlaz, tj. impulsni odziv $g(k)$, sistema prati ulaz (impulsni signal $\delta(k)$, koji je jednak nuli za svaki diskretan trenutak k , sem za $k = 0$). Dakle, za stabilan sistem postoji pozitivna i konačna konstanta M , za koju je ispunjeno $g(\ell) = 0$ za $\ell > M$; $0 < M < \infty$.

Drugim rečima, potrebno je uzeti u razmatranje samo prvih M odbiraka impulsnog odziva, što rezultuje u FIR model

$$y(k) = \sum_{\ell=1}^M g(\ell)u(k-\ell) + e(k) \quad (3.117)$$

Ovakav model predstavlja specijalan slučaj opšte strukture u kojoj je $A(q) = C(q) = D(q) = F(q) = 1$ dok je red n_b polinoma $B(q)$ jednak M , tj. $n_b = M$. Dakle, FIR model se može prikazati i u obliku

$$y(k) = B(q)u(k) + e(k) \quad (3.118)$$

gde je

$$B(q) = \sum_{i=1}^M b_i q^{-i} \quad (3.119)$$

pri čemu parametri b_i imaju značenje odbiraka impulsnog odziva, tj.

$$b_i = g(i); \quad i = 1, \dots, M; \quad g(0) = b_0 = 0 \quad (3.120)$$

Primetimo, takođe, da su poremećaji u FIR modelu opisani belim stohastičkim procesom, tj. predstavljaju beli šum. Drugim rečima, u modelu korelisanog poremećaja ili obojenog šuma

$$v(k) = H(q)e(k), \quad H(q) = 1 \quad (3.121)$$

gde je $e(k)$ beli šum, funkcija prenosa $H(q) = 1$.

3.3.2 ARX model (Auto-regresioni model sa spoljašnjim ulazom)

Ovaj model predstavlja specijalni oblik opšte strukture u kojoj je $C(q) = D(q) = F(q) = 1$, pri čemu su $A(q)$ i $B(q)$ proizvoljni polinomi, tj.

$$A(q)y(k) = B(q)u(k) + e(k) \quad (3.122)$$

Pošto šum $e(k)$ direktno učestvuje u formulisanju ove jednačine, ovakav model se zove i model greške jednačine (eng. Equation Error model ili skraćeno EE model). U navedenom EE modelu greške jednačine $e(k)$ je slučajne prirode i modelovana je kao beli šum. Član $A(q)y(k)$ predstavlja AR deo modela, koji iskazuje činjenicu da izlaz u trenutku k , $y(k)$, zavisi od prethodnih vrednosti izlaza, koje su definisane polinomom $A(q)$, dok signal $u(k)$ predstavlja spoljašnji ulaz ili pobudu. Zaista, polinomijalna jednačina se može prikazati u formi diferentne jednačine

$$\left[1 + \sum_{i=1}^{n_a} a_i q^{-1} \right] y(k) = \sum_{i=1}^{n_b} b_i q^{-i} u(k) + e(k) \quad (3.123)$$

odakle sledi

$$y(k) = - \sum_{i=1}^{n_a} a_i q^{-1} y(k) + \sum_{i=1}^{n_b} b_i q^{-i} u(k) + e(k) \quad (3.124)$$

ili imajući u vidu značenje operatora kašnjenja q^{-1} ,

$$y(k) = - \sum_{i=1}^{n_a} a_i y(k-i) + \sum_{i=1}^{n_b} b_i u(k-i) + e(k) \quad (3.125)$$

Uvodeći vektor kolonu ulazno-izlaznih podataka (regresioni vektor) [2,5]

$$X^T(k-1) = \{-y(k-1) - y(k-2) \dots - y(k-n_a) u(i-1) u(i-2) \dots u(k-n_b)\} \quad (3.126)$$

i vektor kolonu parametara

$$\theta^T = \{ -a_1 - a_2 \dots - a_{n_a} b_1 b_2 \dots b_{n_b} \} \quad (3.127)$$

model se može prikazati u formi linearne regresije[9,78]

$$y(k) = X^T(k-1)\theta + e(k) \quad (3.128)$$

Alternativno se izlaz modela $y(k)$ može prikazati u polinomijalnom obliku

$$y(k) = \frac{B(q)}{A(q)}u(k) + \frac{1}{A(q)}e(k) \quad (3.129)$$

odnosno

$$y(k) = G(q)u(k) + H(q)e(k) \quad (3.130)$$

gde su odgovarajuće funkcije prenosa

$$G(q) = \frac{B(q)}{A(q)}, \quad H(q) = \frac{1}{A(q)} \quad (3.131)$$

Primetimo da je korelisani poremećaj na izlazu modela

$$v(k) = H(q)e(k) = \frac{1}{A(q)}e(k) \quad (3.132)$$

modelovan autoregresionim modelom

3.3.3 ARMAX model (Autoregresioni model pomičnih sredina sa spoljašnjom pobudom)

Ovaj model predstavlja specijalan slučaj opšte strukture u kojoj je $D(q) = F(q) = 1$, pri čemu su $A(q)$, $B(q)$ i $C(q)$ proizvoljno izabrani polinomi, tj.

$$A(q)y(k) = B(q)u(k) + C(q)e(k) \quad (3.133)$$

ili ekvivalentno

$$y(k) = \frac{B(q)}{A(q)}u(k) + \frac{C(q)}{A(q)}e(k) \quad (3.134)$$

odnosno

$$y(k) = G(q)u(k) + H(q)e(k) = G(q)u(k) + v(k) = y_m(k) + v(k) \quad (3.135)$$

gde su odgovarajuće funkcije prenosa

$$G(q) = \frac{B(q)}{A(q)}; \quad H(q) = \frac{C(q)}{A(q)} \quad (3.136)$$

ARMAX model spada u klasu modela tipa greške jednačine (EE modeli), kod kojih je greška jednačine modelovana procesom pomičnih sredina $C(q)e(k)$ (eng. Moving Average ili skraćeno MA-proces). Dakle poremećaj

$$v(k) = H(q)e(k) = \frac{C(q)}{A(q)}e(k) \quad (3.137)$$

je modelovan kao korelisani ili obojeni šum koji je dobijen propuštanjem belog šuma $e(k)$ kroz ARMA sistem (filtar), tj. generisan je ARMA-procesom

$$A(q)v(k) = C(q)e(k) \quad (3.138)$$

gde prvi član $A(q)v(k)$ predstavlja AR deo a drugi član $C(q)e(k)$ MA deo opisanog ARMA-procesa. Poslednja relacija može se napisati i u obliku diferentne jednačine

$$v(k) = -\sum_{i=1}^{n_a} a_i v(k-i) + \sum_{i=1}^{n_c} c_i e(k-i) + e(k) \quad (3.139)$$

Poslednji izraz je dobijen tako što su polinomi $A(q)$ i $C(q)$ zamenjeni odgovarajućim izrazima, uz tumačenje q^{-1} kao operatora jediničnog kašnjenja.

Na sličan način, izlaz procesa $y_m(k)$ generisan je ARX modelom

$$y_m(k) = G(q)u(k) = \frac{B(q)}{A(q)}u(k) \quad (3.140)$$

odnosno polinomijalnom relacijom

$$A(q)y_m(n) = B(q)u(k) \quad (3.141)$$

odakle sledi diferentna jednačina za generisanje izlaza

$$y_m(k) = -\sum_{i=1}^{n_a} a_i y_m(k-i) + \sum_{i=1}^{n_b} b_i u(k-i) \quad (3.142)$$

Primetimo na kraju da se kompletan izlaz

$$y(k) = y_m(k) + v(k) = G(q)u(k) + H(q)e(k) = \frac{B(q)}{A(q)}u(k) + \frac{C(q)}{A(q)}e(k) \quad (3.143)$$

može prikazati u polinomijalnoj ARMAX formi

$$A(q)y(k) = B(q)u(k) + C(q)e(k) \quad (3.144)$$

kojoj odgovara diferencna jednačina

$$y(k) = -\sum_{i=1}^{n_a} a_i y(k-i) + \sum_{i=1}^{n_b} b_i u(k-i) + \sum_{i=1}^{n_c} c_i e(k-i) + e(k) \quad (3.145)$$

ili alternativna jednačina linearne regresije

$$y(k) = X^T(k-1)\theta + e(k) \quad (3.146)$$

gde je regresioni vektor

$$X^T(k-1) = \{-y(k-1) \dots -y(k-n_a) u(k-1) \dots u(k-n_b) e(k-1) \dots e(k-n_c)\} \quad (3.147)$$

a vektor parametara

$$\theta^T = \{a_1 a_2 \dots a_{n_a} b_1 b_2 \dots b_{n_b} c_1 c_2 \dots c_{n_c}\} \quad (3.148)$$

Problem nastaje što su u regresionom vektoru merljivi samo ulazno-izlazni podaci $\{u(i), y(i)\}$, dok beli šum $\{e(i)\}$ nije merljiv signal. Stoga je neophodno da se odbirci belog šuma u regresionom vektoru zamene nekom njihovom procenom, o čemu će više biti reči u poglavlju posvećenom parametarskoj identifikaciji sistema.

3.3.4 OE model (Model greške izlaza)

Ovaj model predstavlja specijalan slučaj opšte strukture modela u kojoj je $A(q) = C(q) = D(q) = 1$, a $B(q)$ i $F(q)$ predstavljaju proizvoljne polinome. Dakle, ova struktura ima formu

$$y(k) = \frac{B(q)}{F(q)} u(k) + e(k) \quad (3.149)$$

ili ako se uvede pomoćni deterministički signal

$$\omega(k) = \frac{B(q)}{F(q)} u(k) \quad (3.150)$$

data se struktura može zapisati kao

$$F(q)\omega(k) = B(q)u(k) \quad (3.151)$$

$$y(k) = \omega(k) + e(k) = y_m(k) + e(k) \quad (3.152)$$

Dobijeni izraz ukazuje na motiv da se ovakav model nazove model greške izlaza, pošto član $e(k)$ u drugoj relaciji ima smisao greške koja se superponira na

izlaz $\omega(k) = y_m(k)$ odgovarajućeg determinističkog sistema u cilju formiranja izlaza $y(k)$. OE model može se zapisati i u obliku

$$y(k) = G(q)u(k) + H(q)e(k) \quad (3.153)$$

gde su odgovarajuće funkcije prenosa

$$G(q) = \frac{B(q)}{F(q)}; \quad H(q) = 1 \quad (3.154)$$

Uočava se da funkcija prenosa procesa (sistema) $G(q)$ i funkcija prenosa poremećaja $H(q)$ nemaju zajedničke parametre, te se kaže da su model procesa i model poremećaja nezavisno parametrizovani.

Polinomijalna relacija kojom je generisan izlaz $\omega(k) = y_m(k)$ determinističkog dela sistema

$$F(q)\omega(k) = B(q)u(k) \quad (3.155)$$

može se zapisati u formi diferencne jednačine

$$\omega(k) = -\sum_{i=1}^{n_f} f_i \omega(k-i) + \sum_{i=1}^{n_b} b_i u(k-i) \quad (3.156)$$

pa se kompletan izlaz može prikazati kao

$$y(k) = \omega(k) + e(k) = -\sum_{i=1}^{n_f} f_i \omega(k-i) + \sum_{i=1}^{n_b} b_i u(k-i) + e(k) \quad (3.157)$$

Navedenoj diferencnoj jednačini odgovara jednačina linearne regresije

$$y(k) = X^T(k-1)\theta + e(k) \quad (3.158)$$

gde je regresioni vektor

$$\theta^T = \{ -f_1 - f_2 \dots - f_{n_f} \quad b_1 b_2 \dots b_{n_b} \} \quad (3.159)$$

Uočava se da se, za razliku od ARX modela koji je EE tipa, kod ovog OE modela u regresionom vektoru pojavljuju zakašnjeni odbirci izlaza ω determinističkog dela sistema, a ne kao kod EE modela zakašnjeni odbirci kompletnog izlaza y . To predstavlja osnovnu razliku između OE i EE tipova modela.

3.3.5 BJ model (Box-Jenkinsov model)

Ovaj model predstavlja specijalan slučaj opšte strukture modela u kojoj je $A(q) = 1$, tako da model ima oblik

$$y(k) = \frac{B(q)}{F(q)}u(k) + \frac{C(q)}{D(q)}e(k) \quad (3.160)$$

ili alternativno

$$y(k) = G(q)u(k) + H(q)e(k) = G(q)u(k) + v(k) \quad (3.161)$$

gde su odgovarajuće funkcije prenosa

$$G(q) = \frac{B(q)}{F(q)}, \quad H(q) = \frac{C(q)}{D(q)}e(k) \quad (3.162)$$

Uočava se da se poremećaj

$$v(k) = H(q)e(k) = \frac{C(q)}{D(q)}e(k) \quad (3.163)$$

dobija propuštanjem belog šuma $e(k)$ kroz odgovarajuću funkciju prenosa $H(q) = C(q)/D(q)$, odnosno da je generisan ARMA-procesom

$$D(q)v(k) = C(q)e(k) \quad (3.164)$$

gde član $D(q)v(k)$ predstavlja AR-proces a član $C(q)e(k)$ opisuje MA-proces u navedenom ARMA-procesu. Ovakva ARMA struktura može se zapisati u obliku diferentne jednačine

$$v(k) = -\sum_{i=1}^{n_d} d_i v(k-1) + \sum_{i=1}^{n_c} c_i e(k-1) + e(k) \quad (3.165)$$

Poslednji izraz je dobijen zamenom odgovarajućih izraza za polinome $D(q)$ i $C(q)$ i korišćenjem definicionog izraza za operator kašnjenja, tj. $q^{-i}x(k) = x(k-i)$. Primitimo, takođe, da funkcija prenosa procesa (sistema) $G(q)$ i funkcija prenosa aditivnog poremećaja $v(k)$ na izlazu sistema $y(k)$ nemaju zajedničkih parametara, tako da su ove funkcije prenosa nezavisno parametrizovane.

Uvodeći oznaku za izlaz determinističkog dela procesa

$$\omega(k) = y_m(k) = G(q)u(k) = \frac{B(q)}{F(q)}u(k) \quad (3.166)$$

ovaj izlaz se može generisati na osnovu polinomijalne jednačine

$$F(q)\omega(k) = B(k)u(k) \quad (3.167)$$

odnosno odgovarajuće diferentne jednačine

$$\omega(k) = -\sum_{i=1}^{n_f} f_i - \omega(k-i) + \sum_{i=1}^{n_b} b_i u(k-i) \quad (3.168)$$

Imajući u vidu da je poremećaj generisan ARMA procesom

$$v(k) = -\sum_{i=1}^{n_d} d_i v(k-i) + \sum_{i=1}^{n_c} c_i e(k-i) + e(k) \quad (3.169)$$

Kompletan ulaz definisan je diferentnom jednačinom

$$\begin{aligned} y(k) &= \omega(k) + v(k) \\ &= -\sum_{i=1}^{n_f} f_i \omega(k-i) + \sum_{i=1}^{n_b} b_i u(k-i) - \sum_{i=1}^{n_d} d_i v(k-i) + \sum_{i=1}^{n_c} c_i e(k-i) + e(k) \end{aligned} \quad (3.170)$$

kojoj odgovara jednačina linearne regresije

$$y(k) = X^T(k-1)\theta + e(k) \quad (3.171)$$

gde je regresioni vektor

$$X^T(k-1) = \left\{ -\omega(k-1) \dots -\omega(k-n_f) u(k-1) \dots u(k-n_b) -v(k-1) \dots -v(k-n_d) e(k-1) \dots e(k-n_c) \right\} \quad (3.172)$$

a vektor parametara

$$\theta^T = \left\{ -f_1 \dots -f_{n_f} b_1 \dots b_{n_b} -d_1 \dots -d_{n_d} c_1 \dots c_{n_c} \right\} \quad (3.173)$$

Primetimo da je i ovo model OE tipa, pošto se u regresionom vektoru pojavljuju zakašnjeni odbirci izlaza ω determinističkog dela sistema, a ne zakašnjeni odbirci kompletnog ulaza y , što je slučaj sa modelima EE tipa. Takođe, u regresionom vektoru su poznati samo deterministički izlaz ω (računa se na osnovu navedene diferencne jednačine), izmereni ulaz u , dok se poremećaj v može proceniti na osnovu merenja izlaza y i izračunatog determinističkog izlaza ω ($y(k) = \omega(k) + v(k)$), pri čemu je beli šum e nemerljivi signal, te se njegovi odbirci moraju na neki način proceniti. O ovom problemu biće više reči u poglavlju posvećenom parametarskoj identifikaciji.

Postupak svođenja opšte strukture (reprezentacije) modela na specifičan oblik, na osnovu pogodnog izbora odgovarajućih polinoma koji definišu opštu strukturu modela, definisan je u sledećoj tabeli. U navedenoj tabeli oznaka “1” označava da je odgovarajući polinom fiksiran na jedinicu (polinom nultog reda čiji je slobodan član jednak 1), dok simbol “x” označava potpuno slobodan izbor odgovarajućeg polinoma.

Polinom Struktura	A	B	C	D	F
OPŠTA	x	x	x	x	x
FIR	1	x	1	1	1
ARX	x	x	1	1	1
ARMAX	x	x	x	1	1
OE	1	x	1	1	x
BJ	1	x	x	x	x

Tabela3.1. Različite strukture (reprezentacije) modela procesa, na osnovu izbora odgovarajućih polinoma u opštoj strukturi modela.

Osnovna razlika između različitih struktura nalazi se u načinu na koji je modelovan aditivni poremećaj (šum) na izlazu sistema. Pri tome, AR i ARMAX strukture pripadaju klasi struktura tipa greške jednačine, tzv. EE strukture, dok OE i BJ strukture pripadaju klasi struktura tipa greške izlaza, tzv. OE strukture. Kod struktura tipa EE model procesa (sistem) i model poremećaja (šuma) sadrže zajedničke parametre, te se ne mogu nezavisno parametrizovati, dok kod strukture OE tipa ova dva modela nemaju zajedničkih parametara, te se mogu nezavisno parametrizovati.

4. PARAMETARSKA IDENTIFIKACIJA

Pod pojmom parametarske identifikacije podrazumeva se estimacija (procena) konačnog broja nepoznatih parametara u unapred usvojenoj strukturi modela sistema. Izrada modela sistema nikad nije sama po sebi cilj, već se uvek ima u vidu namena dobijenog modela. Parametarski modeli se koriste za analizu procesa, njihovu simulaciju i projektovanje kontrolera [6,10,21,37,38,59,61,88,89,95]. Zadatak analize je da se predvidi ponašanje procesa pod određenim okolnostima i analiza se obično zasniva na računarskoj simulaciji. Sa druge strane, zadatak sinteze je da se projektuje kontroler, koji treba da obezbedi da se proces ponaša u skladu sa postavljenim zahtevima. Izvršenje bilo kog od ovih zadataka pretpostavlja postojanje adekvatnog modela procesa koji je u stanju da uspešno predvidi ponašanje procesa pod zadatim uslovima rada. Sa ovakvog stanovišta prirodan kriterijum za ocenu valjanosti modela predstavlja sposobnost modela da predviđa ponašanje procesa pod zadatim okolnostima. Ovakav verbalni kriterijum potrebno je prevesti na jezik matematike, kako bi se definisao matematički kriterijum ili mera kojom će se numerički iskazati kapacitet predikcije parametarskog modela. Jedan mogući pristup sastoji se u izračunavanju budućih izlaza (odziva) procesa za zadate uslove rada (zadate početne uslove, pobudne ulazne signale i poremećaje) i njihovo poređenje sa stvarnim izmerenim izlazima procesa, koji su dobijeni pod istim okolnostima. Ukoliko su greške između izmerenih izlaza i njihovih predikcija dovoljno male, takav model se može smatrati adekvatnim. Metode identifikacije koje se zasnivaju na ovakvom principu nazivaju se **metode greške predikcije** (engleski Prediction Error Methods ili skraćeno PEMs), pošto je greška predikcije, tj. razlika između izmerene i prediktovane vrednosti, usvojena za meru ili kriterijum valjanosti, odnosno kriterijum za ocenu kvaliteta ponašanje modela u odnosu na realan system [14,28,63,68,89]. Da bismo formalizovali izloženi koncept, uvedimo sledeće oznake. Neka $y(k)$ označava izmereni izlaz procesa u diskretnom trenutku vremena $t_k = kT$, $k \in \mathbb{Z}$, gde je T perioda odabiranja, odnosno vremenski interval između dva sukcesivna ekvidistantna diskretna merenja. Pretpostavimo, takođe, da je celokupna prošlost sistema do trenutka k poznata na osnovu raspoložive merne informacije $y(i)$, $i \leq k-1$, te da želimo da predvidimo buduću izlaz sistema $y(k+1)$ u trenutku $k+1$. Ovakva predikcija naziva se jednokoračna predikcija izlaza unapred (engleski one-step ahead prediction) i označava sa $\hat{y}(k+1|k)$. Ako se u trenutku $k+1$ izmeri izlaz i ovo merenje označi sa $y(k+1)$, tada je greška predikcije (engleski prediction error)

$$v(k+1|k) = y(k+1) - \hat{y}(k+1|k) \quad (4.1)$$

Na osnovu koncepta greške predikcije, model sistema je dobar ukoliko je ova greška mala. Međutim, mala greška predikcije u jednom trenutku ne implicira isto tako malu vrednost u nekom narednom trenutku. Zato je bolje kao kriterijum koristiti sekvencu (niz) grešaka predikcija, a da bi se sprečilo da se efekti pozitivnih i

negativnih grešaka potru. Uobičajeno je da se posmatra sekvenca kvadrata grešaka predikcija. Prirodno da se kao empirijski kriterijum za ocenu valjanosti modela usvoji aritmetička sredina grešaka predikcija na posmatranom intervalu $k = 1, 2, \dots, N$ [28,63,68,89].

$$J_N = \frac{1}{N} \sum_{k=1}^N v^2(k|k-1) \quad (4.2)$$

U skladu sa usvojenim kriterijumom, model ima dobar kapacitet predikcije ukoliko je brojna vrednost kriterijuma J_N dovoljno mala pozitivna veličina. Najbolji model je onaj koji minimizira usvojeni kriterijum. Na taj način se zadatak parametarske identifikacije može postaviti kao optimizacioni problem.

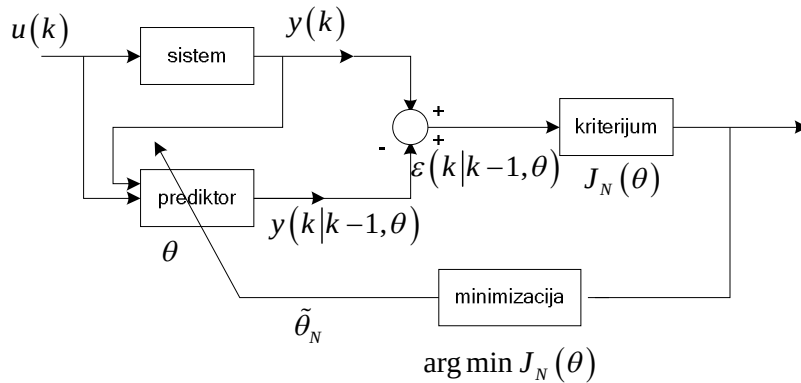
Kao što je istaknuto u prethodnom poglavlju, parametarski model se sastoji od odgovarajuće strukture (na primer, ARX, ARMAX, itd.), skupa celobrojnih promenljivih (označavaju redove odgovarajućih polinoma u usvojenoj strukturi, eventualno vreme kašnjenja, itd.) i skupa realnih promenljivih (označavaju nepoznate parametre u usvojenoj strukturi modela). Ako sve parametre skupimo u vektor parametara θ , tada je optimalan model u okviru usvojene strukture modela onaj čiji vektor parametara $\theta = \hat{\theta}_N$ minimizira usvojeni kriterijum za ocenu valjanosti modela

$$j_N(\theta) = \frac{1}{N} \sum_{k=1}^N v^2(k|k-1, \theta) \quad (4.3)$$

odnosno

$$\hat{\theta}_N = \arg \min_{\theta \in D_M} J_N(\theta) \quad (4.4)$$

Ovde je eksplicitno naznačeno da greške predikcije, a samim tim i kriterijum valjanosti, zavise od vrednosti parametara u usvojenoj strukturi modela, dok D_M označava skup svih mogućih vektora parametara za usvojenu strukturu modela, tj. usvojenu parametrizaciju modela (na primer, D_M može biti skup svih vektora parametara koji konstituišu ARX model drugog reda). Metode greške predikcije (PEMs) zasnivaju se na poslednje dve relacije. Blok šema parametarske identifikacije zasnovana na principu greške predikcije data je na narednoj slici.



Slika 4.1: Blok šema parametarske identifikacije na bazi greške predikcije

Problemu procenjivanja nepoznatog vektora parametara u usvojenoj strukturi modela sistema može se prići i sa generalnog stanovišta minimizacije usrednjenog indeksa performanse

$$J(\theta) = E\{F[v(i, \theta)]\} \quad (4.5a)$$

ili empirijskoj indeks performanse [63,68,72,84,89]

$$J_i(\theta) = \frac{1}{i} \sum_{k=1}^i F[v(k, \theta)] \quad (4.5b)$$

gde je $F(\cdot)$ kriterijumska funkcija, funkcija gubitaka, cene ili rizika, a $v(i, \theta)$ greška jednačine ili rezidual, definisan kao razlika izlaza sistema i njegove optimalne srednje-kvadratne predikcije $\hat{y}(i, \theta) = E\{y(i|1)\}$. Za dovoljno veliko i empirijski kriterijum (indeks performanse) (4.5b) konvergiraće na osnovu zakona velikih brojeva teorijskom kriterijumu optimizacije (4.5a). Procena nepoznatog vektora parametara dobija se iz uslova minimuma kriterijuma (4.5)

$$\nabla_{\theta} J(\theta) = E\{\nabla_{\theta} F[v(i, \theta)]\} = 0 \quad (4.6)$$

Ako se uvede oznaka $\psi(z) = F'(z)$, uslov (4.6) može se napisati u obliku

$$\nabla_{\theta} J(\theta) = E\{\psi[v(i, \theta)] \nabla_{\theta} v(i, \theta)\} = 0. \quad (4.7)$$

Ukoliko je $F(\cdot)$ parna funkcija, njen izvod $\psi(\cdot)$ je neparna funkcija i ukoliko je raspodela poremećaja $\xi(i)$ simetrična funkcija, a slučajne veličine $v(i, \theta)$ i $\nabla v(i, \theta)$ su nekorelisane, rešenje nelinearnog sistema jednačina (4.7) je $\theta = \theta^*$, gde je θ^* tačna vrednost nepoznatog vektora parametara θ koji se procenjuje i koja odgovara minimumu kriterijuma. Zaista, ako su slučajne veličine $v(i, \theta)$ i $\nabla v(i, \theta)$ statističke nekorelisane i $E\{\psi(\xi(i))\} = 0$ tada je

$$\nabla_{\theta} J(\theta^*) = E\{\psi[v(i, \theta^*)] \nabla v(i, \theta^*)\} = E\{\psi(\xi(i))\} E\{\nabla v(i, \theta^*)\} = 0. \quad (4.8)$$

Poslednji izraz jednak je nuli, pošto je $v(i, \theta^*) = y(i) - \hat{y}(i, \theta^*) = \xi(i)$, s obzirom da $\hat{y}(i, \theta^*) = E\{y(i)|i-1\}$ određuje odziv determinističkog dela sistema koji se razlikuje od stvarnog odziva sistema za aditivni šum merenja $\xi(i)$, pri čemu je $E\{\psi(\xi(i))\} = 0$ na osnovu usvojenih pretpostavki (slika 4.1). Ako je, uz to, $F(\cdot)$ konveksna funkcija njen izvod je monotona funkcija i dobijeno rešenje je jedinstveno.

Relacija (4.8) pokazuje da je uslov optimalnosti zadovoljen za proizvoljnu parnu funkciju gubitaka $F(\cdot)$ u (4.5), odnosno da je optimalno rešenje $\theta = \theta^*$ praktično invarijantno na izbor kriterijumske funkcije. Ovakav rezultat imao bi presudnu ulogu u rešavanju problema identifikacije kada bi se optimalno rešenje moglo naći direktno. Međutim, nelinearni sistem jednačina (4.7) generalno se ne može rešiti analitički, već njegovo rešavanje zahteva primenu metoda numeričke

analize [54,55,74,92]. Naime, na osnovu N sekvenci merenja ulaza i izlaza sistema $\{u(i), y(i) | 1 \leq i \leq N\}$ generiše se po nekom algoritmu procena $\theta(N)$, koja kad N neograničeno konvergira ka optimalnom rešenju. Ovakvi algoritmi nazivaju se paketima (odgovarajući engleski termin je batch ili off-line algoritmi). Alternativno, kriterijum optimizacije može se minimizirati iterativno kako evoluira vreme, tako da se procena u i -tom trenutku $\theta(i)$ dobija kao funkcija prethodne procene $\theta(i-1)$ i novo dobijene informacije o ulazu i izlazu sistema, $u(i)$ i $y(i)$, respektivno, tj.

$$\theta(i) = f(\theta(i-1), y(i), u(i)) \quad (4.9)$$

Ovakvi algoritmi nazivaju se rekurzivnim ili algoritmima u realnom vremenu (engleski - real-time ili on-line).

Na osnovu rečenog, ključni korak u ovako formulisanom zadatku identifikacije predstavlja izbor funkcije gubitaka $F(\cdot)$ u indeksu performanse (4.5). Navedenom pitanju u literaturi je posvećeno dosta pažnje [89]. Generalni zaključak svih takvih razmatranja bio bi da je izbor kriterijumske funkcije $F(\cdot)$ u velikoj meri subjektivan i predstavlja kompromis između težnje da se, s jedne strane, u kriterijum što više ugrade fizički uslovi rada i namene sistema, a s druge strane da dobijena rešenja budu što jednostavnija analitički. U inženjerskoj praksi najčešće je korišćen kvadratni kriterijum $F(z) = z^2$, pre svega zbog uprošćavanja potrebnog matematičkog aparata. Naime, u slučaju kvadratnog kriterijuma sistem jednačina (4.7) može da se reši eksplicitno, čime se dobija analitički izraz koji definiše nerekurzivni algoritam najmanjih kvadrata (odgovarajući engleski termin je Least Squares ili skraćeno LS algoritam) [28,63,68,89]. Ako se u tako dobijenoj relaciji matematička očekivanja aproksimiraju odgovarajućim aritmetičkim sredinama i primeni teorema o inverziji matrica dolazi se do rekurzivne forme algoritma najmanjih kvadrata u realnom vremenu [28,63,68,89]. O izvođenju ovih algoritama biće više reči u nastavku ovog poglavlja.

Ukoliko su poremećaji $\xi(i)$ raspodeljeni po Gausu algoritam najmanjih kvadrata je optimalan, u smislu minimalne varijanse greške estimacije [25,48,86]. Ta činjenica, kao i ispunjenje uslova centralno-granične teoreme matematičke statistike u velikom broju praktičnih situacija, uz već ranije navedene razloge vezane za matematičku jednostavnost i teoretsku invarijantnost optimalnog rešenja (4.6) ili (4.7) na izbor kriterijumske funkcije, predstavljali su jak motiv za široko korišćenje algoritma najmanjih kvadrata u inženjerskoj praksi. Konstatovano je, takođe, da algoritam najmanjih kvadrata daje zadovoljavajuće procene i u mnogobrojnim situacijama kada raspodela poremećaja nije gasovska [4,14,63]. Međutim, u izvesnim situacijama karakterisanim malim brojem mernih podataka i pojavom sporadičnih realizacija poremećaja velikog intenziteta (režim “outlier”-a, “bad data” ili anomalnih merenja) kvalitet procena najmanjih kvadrata je nezadovoljavajući [4,12,27,33,34,42,66,75,76,94]. Relativna neosetljivost procena na poremećaje navedenog tipa može se postići pogodnim izborom oblika kriterijumske funkcije. Osnovna ideja sastoji se u zameni kvadratne funkcije gubitaka funkcijom $F(z)$ koja je za male vrednosti argumenta z bliska kvadratnoj, ali je za veće njegove vrednosti njena vrednost manja od z^2 . Jedna mogućnost je da se usvoji [4,10,76]

$$F(z) = \frac{z^2}{1 + az^2} \quad (4.10)$$

Druge dve mogućnosti su [4,40,76]

$$F(z) = \begin{cases} \frac{1}{2}z^2 & |z| \leq a \\ \frac{1}{2}a^2 & |z| > a \end{cases}; \quad F(z) = \begin{cases} \frac{1}{2}z^2 & |z| \leq a \\ a|z| - \frac{1}{2}a^2 & |z| > a \end{cases}; \quad (4.11)$$

koje formalno odgovaraju r -podešenoj sredini i Winzorijanskoj sredini, respektivno. Uopšte, postoji veliki broj različitih mogućnosti, koje su razmatrane u literaturi, a zasnovane su uglavnom na praktičnom iskustvu ili teoriji robusne estimacije [4,9,40,76]. Na osnovu do sada izloženog, kao i blok dijagrama na slici 4.1, problem rekursivne parametarske identifikacije može se raščlaniti u tri ključna zadatka: 1) izbor kriterijuma optimalnosti (4.5); 2) izbor algoritma za predikciju izlaza sistema $\hat{y}(i, \theta)$ u (4.1); 3) izbor algoritma parametarske identifikacije (4.9).

U radu će biti razmatrana optimalna predikcija izlaza u smislu minimuma srednje-kvadratnog kriterijuma (engleski Mean Square Error ili skraćeno MSE kriterijum). Dalje će biti pokazano da se kriterijumska funkcija može izabrati na optimalan način, u određenom smislu, zavisno od raspoložive apriorne informacije o statističkim karakteristikama poremećaja u sistemu. Ukoliko je ova informacija zadata u obliku klase funkcija raspodele verovatnoće poremećaja kojoj pripada i njegova stvarna, ali nepoznata, raspodela, biće pokazano da se kriterijumska funkcija može izabrati na optimalan način, u smislu minimalno moguće greške estimacije parametara u zadatim uslovima stohastičkog okruženja. Predloženi pristup zasniva se na analizi konvergencije i brzine konvergencije procena parametara i primeni Kramer-Raove teoreme. Kada se govori generalno o analizi konvergencije stohastičkih rekursivnih ili rekurentnih procedura parametarske estimacije u literaturi su široko rasprostranjena dva metodološka pristupa. Prvi od njih zasniva se na pridruživanju obične diferencijalne jednačine razmatranoj stohastičkoj rekursivnoj proceduri i primeni teorije stabilnosti Ljapunova za analizu stabilnosti rešenja dobijene obične diferencijalne jednačine. Ovakav pristup skraćeno se označava sa ODE, na engleskom jeziku, Ordinary Differential Equation [63]. Drugi pristup zasniva se na primeni teorije martingala [32,62,70]. Ovakav pristup sastoji se u analizi stohastičke funkcije Ljapunova, koja predstavlja pogodno odabranu kvadratnu formu slučajne greške estimacije parametara i primene osnovnih rezultata o konvergenciji martingala na relevantne slučajne sekvence koji zadovoljavaju definiciju martingala. U izvesnim situacijama moguća je i kombinacija ova dva pristupa. Sa druge strane, Kramer-Raova teorema omogućava komparaciju sa Kramer-Raovom donjom međom, koja definiše najveću moguću tačnost procena nepoznatih parametara. Ako se komparacijom ustanovi da je odgovarajući relevantni pokazatelj greške estimacije veći od Kramer-Raove donje međe, moguće je pod određenim uslovima izvršiti modifikaciju razmatrane rekursivne stohastičke procedure u cilju postizanja maksimalno moguće preciznosti procene parametara pod navedenim uslovima rada, odnosno maksimalno moguće asimptotske brzine konvergencije procene parametara. Ovakav pristup rezultuje u sintezu optimalnih rekursivnih stohastičkih procedura na unapred specificiranoj klasi raspodela poremećaja, kojoj pripada i njegova stvarna, ali nepoznata, funkcija raspodele verovatnoće. Konačno, kao rekursivne stohastičke

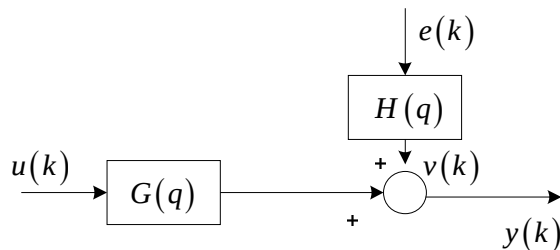
procedure parametarske estimacije biće razmatrane klase algoritama tipa Newton-Raphsonove stohastičke procedure i njene aproksimacije koja se naziva Gauss-Newtonova stohastička procedura, kao i klasa algoritama tipa stohastičkog gradijenta ili stohastičke aproksimacije [50,51,52].

4.1. MODELI PREDIKCIJE: OPTIMALNA SREDNJE – KVADRATNA JEDNOKORAČNA PREDIKCIJA

Izračunavanje greške predikcije $v(k|k-1)$ u (4.1) zahteva izračunavanje jednokoračne predikcije izlaza unapred $\hat{y}(k|k-1)$ i to prvo na primeru standardne (generalne) strukture parametarskog modela, a dobijeni rezultat se lako može prilagoditi na slučajeve specifičnih parametarskih struktura, kao što su FIR, ARX, ARMAX, OE i BJ strukture, koje su opisane u trećem poglavlju.

4.1.1. Model predikcije za generalni parametarski model

Standardni (generalni) model razmatran je u prethodnom poglavlju i prikazan je blok dijagramom na sledećoj slici. Prikazani model je generalan i može biti i parametarski i neparametarski, kao što je pokazano u trećem poglavlju.



Slika 4.2: Standardni (generalni) parametarski model sistema

Na prikazanoj slici $u(k)$, $y(k)$ i $v(k)$ označavaju ulaz, izlaz i poremećaj u diskretnom trenutku k , dok je $e(k)$ beli šum nulte srednje vrednosti. Dalje se pretpostavlja da se raspolaže sa ulazno-izlaznom sekvencom merenja $\{u(\ell), y(\ell)\}$, $\ell = k-1, k-2, \dots$, kao i da su za momenat poznate funkcije prenosa (polinomi) $G(q)$ i $H(q)$. Pod navedenim uslovima mogu se izračunati veličine

$$v(\ell) = y(\ell) - G(q)u(\ell), \quad \ell = k-1, k-2, \dots \quad (4.12)$$

te se ove veličine mogu smatrati poznatim. Takođe, ako je funkcija prenosa (polinom) $G(q)$ striktno kauzalna, tada je njen slobodni član $g_o \equiv 0$, pa je polinom

$$G(q) = \sum_{\ell=1}^{\infty} g_{\ell} q^{-\ell} \quad (4.13)$$

odakle sledi da

$$G(q)u(k) = \sum_{\ell=1}^{\infty} g_{\ell} q^{-\ell} u(k) = \sum_{\ell=1}^{\infty} g_{\ell} u(k-\ell) \quad (4.14)$$

zavisi samo od ulaza $u(i)$, $i = k-1, k-2, \dots$, tako da je i ovaj član poznat. Polazeći dalje od izraza za izlaz sistema (slika 2.2)

$$y(k) = G(q)u(k) + H(q)e(k) = G(q)u(k) + v(k) \quad (4.15)$$

jednokoračna predikcija izlaza može se definisati sa

$$\hat{y}(k|k-1) = G(q)u(k) + \hat{v}(k|k-1) \quad (4.16)$$

Dakle, rešenje postavljenog jednokoračnog prediktora zahteva da se izračuna jednokoračna predikcija $\hat{v}(k|k-1)$ nemerljivog šuma $v(k)$. Da bi ovaj problem bio rešiv funkcija prenosa (filter) ili polinoma šuma $H(q)$ mora da zadovoljava sledeća tri uslova:

1. $H(q)$ je stabilna funkcija prenosa.
2. Inverzna funkcija prenosa $H^{-1}(q) = 1/H(q)$ je stabilna.
3. $H(q)$ je monik funkcija prenosa, odnosno $h_0 = 1$ odakle sledi

$$H(q) = 1 + \sum_{\ell=1}^{\infty} h_{\ell} q^{-\ell} \quad (4.17)$$

Na osnovu trećeg uslova, korelisani (obojeni) aditivni šum na izlazu sistema je

$$v(k) = H(q)e(k) = e(k) + \sum_{\ell=1}^{\infty} h(\ell)e(k-\ell) \quad (4.18)$$

Sada se postavlja problem kako generisati predikciju $v(k)$, ako su izračunate prethodne vrednosti $v(\ell)$, $\ell \leq k-1$. Najbolju (optimalnu srednje-kvadratnu) predikciju prirodno predstavlja očekivana (srednja) vrednost $v(k)$, tj. uslovno matematičko očekivanje [25, 41, 48, 60, 77, 86].

$$\hat{v}(k|k-1) = \bar{E}\{v(k)\} = \bar{E}\{e(k)\} + \bar{E}\left\{\sum_{\ell=1}^{\infty} h(\ell)e(k-\ell)\right\} \quad (4.19)$$

gde je $\bar{E}\{\cdot\}$ uslovno matematičko očekivanje, tj. $\bar{E}\{v(k)\} = E\{v(k)|k-1\}$. Pošto je na osnovu pretpostavke $\{e(k)\}$ beli šum nulte srednje vrednosti, to je $\bar{E}\{e(k)\} = 0$, dok je drugi sabirak deterministička veličina, pošto je $v(\ell)$, a samim time i $e(\ell)$, poznato za $\ell \leq k-1$. Na taj način se dobija

$$\hat{v}(k|k-1) = \sum_{\ell=1}^{\infty} h_{\ell} e(k-\ell) = [H(q) - 1]e(k) \quad (4.20)$$

Imajući u vidu da je na osnovu slike 4.2

$$v(k) = H(q)e(k) \Rightarrow e(k) = H^{-1}(q)v(k) = v(k)/H(q) \quad (4.21)$$

konačno se dobija traženi izraz za predikciju poremećaja

$$\hat{v}(k|k-1) = \frac{H(q)-1}{H(q)}v(k) = \left[1 - \frac{1}{H(q)}\right]v(k) \quad (4.22)$$

Da bi se izračunao prediktor $\hat{v}(k|k-1)$ neophodne su samo vrednosti $v(\ell)$, $\ell \leq k-1$, koje su generisane na osnovu relacije (4.12)

Zamenom prediktora poremećaja (šuma) (4.22) u izraz za predikciju izlaza (4.16), kao i korišćenjem (4.12), dobija se relacija

$$\begin{aligned} \hat{y}(k|k-1) &= G(q)u(k) + [1 - H^{-1}(q)]v(k) \\ &= G(q)u(k) + [1 - H^{-1}(q)][y(k) - G(q)u(k)] \end{aligned} \quad (4.23)$$

na osnovu koje sledi

$$\hat{y}(k|k-1) = H^{-1}(q)G(q)u(k) + [1 - H^{-1}(q)]y(k) \quad (4.24)$$

ili u alternativnom obliku

$$H(q)\hat{y}(k|k-1) = G(q)u(k) + [H(q)-1]y(k) \quad (4.25)$$

Zamenom izvedenog prediktora izlaza u definicioni izraz za grešku predikcije dalje se dobija

$$\varepsilon(k|k-1) = y(k) - \hat{y}(k|k-1) = -H^{-1}(q)G(q)u(k) + H^{-1}(q)y(k) \quad (4.26a)$$

Primetimo da veličine $\varepsilon(k|k-1)$ u (4.26) odgovara veličini $v(k|k-1)$ u relaciji (4.1). Ovde je korišćena različita oznaka, da bi se rezidual jasno razlikovao od šuma merenja. Takođe, beli šum $e(i)$ odgovara varijabli $\xi(i)$ u (4.8).

Prilikom prethodnih izvođenja pretpostavljeno je da su funkcije prenosa $G(q)$ i $H(q)$ poznate, što inače nije ispunjeno, pošto je model poznat sa tačnošću do nepoznatih parametara, koje treba estimirati na bazi raspoloživih merenja ulaza i izlaza. Dakle, funkcije prenosa $G(q)$ i $H(q)$ nisu poznate, ali su parametrizovane sa nekim vektorom parametara θ , tako da je u suštini greška predikcije funkcija ovog vektora parametara, odnosno

$$\varepsilon(k|k-1, \theta) = H^{-1}(q, \theta)[y(k) - G(q, \theta)u(k)] \quad (4.26b)$$

Varirajući vektor parametara unutar skupa D_M može se izračunati greška predikcije ε i rezultujuća vrednost kriterijuma valjanosti modela (4.2) ili (4.5). Izvedeni prediktor je primenljiv na bilo koji model koji ima standardnu (generalnu) strukturu prikazanu gornjim blok dijagramom. Dobijeni prediktor može se primeniti na specifične parametrizovane modele, koji su razmatrani u prethodnom poglavlju, kao što su FIR, ARX i ARMAX strukture, a koje predstavljaju posebne slučajeve razmatrane standardne (generalne) reprezentacije modela sistema.

4.1.1. Model predikcije za FIR reprezentaciju sistema

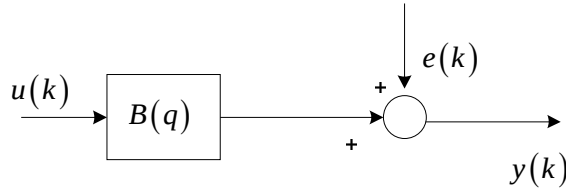
FIR struktura modela sistema predstavlja specijalan slučaj prethodno razmatrane standardne (generalne) strukture, kada se usvoji

$$G(q, \theta) = B(q) = \sum_{i=1}^{n_b} b_i q^{-i}, \quad H(q, \theta) = 1 \quad (4.27)$$

tako da je vektor parametara u ovako parametrizovanom modelu

$$\theta = \{b_1 \ b_2 \dots b_{n_b}\}^T. \quad (4.28)$$

Ovakva struktura prikazana je blok dijagramom na narednoj slici 4.3



Slika 4.3: FIR struktura modela sistema

Na taj način, izlazni signal može se prikazati relacijom

$$y(k) = B(q)u(k) + e(k) \quad (4.29)$$

gde je $u(k)$ pobudni signal u k -tom diskretnom trenutku, a $e(k)$ je aditivni šum na izlazu sistema, koji predstavlja beli šum nulte srednje vrednosti. Koristeći prethodno dobijeni izraz za prediktor izlaza u slučaju standardne reprezentacije modela sistema (4.23) i uvrštavajući izraze za funkcije prenosa u FIR strukturi $G(q) = B(q)$, $H(q) = 1$ dobija se za prediktor izlaza FIR strukture

$$\hat{y}(k|k-1, \theta) = B(q)u(k) = \sum_{i=1}^{n_b} b_i q^{-i} u(k) = \sum_{i=1}^{n_b} b_i u(k-i) \quad (4.30)$$

Ako se merenja ulaza $u(\ell)$, $\ell \leq k-1$ sakupe u regresioni vektor

$$\varphi^T(k) = \{u(k-1), u(k-2), \dots, u(k-n_b)\} \quad (4.31)$$

tada se prediktor izlaza može zapisati u formi linearne regresije

$$\hat{y}(k|k-1, \theta) = \varphi^T(k) \theta = \theta^T \varphi(k) \quad (4.32)$$

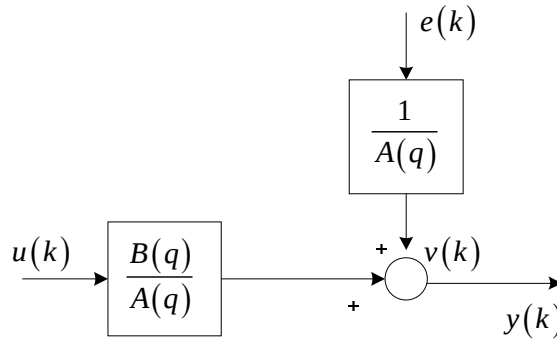
Naziv linearna regresija potiče od činjenice da je relacija za predikciju izlaza linearna funkcija nepoznatog vektora parametara θ .

4.1.2. Model predikcije za ARX reprezentacije modela sistema

ARX struktura predstavlja poseban oblik standardne strukture na slici 4.2 u kojoj je usvojeno

$$G(q, \theta) = \frac{B(q)}{A(q)}, \quad A(q) = 1 + \sum_{i=1}^{n_a} a_i q^{-i}, \quad B(q) = \sum_{i=1}^{n_b} b_i q^{-i}, \quad H(q, \theta) = \frac{1}{A(q)} \quad (4.33)$$

Ovakva struktura predstavljena je sledećim blok dijagramom na slici 4.4



Slika 4.4: ARX struktura modela sistema.

Parametri polinoma $A(q)$ i $B(q)$ definišu vektor parametara

$$\theta^T = \{a_1 \ a_2 \dots a_{n_a} \ b_1 \ b_2 \dots b_{n_b}\} \quad (4.34)$$

dok je izlaz sistema opisan polinomijalnom relacijom, odnosno odgovarajućom diferencnom jednačinom

$$A(q)y(k) = B(q)u(k) + e(k) \quad (4.35)$$

Koristeći izraz za prediktor standardne strukture (4.24), uz uvrštavanje izraza za odgovarajuće funkcije prenosa u ARX strukturi $G(q, \theta) = B(q)/A(q)$; $H(q, \theta) = 1/A(q)$, dobija se konačan izraz za prediktor izlaza

$$\begin{aligned} \hat{y}(k|k-1, \theta) &= B(q)u(k) + [1 - A(q)]y(k) \\ &= \sum_{i=1}^{n_b} b_i q^{-i} u(k) - \sum_{i=1}^{n_a} a_i q^{-i} y(k) \end{aligned}$$

$$= \sum_{i=1}^{n_b} b_i u(k-i) - \sum_{i=1}^{n_a} a_i y(k-i) \quad (4.36)$$

Uvodeći dalje regresioni vektor

$$\varphi^T(k) = \{-y(k-1) - y(k-2) \dots - y(k-n_a) u(k-1) u(k-2) \dots u(k-n_b)\} \quad (4.37)$$

izvedeni izraz za prediktor izlaza ARX strukture modela može se napisati u formi modela linearne regresije

$$\hat{y}(k|k-1, \theta) = \theta^T \varphi(k) = \varphi^T(k) \theta. \quad (4.38)$$

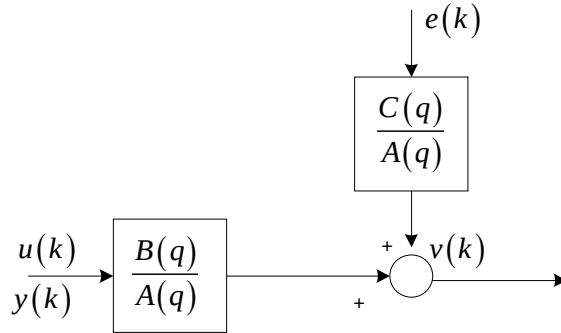
4.1.3. Model predikcije za ARMAX reprezentaciju modela sistema

ARMAX reprezentacija modela sistema predstavlja specijalan slučaj standardne strukture modela na slici 4.2 za specifičan izbor funkcija prenosa u standardnoj strukturi (reprezentaciji) modela

$$G(q, \theta) = \frac{B(q)}{A(q)}, \quad B(q) = \sum_{i=1}^{n_b} b_i q^{-i}, \quad A(q) = 1 + \sum_{i=1}^{n_a} a_i q^{-i},$$

$$H(q, \theta) = \frac{C(q)}{A(q)}, \quad C(q) = 1 + \sum_{i=1}^{n_c} c_i q^{-i}. \quad (4.39)$$

ARMAX struktura modela prikazana je na sledećoj slici



Slika 4.5: ARMAX struktura modela sistema

Na osnovu slike 4.5, izlazni signal kod ARMAX strukture opisan je relacijom

$$A(q)y(k) = B(q)u(k) + C(q)e(k) \quad (4.40)$$

gde je $u(k)$ pobudni (ulazni) signal a $e(k)$ diskretan beli šum nulte srednje vrednosti. Parametri (koeficijenti) polinoma A , B , C definišu vektor parametara θ u ovako usvojenoj parametrizaciji modela, odnosno

$$\theta^T = \{a_1 \ a_2 \dots a_{n_a} \ b_1 \ b_2 \dots b_{n_b} \ c_1 \ c_2 \dots c_{n_c}\} \quad (4.41)$$

Zamenom usvojenih izraza za funkcije prenosa $G = B/A$ i $H = C/A$ u ARMAX strukturi, uz korišćenje formule za prediktor izlaza opšte strukture (4.24), dobija se formula za prediktor izlaza ARMAX strukture

$$\hat{y}(k|k-1, \theta) = \frac{B(q)}{C(q)}u(k) + \left[1 - \frac{A(q)}{C(q)}\right]y(k) \quad (4.42)$$

Množenjem dobijenog izraza sa $C(q)$ i dodavanjem člana $[1 - C(q)]\hat{y}(k|k-1, \theta)$ levoj i desnoj strani poslednjeg izraza, kao i nultog člana $y(k) - y(k)$ desnoj strani izraza (4.42), može se pisati

$$\begin{aligned} C(q)\hat{y}(k|k-1, \theta) + [1 - C(q)]\hat{y}(k|k-1, \theta) &= B(q)u(k) + [C(q) - A(q)]y(k) \\ &+ [1 - C(q)]\hat{y}(k|k-1, \theta) + y(k) - y(k) \end{aligned} \quad (4.43)$$

odakle sledi izraz za prediktor izlaza ARMAX strukture

$$\hat{y}(k|k-1, \theta) = B(q)u(k) + [1 - A(q)]y(k) + [C(q) - 1][y(k) - \hat{y}(k|k-1, \theta)] \quad (4.44)$$

Uvodeći definicioni izraz za grešku predikcije (4.1), odnosno

$$\varepsilon(k|k-1, \theta) = y(k) - \hat{y}(k|k-1, \theta) \quad (4.45)$$

dobija se konačan izraz za prediktor izlaza ARMAX modela

$$\begin{aligned} \hat{y}(k|k-1, \theta) &= B(q)u(k) + [1 - A(q)]y(k) + [C(q) - 1]\varepsilon(k|k-1, \theta) \\ &= \sum_{\ell=1}^{n_b} b_{\ell} q^{-\ell} u(k) - \sum_{i=1}^{n_a} a_i q^{-i} y(k) + \sum_{j=1}^{n_c} c_j q^{-j} \varepsilon(k|k-1, \theta) \\ &= \sum_{\ell=1}^{n_b} b_{\ell} u(k - \ell) - \sum_{i=1}^{n_a} a_i y(k - i) + \sum_{j=1}^{n_c} c_j \varepsilon(k - j|k - j - 1, \theta) \end{aligned} \quad (4.46)$$

Uvodeći dalje regresioni vektor

$$\varphi^T(k, \theta) = \{-y(k-1) \dots -y(k-n_a) u(k-1) \dots u(k-n_b) \varepsilon(k-1|k-2, \theta) \dots \varepsilon(k-n_c|k-n_c-1, \theta)\} \quad (4.47)$$

i imajući u vidu definicioni izraz za vektor parametara θ u razmatranoj ARMAX strukturi (4.41), izraz za prediktor izlaza može se napisati u alternativnom obliku

$$\hat{y}(k|k-1, \theta) = \varphi^T(k, \theta)\theta = \theta^T \varphi(k, \theta) \quad (4.48)$$

Primetimo da je u ovom slučaju, za razliku od FIR i ARX struktura, vektor regresije φ funkcija vektora parametara θ , pošto greška predikcije ε predstavlja funkciju θ ,

tako da izvedeni izraz za prediktor izlaza ne predstavlja linearnu funkciju vektora parametara. Stoga se poslednji izraz naziva model pseudo-linearne regresije, jer formalno podseća na model linearne regresije kod koga regresioni vektor φ ne zavisi od vektora parametara θ .

4.2. PARAMETARSKA IDENTIFIKACIJA SISTEMA: ALGORITMI ESTIMACIJE PARAMETARA

Metode parametarske identifikacije koje će se razmatrati u ovom poglavlju pripadaju familiji metoda tipa greške predikcije (PEM metode). Ove metode se zasnivaju na minimizaciji kriterijumske funkcije (4.2), koju sačinjava suma kvadrata grešaka predikcija na nekom intervalu vremena ili, u opštem slučaju, predstavlja empirijski nelinearni indeks performanse (kriterijum kvaliteta ponašanja) (4.5b). Rešavanje postavljenog optimizacionog problema zahteva primenu određenih matematičkih alata, a jedan uobičajen postupak je metod najmanjih kvadrata (odgovarajući naziv na engleskom jeziku je Least Squares Method ili skraćeno LS) u slučaju srednje-kvadratnog kriterijuma (4.2), odnosno metode tipa stohastičkog Newton-Raphsonovog algoritma ili algoritma dinamičke stohastičke aproksimacije u slučaju generalnog nelinearnog indeksa performanse (4.5) [63,68,89]. Ove metode biće opisane u nastavku, a biće takođe pokazano kako se one mogu primeniti za estimaciju parametara u specifičnoj strukturi modela sistema. Ovakav zadatak često se naziva i parametarska estimacija (engleski parameter estimation).

4.2.1 Linearni algoritmi estimacije parametara

Kao što je već istaknuto, metod najmanjih kvadrata predstavlja matematički postupak za rešavanje problema identifikacije (estimacije) parametara u usvojenoj strukturi modela sistema, pri čemu je ovaj problem postavljen kao zadatak optimizacije (minimizacije) određene kvadratne kriterijumske funkcije (4.2). Naime, problem parametarske identifikacije može se formulisati na sledeći način: za zadati skup merenja ulaza i izlaza sistema $\{u(k), y(k)\}$, $k = 1, 2, \dots, N$ i usvojenu strukturu modela sistema, koja je poznata sa tačnošću do nepoznatog vektora parametara θ , proceniti (estimirati) vektor θ tako da izvedeni model aproksimira što je moguće bolje realan sistem. Naravno, verbalni iskaz “što je moguće bolje” je subjektivne prirode i neophodno je da se uvede matematički kriterijum ili mera za ocenu kvaliteta ponašanja ili validnosti modela. Jednu moguću takvu matematičku meru predstavlja kriterijum koji se sastoji od sume kvadrata grešaka predikcija na posmatranom intervalu posmatranja (4.2), odnosno

$$J_N(\theta) = \frac{1}{N} \sum_{k=1}^N \varepsilon^2(k|k-1, \theta) \quad (4.49)$$

Tada je optimalna procena vektora parametara $\theta = \hat{\theta}_N$ ona koja minimizira usvojeni kriterijum i data je sa (4.4), odnosno

$$\hat{\theta}_N = \arg \min_{\theta} J_N(\theta) \quad (4.50)$$

Jednokoračna greška predikcije definisana je izrazom (4.1), tj.

$$\varepsilon(k|k-1, \theta) = y(k) - \hat{y}(k|k-1, \theta) \quad (4.51)$$

gde je $\hat{y}(k|k-1, \theta)$ jednokoračni prediktor izlaza sistema $y(k)$ na bazi raspoložive informacije o sistemu, kroz merenja, do trenutka $k-1$ i za neki usvojeni vektor parametara θ . U prethodnom poglavlju je pokazano da se jednokoračni prediktor izlaza sistema, zavisno od usvojene strukture modela, može prikazati u formi jednačine linearne regresije (FIR i ARX strukture)

$$\hat{y}(k|k-1, \theta) = \varphi^T(k) \theta = \theta^T \varphi(k) \quad (4.52)$$

gde regresioni vektor $\varphi(k)$ ne zavisi od nepoznatog vektora parametara koji se estimira, ili u formi jednačine pseudolinearne regresije (ARMAX strukture)

$$\hat{y}(k|k-1, \theta) = \varphi^T(k, \theta) \theta = \theta^T \varphi(k, \theta) \quad (4.53)$$

kada regresioni vektor $\varphi(k, \theta)$ predstavlja funkciju vektora parametara θ . U prvom slučaju primena metoda najmanjih kvadrata za rešavanje postavljenog optimizacionog problema rezultuje u linearan sistem jednačina (4.6) ili (4.7), gde je $F(\cdot)$ kvadratna funkcija po nepoznatim elementima vektora parametara θ i takav metod naziva se linearan metod najmanjih kvadrata (engleski Linear Least Squares Method ili skraćeno LLS metod). U drugom slučaju, primena tehnike najmanjih kvadrata za minimizaciju usvojenog kriterijuma rezultuje u nelinearan sistem jednačina (4.6) ili (4.7), gde je $F(\cdot)$ kvadratna funkcija po nepoznatim elementima vektora parametara θ , te je neophodno primeniti neku iterativnu numeričku proceduru za rešavanje tako dobijenog nelinearnog sistema algebarskih jednačina. U poslednjem slučaju postupak najmanjih kvadrata se naziva pseudo-linearan metod najmanjih kvadrata (engleski Pseudo-Linear Least Square Method ili skraćeno PLLS metod).

4.2.1.1. Nerekurzivni linearan metod najmanjih kvadrata: LLS metod

Kao što je izvedeno u prethodnom poglavlju, jednokoračna predikcija izlaza za FIR i ARX reprezentaciju (strukturu) modela sistema definisana je izrazom (4.52) koji predstavlja linearnu funkciju vektora parametara θ , pošto regresioni vektor $\varphi(k)$ ne zavisi od nepoznatog vektora parametara θ . Kao što je ranije pokazano u (4.31), za FIR strukturu modela regresioni vektor sadrži samo zakašnjene (prethodne) ulazne signale, odnosno

$$\varphi^T(k) = \{u(k-1)u(k-2)..u(k-n_o)\} \quad (4.54)$$

dok za slučaj ARX strukture modela, koja predstavlja specifičnu IIR strukturu, ovaj vektor dodatno sadrži i zakašnjene (prethodne) izlazne signale, kao što je prikazano u (4.37), tj.

$$\varphi^T(k) = \{-y(k-1)..-y(k-n_a)u(k-1)..u(k-n_b)\} \quad (4.55)$$

Na taj način, jednokoračna greška predikcije (4.51) postaje

$$\varepsilon(k|k-1, \theta) = y(k) - \hat{y}(k|k-1, \theta) = y(k) - \varphi^T(k)\theta \quad (4.56)$$

te se usvojeni kriterijum valjanosti modela (4.2) svodi na

$$J_N(\theta) = \frac{1}{N} \sum_{k=1}^N [y(k) - \varphi^T(k)\theta]^2 \quad (4.57)$$

U tački minimuma razmatranog kriterijuma $\theta = \hat{\theta}_N$ gradijent kriterijumske funkcije jednak je nuli, kao što je naznačeno u (4.6), odnosno

$$\text{grad}_{\theta} J_N(\hat{\theta}_N) = \nabla_{\theta} J_N(\hat{\theta}_N) = \frac{\partial J_N(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}_N} = \left\{ \frac{\partial J_N(\theta)}{\partial \theta_1} \frac{\partial J_N(\theta)}{\partial \theta_2} \dots \frac{\partial J_N(\theta)}{\partial \theta_p} \right\}^T \Big|_{\theta=\hat{\theta}_N} = 0 \quad (4.58)$$

gde je p dimenzija vektora parametara $\theta = \{\theta_1 \theta_2 \dots \theta_p\}^T$, a 0 je $p \times 1$ dimenzioni nula vektor, čiji su svi elementi jednaki nuli. Primenom operatora parcijalnog diferenciranja na zadati kriterijum dobija se

$$\begin{aligned} \frac{\partial J_N(\theta)}{\partial \theta} &= \frac{\partial}{\partial \theta} \left\{ \frac{1}{N} \sum_{k=1}^N [y(k) - \varphi^T(k)\theta]^2 \right\} \\ &= \frac{1}{N} \sum_{k=1}^N \frac{\partial}{\partial \theta} \{ [y(k) - \varphi^T(k)\theta]^2 \} \\ &= \frac{1}{N} \sum_{k=1}^N 2[y(k) - \varphi^T(k)\theta] \frac{\partial [y(k) - \varphi^T(k)\theta]}{\partial \theta} \\ &= \frac{1}{N} \sum_{k=1}^N (-2)[y(k) - \varphi^T(k)\theta] \frac{\partial [\varphi^T(k)\theta]}{\partial \theta} \end{aligned} \quad (4.59)$$

Prilikom dobijanja poslednjeg izraza uzeto je u obzir da izlaz sistema $y(k)$ ne zavisi od vektora parametara θ . Koristeći dalje pravila linearne algebre koja se odnose na diferenciranje skalarnog proizvoda po vektoru [2,5]

$$\frac{\partial y^T x}{\partial x} = y, \quad \frac{\partial x^T y}{\partial x} = y \quad (4.60)$$

gde su x i y vektori kolone odgovarajućih dimenzija, a y^T i x^T označavaju vektore vrste, kao i usvajajući $y = \varphi(k)$ i $x = \theta$, zaključuje se da je

$$\frac{\partial [\varphi^T(k)\theta]}{\partial \theta} = \varphi(k) \quad (4.61)$$

Zamenom izvedenog izraza u relaciju za gradijent kriterijumske funkcije (4.59), uz izjednačavanje poslednjeg izraza sa nulom, konačno se dobija linearna vektorska jednačina za izračunavanje vektora parametra θ

$$\sum_{k=1}^N [y(k) - \varphi^T(k)\theta] \varphi(k) = 0 \quad (4.62)$$

odakle sledi da se procena najmanjih kvadrata $\theta = \hat{\theta}_N$ dobija kao rešenje linearne vektorske jednačine kojoj odgovara sistem linearnih jednačina po nepoznatim elementima vektora θ , a koje se nazivaju i normalne jednačine (engleski Normal equations) [63,68]

$$\left[\frac{1}{N} \sum_{k=1}^N \varphi^T(k) \varphi(k) \right] \hat{\theta}_N = \left[\frac{1}{N} \sum_{k=1}^N \varphi(k) y(k) \right]$$

odnosno

$$\hat{\theta}_N = \left[\frac{1}{N} \sum_{k=1}^N \varphi(k) \varphi^T(k) \right]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \varphi(k) y(k) \right] \quad (4.63)$$

pretpostavljajući da postoji inverzija navedene matrice. Ovaj uslov vezan je sa pojmom perzistentne eksitacije, o čemu će više reči biti kasnije.

U literaturi je uobičajeno da se LS procena prikaže u matricno-vektorskoj notaciji. U tom cilju definišimo $N \times p$ matricu Φ i $N \times 1$ vektor Y kao

$$\Phi = \begin{bmatrix} \varphi^T(1) \\ \varphi^T(2) \\ \vdots \\ \varphi^T(N) \end{bmatrix}; \quad Y = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} \quad (4.64)$$

tako da se usvojeni skalarni kriterijum $J_N(\theta)$ može prikazati u obliku

$$J_N(\theta) = \|Y - \Phi\theta\|^2 = (Y - \Phi\theta)^T (Y - \Phi\theta) \quad (4.65)$$

gde $\|x\|$ označava Euclidian-ovu normu vektora x . Dalje se može pisati

$$\begin{aligned} J_N(\theta) &= (Y^T - \theta^T \Phi^T) (Y - \Phi\theta) \\ &= Y^T Y - \theta^T (\Phi^T Y) - (Y^T \Phi) \theta + \theta^T (\Phi^T \Phi) \theta \end{aligned} \quad (4.66)$$

odakle sledi potreban uslov za pronalaženje minimuma kriterijuma

$$\text{grad}_{\theta} J_N(\theta) = \frac{\partial J_N(\theta)}{\partial \theta} = -\frac{\partial \theta^T (\Phi^T Y)}{\partial \theta} - \frac{\partial (Y^T \Phi) \theta}{\partial \theta} + \frac{\partial \theta^T (\Phi^T \Phi) \theta}{\partial \theta} = 0 \quad (4.67)$$

Prva dva člana u poslednjem izrazu dobijaju se korišćenjem prethodno definisanih pravila za diferenciranje skalarnog proizvoda vektora po vektoru (4.60), odnosno

$$\frac{\partial \theta^T (\phi^T Y)}{\partial \theta} = \phi^T Y, \quad \frac{\partial \theta^T (Y^T \phi)}{\partial \theta} = (Y^T \phi)^T = \phi^T Y \quad (4.68)$$

kao i dodatnog pravila za diferenciranje skalarne kvadratne forme po vektoru [2,5]

$$\frac{\partial x^T A x}{\partial x} = 2 A x \quad (4.69)$$

gde je x vektor kolona odgovarajućih dimenzija a A kvadratna i simetrična matrica ($A = A^T$), odakle se zaključuje, usvajajući $x = \theta$ i $A = \phi^T \phi$,

$$\frac{\partial \theta^T (\phi^T \phi) \theta}{\partial \theta} = 2 (\phi^T \phi) \theta \quad (4.70)$$

Zamenom izvedenih izraza (4.68), (4.70) za parcijalne izvode u relaciju za gradijent kriterijumske funkcije (4.67) dobija se linearna vektorska jednačina, tzv. Normalna jednačina, za izračunavanje LS procene vektora parametara $\theta = \hat{\theta}_N$

$$[\phi^T \phi] \hat{\theta}_N = \phi^T Y \quad (4.71)$$

odakle sledi

$$\hat{\theta}_N = [\phi^T \phi]^{-1} \phi^T Y = \left[\frac{1}{N} \phi^T \phi \right]^{-1} \left[\frac{1}{N} \phi^T Y \right] \quad (4.72)$$

pretpostavljajući da odgovarajuća inverzna matrica postoji. Kao što je već istaknuto, egzistencija ove matrice vezana je sa pojmom perzistentne eksitacije, o čemu će biti više reči u nastavku.

Primetimo, takođe, da je matrica čija se inverzija traži u stvari kovarijaciona matrica regresionog p -dimenzionog vektora $\varphi(k)$ ($p = n_a + n_b$ kod ARX modela a $p = n_b$ kod FIR modela), tako da se može pisati

$$R_\varphi^p(N) = \frac{1}{N} \sum_{k=1}^N \varphi(k) \varphi^T(k) = \frac{1}{N} \Phi^T \Phi \quad (4.73)$$

Posebno, za ARX strukturu modela važi ($p = n_a + n_b$)

$$R_{\varphi}^p(N) = \frac{1}{N} \sum_{k=1}^N \begin{bmatrix} -y(k-1) \\ -y(k-2) \\ \vdots \\ -y(k-n_a) \\ u(k-1) \\ u(k-2) \\ \vdots \\ u(k-n_b) \end{bmatrix} [-y(k-1) \dots -y(k-n_a) u(k-1) \dots u(k-n_b)]$$

$$= \begin{bmatrix} \frac{1}{N} \sum_{k=1}^N y^2(k-1) & \frac{1}{N} \sum_{k=1}^N y(k-1)y(k-2) & \dots & \frac{1}{N} \sum_{k=1}^N y(k-1)y(k-n_a) & -\frac{1}{N} \sum_{k=1}^N y(k-1)u(k-1) & \dots & -\frac{1}{N} \sum_{k=1}^N y(k-1)u(k-n_b) \\ \frac{1}{N} \sum_{k=1}^N y(k-2)y(k-1) & \frac{1}{N} \sum_{k=1}^N y^2(k-2) & \dots & \frac{1}{N} \sum_{k=1}^N y(k-2)y(k-n_a) & -\frac{1}{N} \sum_{k=1}^N y(k-2)u(k-1) & \dots & -\frac{1}{N} \sum_{k=1}^N y(k-2)u(k-n_b) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{N} \sum_{k=1}^N y(k-n_a)y(k-1) & \frac{1}{N} \sum_{k=1}^N y(k-n_a)y(k-2) & \dots & \frac{1}{N} \sum_{k=1}^N y^2(k-n_a) & -\frac{1}{N} \sum_{k=1}^N y(k-n_a)u(k-1) & \dots & -\frac{1}{N} \sum_{k=1}^N y(k-n_a)u(k-n_b) \\ -\frac{1}{N} \sum_{k=1}^N u(k-1)y(k-1) & -\frac{1}{N} \sum_{k=1}^N u(k-1)y(k-2) & \dots & -\frac{1}{N} \sum_{k=1}^N u(k-1)y(k-n_a) & \frac{1}{N} \sum_{k=1}^N u^2(k-1) & \dots & \frac{1}{N} \sum_{k=1}^N u(k-1)u(k-n_b) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{N} \sum_{k=1}^N u(k-n_b)y(k-1) & -\frac{1}{N} \sum_{k=1}^N u(k-n_b)y(k-2) & \dots & -\frac{1}{N} \sum_{k=1}^N u(k-n_b)y(k-n_a) & \frac{1}{N} \sum_{k=1}^N u(k-n_b)u(k-1) & \dots & \frac{1}{N} \sum_{k=1}^N u^2(k-n_b) \end{bmatrix}$$

$$= \begin{bmatrix} \hat{R}_{yy}(0) & \hat{R}_{yy}(1) & \dots & \hat{R}_{yy}(n_a-1) & \hat{R}_{yu}(0) & \hat{R}_{yu}(1) & \dots & \hat{R}_{yu}(n_b-1) \\ \hat{R}_{yy}(1) & \hat{R}_{yy}(0) & \dots & \hat{R}_{yy}(n_a-2) & \hat{R}_{yu}(-1) & \hat{R}_{yu}(0) & \dots & \hat{R}_{yu}(n_b-2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{yy}(n_a-1) & \hat{R}_{yu}(n_a-2) & \dots & \hat{R}_{yy}(0) & \hat{R}_{yu}(1-n_a) & \hat{R}_{yu}(2-n_a) & \dots & \hat{R}_{yu}(n_b-n_a) \\ -\hat{R}_{yu}(0) & -\hat{R}_{yu}(-1) & \dots & -\hat{R}_{yu}(-n_a+1) & \hat{R}_{uu}(0) & \hat{R}_{uu}(1) & \dots & \hat{R}_{uu}(n_b-1) \\ -\hat{R}_{yu}(1) & -\hat{R}_{yu}(0) & \dots & -\hat{R}_{yu}(-n_a+2) & \hat{R}_{uu}(1) & \hat{R}_{uu}(0) & \dots & \hat{R}_{uu}(n_b-2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -\hat{R}_{yu}(-1+n_b) & -\hat{R}_{yu}(-2+n_b) & \dots & -\hat{R}_{yu}(-n_a+n_b) & \hat{R}_{uu}(n_b-1) & \hat{R}_{uu}(n_b-2) & \dots & \hat{R}_{uu}(0) \end{bmatrix} \quad (4.74)$$

Prilikom definisanja ove matrice iskorišćena je osobina kros-korelacione funkcije da je $R_{uy}(\tau) = R_{yu}(-\tau)$, kao i da je autokorelaciona funkcija parna, odnosno $R_{xx}(-\tau) = R_{xx}(\tau)$.

Dakle, elementi $p \times p$ dimenzione korelacione matrice p -dimenzionog ($p = n_a + n_b$) regresionog vektora $\varphi(k)$ su procene auto-kovarijacionih i kros-kovarijacionih funkcija ulaznog i izlaznog signala, odnosno

$$R_{\varphi}^p(N) = \{R_{ij}(N)\}_{p \times p} \quad (4.75)$$

gde je

$$R_{ij}(N) = \hat{R}_{yy}(j-i) = \frac{1}{N} \sum_{k=1}^N y(k-i)y(k-j), \quad 1 \leq i, j \leq n_a$$

$$R_{ij}(N) = \hat{R}_{yu}(j-n_a-i) = -\frac{1}{N} \sum_{k=1}^N y(k-i)u(k-j), \quad 1 \leq i, j \leq n_a, \quad n_a+1 \leq j \leq p = n_a + n_b$$

$$R_{ij}(N) = \hat{R}_{uy}(j-i+n_a) = -\frac{1}{N} \sum_{k=1}^N u(k-i)y(k-j), \quad n_a+1 \leq i \leq p, \quad 1 \leq i, j \leq n_a$$

$$R_{ij}(N) = \hat{R}_{uu}(j-i) = \frac{1}{N} \sum_{k=1}^N u(k-i)u(k-j), \quad n_a+1 \leq i, j \leq p \quad (4.76)$$

U navedenim izrazima simbol “ $\wedge$ ” označava procenu, a R_{xx} i R_{xy} su autokorelacione i kros-korelacione funkcije razmatranih signala x i y , respektivno. Kod FIR strukture modela postoji samo blok sa procenama auto-korelacionih funkcija ulaznog signala, tj. za FIR strukturu $p = n_b$ i

$$R_{ij}(N) = \hat{R}_{uu}(j-i) = \frac{1}{N} \sum_{k=1}^N u(k-i)u(k-j), \quad 1 \leq i, j \leq p = n_b \quad (4.77)$$

Na sličan način, p -dimenzioni vektor kolona

$$f_\varphi(N) = \frac{1}{N} \sum_{k=1}^N \varphi(k)y(k) = \phi^T Y \quad (4.78)$$

može se za ARX strukturu (reprezentaciju) modela napisati u razvijenom obliku

$$f_\varphi(N) = \frac{1}{N} \sum_{k=1}^N \begin{bmatrix} -y(k-1) \\ -y(k-2) \\ \vdots \\ -y(k-n_a) \\ u(k-1) \\ \vdots \\ u(k-n_b) \end{bmatrix} y(k) = -\frac{1}{N} \sum_{k=1}^N \begin{bmatrix} y(k)y(k-1) \\ y(k)y(k-2) \\ \vdots \\ y(k)y(k-n_a) \\ y(k)u(k-1) \\ \vdots \\ y(k)u(k-n_b) \end{bmatrix} = \begin{bmatrix} -\hat{R}_{yy}(1) \\ -\hat{R}_{yy}(2) \\ \vdots \\ -\hat{R}_{yy}(n_a) \\ \hat{R}_{yu}(1) \\ \vdots \\ \hat{R}_{yu}(n_b) \end{bmatrix} \quad (4.79)$$

tako da su elementi ovog vektora

$$f_i(N) = -\hat{R}_{yy}(i) = -\frac{1}{N} \sum_{k=1}^N y(k)y(k-i), \quad 1 \leq i \leq n_a$$

$$f_i(N) = \hat{R}_{yu}(i-n_a) = \frac{1}{N} \sum_{k=1}^N y(k)u(k-i), \quad n_a+1 \leq i \leq p = n_a + n_b \quad (4.80)$$

4.2.1.2. Nerekurzivni linearni metod težinskih (ponderisanih) najmanjih kvadrata: WLS metod

Ako se umesto kriterijuma najmanjih kvadrata kao optimizacioni kriterijum za ocenu valjanosti modela usvoji sledeća njegova modifikacija, koja se naziva

kriterijum težinskih (ponderisanih) najmanjih kvadrata (engleski Weighted Least Squares ili skraćeno WLS ili WLLS ako se radi o linearnom algoritmu) [14,23,28,63,68]

$$J_N(\theta) = \|Y - \phi\theta\|_W^2 = (Y - \phi\theta)^T W (Y - \phi\theta) \quad (4.81)$$

gde je W pozitivno-definitna ($W > 0$) i simetrična ($W = W^T$) težinska matrica, tada je

$$J_N(\theta) = Y^T W Y - \theta^T \phi^T W Y - Y^T W \phi \theta + \theta^T \phi^T W \phi \theta \quad (4.82)$$

pa se iz uslova minimuma kriterijuma (4.82) dobija

$$\text{grad}_\theta J_N(\theta) = \frac{\partial J_N(\theta)}{\partial \theta} = -\frac{\partial \theta^T (\phi^T W Y)}{\partial \theta} - \frac{\partial (Y^T W \phi) \theta}{\partial \theta} + \frac{\partial \theta^T (\phi^T W \phi) \theta}{\partial \theta} = 0 \quad (4.83)$$

Primenom navedenih pravila za diferenciranje skalara po vektoru (4.60) i (4.69) dalje se dobija

$$-(\phi^T W Y) - (Y^T W \phi)^T + 2(\phi^T W \phi) \theta = 0 \quad (4.84)$$

odakle direktno sledi WLS procena težinskih najmanjih kvadrata

$$\hat{\theta}_N = [\phi^T W \phi]^{-1} \phi^T W Y \quad (4.85)$$

Za razmatranu ARX struktura modela matrice ϕ je dimenzije $N \times p$, $p = n_a + n_b$, vektor Y je kolona vektor dimenzije $N \times 1$, tako da je W matrica dimenzija $N \times N$. Posebno, ako se usvoji $W = cI$, gde je $c \geq 0$ a I jedinična matrica dimenzije $N \times N$, WLS metod se svodi na originalni nerekurzivni LLS metod. Ako se usvoji

$$W = \text{diag}\{\rho^{N-1} \rho^{N-2} \dots \rho \rho^0\} \quad (4.86)$$

gde $\text{diag}\{\cdot\}$ označava dijagonalnu matricu, to će se za konstantu $\rho < 1$ u proizvodu WY , koji definiše otežnjeni vektor merenja Y , veće težine dodeljivati potonjim merenjima $y(N), y(N-1), \dots$ u odnosu na početna merenja $y(1), y(2), \dots$, odnosno

$$WY = \begin{bmatrix} \rho^{N-1} y(1) \\ \rho^{N-2} y(2) \\ \vdots \\ \rho y(N-1) \\ \rho^0 y(N) \end{bmatrix} \quad (4.87)$$

dok će se za $\rho > 1$ veće težine dodeljivati početnim (prethodnim) merenjima. Takođe, ako se usvoji da je težinska matrica W jednaka inverznoj matrici matrice kovarijanse aditivnog šuma na izlazu sistema $e(i)$ u FIR i ARX strukturi modela

$$W = R_e^{-1} = [E\{ee^T\}]^{-1}; \quad e^T = \{e(1)e(2)..e(N)\} \quad (4.88)$$

gde $E\{\cdot\}$ označava matematičko očekivanje, WLS algoritam se naziva algoritam minimalne varijanse greške estimacije (skraćeno se označava sa MV), pošto se u navedenom slučaju optimizacioni kriterijum svodi na totalnu varijansu greške estimacije, kao sumu varijansi grešaka estimacije pojedinih parametara u vektoru parametara koji se estimira [3,23,63,68].

4.2.1.3. Nerekurzivni pseudo-linearni metod najmanjih kvadrata: PLLS metod

Prilikom izvođenja linearnog metoda najmanjih kvadrata (LLS) ili linearnog metoda ponderisanih (težinskih) najmanjih kvadrata (WLLS) ključna pretpostavka je bila da je regresioni vektor $\varphi(k)$ u linearnom modelu predikcije $\hat{y}(k|k-1, \theta) = \varphi^T(k)\theta$ nezavisan od nepoznatog vektora parametara, što je podrazumevalo FIR ili ARX reprezentaciju (strukturu) modela. Ukoliko ova pretpostavka nije ispunjena, već je i regresioni vektor funkcija nepoznatog vektora parametara, $\varphi(k, \theta)$, što je slučaj kod ARMAX reprezentacija modela, tada je model prediktora $\hat{y}(k|k-1, \theta) = \varphi^T(k, \theta)\theta$ nelinearna funkcija vektora parametara θ , te se ovakav prediktor naziva pseudo-linearan prediktor. Kriterijum (4.2) koji se minimizira tada postaje

$$J_N(\theta) = \frac{1}{N} \sum_{k=1}^N [y(k) - \varphi^T(k, \theta)\theta]^2 \quad (4.89)$$

a minimizacijom ovog kriterijuma dobija se procena $\hat{\theta}_N$ nepoznatog vektora parametara θ

$$\theta = \hat{\theta}_N = \arg \min_{\theta \in D_M} J_N(\theta); \quad J_N(\hat{\theta}_N) \leq J_N(\theta) \quad \text{za} \quad \forall \theta \in D_M \quad \text{i} \quad \theta \neq \hat{\theta}_N. \quad (4.90)$$

gde je D_M skup svih mogućih vrednosti nepoznatog vektora parametara θ . Rešenje ovako postavljenog optimizacionog problema, koji se naziva i pseudo-linearan problem najmanjih kvadrata (skraćenica PLLS), zahteva primenu iterativnih nelinearnih numeričkih algoritama [9,74]. Ovakav postupak je računski složen, a i sama konvergencija procena parametara koje su generisane nelinearnim numeričkim algoritmima ka stvarnom (tačnom) vektoru parametara nije u opštem slučaju garantovana.

Za razliku od klasičnog zadatka optimizacije, gde je u opštem slučaju poznata samo kriterijumska funkcija, ovde je pored kriterijuma $J_N(\theta)$ poznat i izraz za njen argument, odnosno grešku predikcije $\varepsilon(k, \theta)$, tj.

$$J_N(\theta) = \frac{1}{2N} \sum_{k=1}^N \varepsilon^2(k, \theta), \quad \varepsilon(k, \theta) = y(k) - \varphi^T(k, \theta)\theta \quad (4.91)$$

što će omogućiti da se izvedu optimizacione (minimizacione) metode koje imaju relativno dobra konvergentna svojstva. Neke od najčešće korišćenih metoda biće razmatrane u nastavku teksta.

Razvojem kriterijuma $J_N(\theta)$ oko neke tačke u prostoru parametara $\theta = \theta_0$ dobija se

$$J_N(\theta) = J_N(\theta_0) + g^T(\theta_0)(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)^T H(\theta_0)(\theta - \theta_0) + o(\|\theta - \theta_0\|) \quad (4.92)$$

gde je

$$g(\theta_0) = \text{grad} J_N(\theta_0) = \frac{\partial J_N(\theta_0)}{\partial \theta}; \quad H(\theta_0) = \frac{\partial^2 J_N(\theta_0)}{\partial \theta^2}, \quad \lim_{\theta \rightarrow \theta_0} \frac{o(\|\theta - \theta_0\|)}{\|\theta - \theta_0\|} = 0 \quad (4.93)$$

pri čemu simbol $\|\cdot\|$ označava Euklidovu normu vektora. Ako se dalje pretpostavi da je θ_0 blisko minimumu θ^* kriterijuma, tada je $J_N(\theta_0) \approx 0$, te je potreban uslov minimuma definisan relacijom

$$\frac{\partial J_N(\theta)}{\partial \theta} = \frac{\partial [g^T(\theta_0)(\theta - \theta_0)]}{\partial \theta} + \frac{1}{2} \frac{\partial [(\theta - \theta_0)^T H(\theta_0)(\theta - \theta_0)]}{\partial \theta} = 0 \quad (4.94)$$

Koristeći prethodno definisana pravila za diferenciranje skalara po vektoru (4.60) i (4.69), relacija (4.94) se svodi na

$$g(\theta_0) + \frac{1}{2} 2H(\theta_0)(\theta - \theta_0) = 0 \quad (4.95)$$

odakle sledi

$$\theta = \theta_0 - H^{-1}(\theta_0)g(\theta_0) \quad (4.96)$$

Ukoliko je $J_N(\theta)$ kvadratna funkcija ovako određeno θ odgovaraće minimumu kriterijumske funkcije θ^* i algoritam će u jednom koraku odrediti traženu tačku minimuma θ^* . U svakom drugom slučaju algoritam se može primeniti u više koraka, u cilju iterativnog pronalaženja minimuma kriterijuma θ^* . Na taj način, ako je θ_i procena vektora parametara u i -tom koraku ($i = 0, 1, 2, \dots$), tada je procena vektora parametara u narednom koraku (iteraciji)

$$\theta_{i+1} = \theta_i - H^{-1}(\theta_i)g(\theta_i); \quad i = 0, 1, 2, \dots; \quad \theta_0 \text{-početna procena} \quad (4.97)$$

Ovakav iterativan postupak naziva se deterministički Newton-ov ili Newton-Raphsonov algoritam [7,63]. Algoritam može zahtevati nekoliko ili čak relativno mnogo iteracija da bi se odredio minimum θ^* kriterijuma. Kao mogući kriterijum za zaustavljanje više koračne procedure može se koristiti jedan od sledećih uslova:

$$\|g(\theta_{i+1})\| \leq \varepsilon, \quad |J(\theta_{i+1}) - J(\theta_i)| \leq \varepsilon, \quad \|\theta_{i+1} - \theta_i\| \leq \varepsilon \quad (4.98)$$

gde je ε mala pozitivna konstanta.

Ako je kriterijumska funkcija $J_N(\theta)$ konveksna (konveksna kriterijumska funkcija J zadovoljava Jensenovu nejednakost $J(\lambda\theta_1 + (1-\lambda)\theta_2) \leq \lambda J(\theta_1) + (1-\lambda)J(\theta_2)$ za $0 < \lambda < 1$) tada je i unimodalna, tj. poseduje jedan ekstremum i algoritam će konvergirati ka tom jedinstvenom minimumu. Diferencijabilna funkcija $J_N(\theta)$ je konveksna ukoliko je Hessian matrica drugih izvoda $H(\theta) = \partial^2 J_N(\theta) / \partial \theta^2$ pozitivno-definitna ($H > 0$). Međutim, unimodalna funkcija u opštem slučaju ne mora biti i konveksna. Ukoliko se ne može dokazati ni konveksnost ni unimodalnost kriterijumske funkcije, tada ona poseduje nekoliko minimuma, a minimum sa najmanjom vrednošću naziva se globalnim, dok se ostali minimumi nazivaju lokalnim. U navedenom slučaju ne postoji garancija da će algoritam konvergirati ka globalnom minimumu, odnosno kada algoritam dostigne tačku minimuma ne postoji mogućnost da se ispita da li je dostignuti minimum zaista i globalni. Inače, generalno je vrlo teško da se dokaže bilo konveksnost bilo unimodalnost kriterijumske funkcije i jedino u slučaju kriterijumske funkcije u formi sume kvadrata grešaka predikcije koje su linearne funkcije vektora parametara (FIR i ARX reprezentacije modela) može se dokazati konveksnost kriterijuma $J_N(\theta)$.

Primetimo, takođe, da su važne osobine više koračne (iterativne) procedure brzina konvergencije i robusnost. Brzina konvergencije odnosi se na sposobnost algoritma da locira tačku ekstremuma u što je moguće kraćem vremenskom intervalu, odnosno za što je moguće manji broj koraka (iteracija), dok robusnost označava sposobnost algoritma da odredi tačku minimuma za proizvoljan oblik kriterijumske funkcije.

Brzina konvergencije algoritma određuje se na osnovu faktora konvergencije

$$\beta = \frac{\|\theta_{i+1} - \theta^*\|}{\|\theta_i - \theta^*\|} \quad (4.99)$$

i ukoliko je $\beta \in (0,1)$ takav iterativni algoritam je linearno konvergentan. Iterativni algoritmi za koje je $\beta \approx 0$ konvergiraju vrlo brzo, pošto je tada $\theta_{i+1} \approx \theta^*$. Posebno, za $\beta = 0$ algoritam će posedovati super brzu linearnu konvergenciju. Newtonov iterativni metod poseduje kvadratnu konvergenciju, pošto kod njega β^2 teži ka pozitivnoj konstanti. Inače, Newtonov algoritam će konvergirati vrlo brzo u okolini tačke minimuma, pošto je tada izvršena aproksimacija kriterijumske funkcije odgovarajućom kvadratnom funkcijom (4.92) adekvatna. Takođe, ukoliko nije moguće da se Hessian izračuna egzaktno, ne može se garantovati konvergencija ovog algoritma, te se u takvim situacijama uvodi pozitivna skalarna veličina s_i kojom se utiče na priraštaj vektora parametara, odnosno veličinu koraka algoritma, kako bi se obezbedila njegova konvergencija [7,63,88,89]

$$\theta_{i+1} = \theta_i - s_i H^{-1}(\theta_i) g(\theta_i), \quad i = 0,1,2,\dots, \quad \theta_0 \text{-dato} \quad (4.100)$$

Faktor s_i može biti fiksni ili promenljiv. Na primer, dok god je ispunjen uslov $J(\theta_{i+1}) < J(\theta_i)$ može se koristiti konstantan faktor $s_i > 1$ (na primer, $s_i = 1.3$), a kada

dođe do neadekvatnog koraka (otkaza algoritma) koji je okarakterisan uslovom $J(\theta_{i+1}) > J(\theta_i)$ veličinu koraka treba smanjiti njegovim množenjem sa nekom konstantom manjom od 1 (na primer, 0.2), pa nastaviti proceduru uz testiranje vrednosti kriterijumske funkcije.

Posebno, ukoliko se Hessian aproksimira sa jediničnom matricom, $H \approx I$, tada se Newton algoritam svodi na gradijentni algoritam [7,63,88,89]

$$\theta_{i+1} = \theta_i - s_i g(\theta_i), \quad i = 0, 1, 2, \dots, \quad \theta_0 \text{-dato} \quad (4.101)$$

kod koga se približavanje procena minimumu kriterijumske funkcije θ^* vrši u anti-gradijentnom smeru $-g(\theta_i)$, dok s_i definiše veličinu koraka u navedenom smeru. Ovakav metod naziva se metod najbržeg spusta (engleski steepest descent method) [7,63].

Naravno, primena Newtonovog algoritma zahteva izračunavanje gradijenta i Hessiana

$$g(\theta) = \frac{\partial J(\theta)}{\partial \theta} = \begin{bmatrix} \frac{\partial J(\theta)}{\partial \theta_1} \\ \frac{\partial J(\theta)}{\partial \theta_2} \\ \vdots \\ \frac{\partial J(\theta)}{\partial \theta_p} \end{bmatrix} = \begin{bmatrix} g_1(\theta) \\ g_2(\theta) \\ \vdots \\ g_p(\theta) \end{bmatrix} \quad H(\theta) = \begin{bmatrix} \frac{\partial g_1(\theta)}{\partial \theta_1} & \frac{\partial g_1(\theta)}{\partial \theta_2} & \dots & \frac{\partial g_1(\theta)}{\partial \theta_p} \\ \frac{\partial g_2(\theta)}{\partial \theta_1} & \frac{\partial g_2(\theta)}{\partial \theta_2} & \dots & \frac{\partial g_2(\theta)}{\partial \theta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_p(\theta)}{\partial \theta_1} & \frac{\partial g_p(\theta)}{\partial \theta_2} & \dots & \frac{\partial g_p(\theta)}{\partial \theta_p} \end{bmatrix} \quad (4.102)$$

polazeći od izraza za kriterijumsku funkciju

$$J_N(\theta) = \frac{1}{2N} \sum_{k=1}^N \varepsilon^2(k, \theta); \quad \varepsilon(k, \theta) = y(k) - \varphi^T(k, \theta) \theta \quad (4.103)$$

i diferenciranjem po elementu θ_i , $i=1, 2, \dots, p$, vektora parametara θ . Da bi se obezbedila konvergencija algoritma za proizvodnju početno pogađanje vektora parametara koji se estimira, uobičajeno je da se Hessian aproksimira odgovarajućom pozitivno-definitnom matricom i takav algoritam se naziva Gauss-Newtonov metod []. S obzirom da je ovaj rad primarno posvećen rekurzivnim algoritmima parametarske identifikacije, ovom problemu neće biti posvećena dalja pažnja.

4.2.1.4. Rekurzivni linearni metod najmanjih kvadrata

Osnovni nedostatak prethodno razmatranog nerekurzivnog metoda najmanjih kvadrata je što se sa povećanjem merne populacije $z^N = \{u(i), y(i); i=1, \dots, N\}$ za jedno merenje $z(N+1) = \{u(N+1), y(N+1)\}$ celokupni proračun mora da ponovi u cilju generisanja nove procene parametara $\hat{\theta}_{N+1}$, pri čemu se povećavaju i računarski zahtevi vezani za potreban memorijski prostor i vreme izračunavanja. Stoga se prikazani metod najmanjih kvadrata naziva i metod paketne obrade (engleski batch

processing method) podataka ili nerekurzivni (engleski off-line) metod. Za razliku od nerekurzivnih metoda, rekurzivne metode parametarske identifikacije generišu novu procenu parametara $\hat{\theta}_{N+1}$ na bazi prethodno izračunate procene $\hat{\theta}_N$ i novo dobijene informacije kroz merenja $z(N+1) = \{u(N+1), y(N+1)\}$, kao što je naznačeno u (4.9). Rekurzivna varijanta metode najmanjih kvadrata obično se izvodi iz nerekurzivnog metoda koristeći alat matrične linearne algebra [2,5]. U tom cilju usvojimo reprezentaciju modela u formi jednačine linearne regresije

$$y(k) = \phi^T(k)\theta + e(k); \quad k = 1, 2, \dots, N \quad (4.104)$$

gde je $y(k)$ vektor izlaza u k -tom diskretnom trenutku, $\phi(k)$ je regresioni vektor, $e(k)$ predstavlja beli šum nulte srednje vrednosti na izlazu sistema, a θ je nepoznati vektor parametara koji se estimira. Sistem od N relacija može se napisati u matrično-vektorskoj formi

$$Y(N) = \phi(N)\theta + E(N) \quad (4.105)$$

gde su odgovarajući vektori i matrice definisani sa

$$Y(N) = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}; \quad \phi(N) = \begin{bmatrix} \phi^T(1) \\ \phi^T(2) \\ \vdots \\ \phi^T(N) \end{bmatrix}; \quad E(N) = \begin{bmatrix} e(1) \\ e(2) \\ \vdots \\ e(N) \end{bmatrix} \quad (4.106)$$

Linearni algoritam težinskih najmanjih kvadrata (engleski Weighted Linear Least Squares ili WLLS) minimizira kvadratni optimizacioni kriterijum

$$J_N(\theta) = \frac{1}{2N} \varepsilon^T(N, \theta) W(N) \varepsilon(N, \theta) = \frac{1}{2N} \sum_{i=1}^N w(i) \varepsilon^2(i, \theta) \quad (4.107)$$

gde je vektor grešaka predikcija ili reziduala merenja

$$\varepsilon(N, \theta) = Y(N) - \phi(N)\theta \quad (4.108)$$

a težinska kriterijumska matrica je dijagonalna

$$W(N) = \text{diag}\{w(1) w(2) \dots w(N)\} \quad (4.109)$$

Procena najmanjih kvadrata koji minimizira dati kriterijum je

$$\begin{aligned} \hat{\theta}_N &= [\phi^T(N) W(N) \phi(N)]^{-1} \phi^T(N) W(N) Y(N) \\ &= \left[\frac{1}{N} \sum_{i=1}^N w(i) \phi(i) \phi^T(i) \right]^{-1} \left[\frac{1}{N} \sum_{i=1}^N w(i) \phi(i) y(i) \right] \end{aligned} \quad (4.110)$$

Na identičan način, procena najmanjih kvadrata zasnovana na $(N+1)$ merenju ulaza i izlaza sistema data je sa

$$\hat{\theta}_{N+1} = [\phi^T(N+1)W(N+1)\phi(N+1)]^{-1} \phi^T(N+1)W(N+1)Y(N+1) \quad (4.111)$$

pri čemu je

$$Y(N+1) = \begin{bmatrix} Y(N) \\ y(N+1) \end{bmatrix}; \quad \phi(N+1) = \begin{bmatrix} \phi(N) \\ \phi^T(N+1) \end{bmatrix}; \quad E(N+1) = \begin{bmatrix} E(N) \\ e(N+1) \end{bmatrix};$$

$$W(N+1) = \begin{bmatrix} W(N) & 0 \\ 0 & w(N+1) \end{bmatrix} \quad (4.112)$$

Posebno, ako se izabere da su dijagonalni elementi težinske matrice $W(N)$ dati u formi eksponencijalnih faktora [3,23,63,68]

$$w(i) = \rho^{N-i}, \quad i = 1, 2, \dots, N \quad (4.113)$$

gde je ρ pozitivna konstanta, tada je

$$W(N+1) = \begin{bmatrix} \rho W(N) & 0 \\ 0 & 1 \end{bmatrix}. \quad (4.114)$$

Dalje se može pisati

$$\begin{aligned} \hat{\theta}_{N+1} &= \left\{ [\phi^T(N)\phi(N+1)] \begin{bmatrix} W(N) & 0 \\ 0 & w(N+1) \end{bmatrix} \begin{bmatrix} \phi(N) \\ \phi^T(N+1) \end{bmatrix} \right\}^{-1} \times \\ &\quad \left\{ [\phi^T(N)\phi(N+1)] \begin{bmatrix} W(N) & 0 \\ 0 & w(N+1) \end{bmatrix} \begin{bmatrix} Y(N) \\ y(N+1) \end{bmatrix} \right\} \\ &= [\phi^T(N)W(N)\phi(N) + w(N+1)\phi(N+1)\phi^T(N+1)]^{-1} \cdot \\ &\quad [w(N+1)\phi(N+1)y(N+1) + \phi^T(N)W(N)Y(N)] \end{aligned} \quad (4.115)$$

ili ako su elementi težinske matrice $W(N)$ eksponencijalni faktori $w(i) = \rho^{N-i}$, $i = 1, \dots, N$

$$\begin{aligned} \hat{\theta}_{N+1} &= [\rho \phi^T(N)W(N)\phi(N) + \phi(N+1)\phi^T(N+1)]^{-1} \cdot \\ &\quad [\phi(N+1)y(N+1) + \rho \phi^T(N)W(N)Y(N)] \end{aligned} \quad (4.116)$$

Usvajajući dalje oznaku za matricu pojačanja algoritma

$$P(N) = [\phi^T(N)W(N)\phi(N)]^{-1} \quad (4.117)$$

može se pisati

$$\hat{\theta}_N = P(N)\phi^T(N)W(N)Y(N) \quad (4.118)$$

pri čemu je

$$\hat{\theta}_{N+1} = P(N+1)[w(N+1)\phi(N+1)y(N+1) + \phi^T(N)W(N)Y(N)] \quad (4.119)$$

$$P^{-1}(N+1) = P^{-1}(N) + w(N+1)\phi(N+1)\phi^T(N+1) \quad (4.120)$$

Na osnovu poslednje dve relacije se zaključuje da je

$$P^{-1}(N) = P^{-1}(N+1) - w(N+1)\phi(N+1)\phi^T(N+1) \quad (4.121)$$

$$\phi^T(N)W(N)Y(N) = P^{-1}(N)\hat{\theta}_N \quad (4.122)$$

odakle sledi

$$\phi^T(N)W(N)Y(N) = P^{-1}(N+1)\hat{\theta}_N - w(N+1)\phi(N+1)\phi^T(N+1)\hat{\theta}_N \quad (4.123)$$

Zamenom izraza (4.123) u relaciju (4.119) dalje sledi

$$\begin{aligned} \hat{\theta}_{N+1} &= P(N+1)[w(N+1)\phi(N+1)y(N+1) + P^{-1}(N+1)\hat{\theta}_N - w(N+1)\phi(N+1)\phi^T(N+1)\hat{\theta}_N] \\ &= \hat{\theta}_N + P(N+1)\phi(N+1)w(N+1)[y(N+1) - \phi^T(N+1)\hat{\theta}_N] \end{aligned} \quad (4.124)$$

Na taj način, definitivni rekursivni oblik linearnog metoda težinskih najmanjih kvadrata, koji se skraćeno označava i sa RWLLS, dat je sa

$$\hat{\theta}_{N+1} = \hat{\theta}_N + K(N+1)[y(N+1) - \phi^T(N+1)\hat{\theta}_N]; \quad N = 0, 1, \dots$$

$$K(N+1) = P(N+1)\phi(N+1)w(N+1)$$

$$P^{-1}(N+1) = P^{-1}(N) + w(N+1)\phi(N+1)\phi^T(N+1) \quad (4.125)$$

Posebno, usvajajući da je težinska matrica $W(N) = I$, gde je I jedinična matrica, zaključuje se da je $w(N+1) = 1$ a odgovarajući algoritam je rekursivni linearni metod najmanjih kvadrata, koji se označava sa RLLS. Takođe, ukoliko se usvoji da je težinska matrica $W(N)$ jednaka inverznoj kovarijacionoj matrici belog šuma nulte srednje vrednosti $\{e(k)\}$, tada je $W(N) = R_e^{-1} = \sigma_e^{-2}I$, gde je σ_e^2 varijansa šuma, tako da je $w(N+1) = \sigma_0^{-2}$, a odgovarajući algoritam je rekursivni linearni metod minimalne varijanse greške, koji se skraćeno označava sa RLMV. Kod algoritma sa eksponencijalnim ponderisanjem merenja, gde su težinski koeficijenti $w(i) = \rho^{N-i}$,

$i = 1, \dots, N$, u dobijenim izrazima treba matricu $W(N)$ zameniti sa $\rho W(N)$ i usvojiti $w(N+1)=1$. Naime, na osnovu (4.110), (4.116) i (4.117) se zaključuje da je

$$\begin{aligned} P^{-1}(N+1) &= \rho P^{-1}(N) + \varphi(N+1)\varphi^T(N+1) \\ &= \rho \phi^T(N)W(N)\phi(N) + \varphi(N+1)\varphi^T(N+1) \end{aligned} \quad (4.126)$$

pri čemu je po osnovu (4.116)

$$\hat{\theta}_{N+1} = P(N+1) \left[\rho \phi^T(N)W(N)Y(N) + \varphi(N+1)y(N+1) \right] \quad (4.127)$$

gde je procena najmanjih kvadrata na bazi mernog uzorka dužine N , odnosno u N -tom diskretnom trenutku, definisana sa (4.117) i (4.118)

$$\phi^T(N)W(N)Y(N) = P^{-1}(N)\hat{\theta}_N \quad (4.128)$$

Analogno kao u prethodnom slučaju, zamenom (4.128) u (4.116), uz korišćenje (4.126), dobija se

$$\begin{aligned} \hat{\theta}_{N+1} &= P(N+1) \left[\varphi(N+1)y(N+1) + \rho P^{-1}(N)\hat{\theta}_N \right] \\ &= P(N+1) \left\{ \varphi(N+1)y(N+1) + \left[P^{-1}(N+1) - \varphi(N+1)\varphi^T(N+1) \right] \hat{\theta}_N \right\} \\ &= \hat{\theta}_N + P(N+1)\varphi(N+1) \left[y(N+1) - \varphi^T(N+1)\hat{\theta}_N \right] \end{aligned} \quad (4.129)$$

što se i direktno dobija iz prethodno izvedenog izraza (4.124) usvajajući $w(N+1)=1$.

Na taj način rekurzivna varijanta linearnog metoda najmanjih kvadrata sa eksponencijalnim ponderisanjem observacije data je sa

$$\begin{aligned} \hat{\theta}_{N+1} &= \hat{\theta}_N + K(N+1) \left[y(N+1) - \varphi^T(N+1)\hat{\theta}_N \right]; \quad N = 0, 1, \dots \\ K(N+1) &= P(N+1)\varphi(N+1) \\ P^{-1}(N+1) &= \rho P^{-1}(N) + \varphi(N+1)\varphi^T(N+1) \end{aligned} \quad (4.130)$$

Primena algoritma zahteva određivanje inverzne matrice prilikom određivanja matrice pojačanja algoritma, što se može izbeći korišćenjem leme o inverziji matrica, na osnovu koje se matična relacija

$$P_2^{-1} = P_1^{-1} + H^T R^{-1} H \quad (4.131)$$

gde su R i P_1 regularne matrice, može napisati u alternativnom obliku [23,63,68,88]

$$P_2 = P_1 - P_1 H^T (H P_1 H^T + R)^{-1} H P_1. \quad (4.132)$$

Usvajajući u (4.125) da je

$$P_2 = P(N+1), \quad P_1 = P(N), \quad H^T = \varphi(N+1), \quad H = \varphi^T(N+1), \\ R^{-1} = w(N+1)$$

dobija se alternativan izraz za matricu pojačanja algoritma

$$P(N+1) = P(N) - P(N)\varphi(N+1)\left[\varphi^T(N+1)P(N)\varphi(N+1) + w^{-1}(N+1)\right]^{-1}\varphi^T(N+1)P(N) \\ (4.133)$$

Primetimo da je član u srednjim zagradama skalarna veličina u slučaju sistema sa jednim ulazom i jednim izlazom (engleski single input-single output ili SISO sistemi), tako da poslednja relacija ne zahteva inerziju matrice.

U slučaju rekurzivnog težinskog linearnog algoritma najmanjih kvadrata sa eksponencijalnim ponderisanjem observacija treba usvojiti u (4.130)

$$P_2 = P(N+1), \quad P_1 = \rho^{-1}P(N), \quad H^T = \varphi(N+1), \quad H = \varphi^T(N+1), \quad R^{-1} = 1$$

pa se na osnovu leme o inverziji matrice (4.131) i (4.132) može pisati

$$P(N+1) = \frac{1}{\rho} \left\{ P(N) - P(N)\varphi(N+1)\left[\varphi^T(N+1)P(N)\varphi(N+1) + \rho\right]^{-1}\varphi^T(N+1)P(N) \right\} \\ (4.134)$$

Kao što je već istaknuto, izborom faktora $\rho < 1$, koji se u navedenom slučaju naziva i faktor zaboravljanja (engleski forgetting factor) u algoritmu za estimaciju parametara manje težine dodeljuju se prethodnim merenjima izlaza u odnosu na potonja merenja, čime je u algoritam implicitno ugrađena eksponencijalno opadajuća memorija [3,1,3]. Na taj način, izborom faktora zaboravljanja $\rho < 1$, ali koji je blizak jediničnoj vrednosti, mogu se pratiti vremenski promenljivi parametri modela, čija je dinamika promene relativno spora. Međutim, praćenje naglih promena parametara, kao i vremenski brzo promenljivih parametara, zahteva da se faktor ρ menja adaptivno u svakom vremenskom koraku. Naime, u opštem slučaju, što je vrednost ρ više udaljena od 1 algoritam će bolje pratiti brze promene parametara, ali će lošije pratiti konstantne i sporo promenljive parametre, dok će obrnuto vrednost ρ bliska 1 omogućiti dobro praćenje konstantnih i sporo promenljivih parametara, uz lošije praćenje relativno brzo promenljivih parametara modela. Na taj način, zavisno od dinamike promene parametara treba podešavati faktor ρ u svakom koraku rada algoritma, ali ovakva razmatranja izlaze iz konteksta ovog rada. Takođe, pošto se za vrednosti ρ bliske 1 može pisati

$$\rho^k = e^{k \ln \rho} = e^{k \ln(1+\rho-1)} \approx e^{-k(1-\rho)} \\ (4.135)$$

ili

$$\rho^k = e^{k/\tau}, \quad \tau = \frac{1}{1-\rho} \quad (4.136)$$

zaključuje se da je za $\rho < 1$, pri čemu je ρ blisko 1, efektivna memorija algoritma

$$\tau = \frac{1}{\log \rho} \approx \frac{1}{1-\rho} \quad (4.137)$$

Pošto je za poslednje merenje $k = N$ (u težinskoj matrici $W(N)$ element na glavnoj dijagonali je ρ^{N-k} , $k=1, \dots, N$), odnosno N -tom merenju se dodeljuje težina $\rho^0 = 1$, izvedeni izraz pokazuje da će se merenjima starijim od $\tau (k > \tau)$ dodeliti težine manje od $e^{-1} \approx 0.36 (k = \tau)$. Drugim rečima, ovako izabran težinski faktor odgovara eksponencijalno opadajućoj memoriji algoritma, pri čemu vremenska konstanta τ eksponencijalne krive odgovara približno efektivnoj dužini memorije algoritma u usvojenim jedinicama vremena ili broju perioda diskretnih merenja. Na taj način, izbor dužine memorije algoritma τ , kojim je određena direktno vrednost faktora zaboravljanja ρ , zavisi od očekivane dinamike promene parametara i treba je tako usvojiti da parametri budu približno konstantni na intervalu dužine τ . Na nestacionarnim intervalima merenja, kojima odgovaraju brzo promenljivi parametri, treba usvojiti malo τ , kome odgovara $\rho < 1$, dok na stacionarnim intervalima merenja, kojima odgovaraju približno konstantni parametri, treba usvojiti relativno veliko τ (teorijski $\tau \rightarrow \infty$), kome odgovara $\rho \approx 1$ (teorijski $\rho = 1$). U praksi se ponekad usvaja da je faktor ρ vremenski promenljiv na intervalu kvazistacionarnosti i da asimptotski eksponencijalno raste ka vrednosti 1 (kada $k \rightarrow \infty$), odnosno u okviru takvog intervala merenja može se usvojiti [3,53,63]

$$\rho(k) = \rho_o \rho(k-1) + (1-\rho_o), \quad k=1,2,\dots,K \quad (4.138)$$

pri čemu je uobičajen izbor $\rho_o = 0.99$ i $\rho(0) = 0.95$.

Za startovanje ($N=0$) bilo koje od predloženih varijanti rekursivnih algoritama potrebno je usvojiti početne procene parametara $\hat{\theta}_o$ i početnu vrednost matrice pojačanja algoritma $P(0)$ (ove vrednosti se nazivaju i početna pogađanja) i ukoliko ne postoji nikakva apriorna informacija o očekivanim vrednostima parametara modela uobičajeno je da se usvoji [23,47,63,68]

$$\hat{\theta}_o = 0, \quad P(0) = c^2 I, \quad c^2 \gg 1. \quad (4.139)$$

gde je I jedinična matrica. Takođe, početno pogađanje $\hat{\theta}_o$ može se generisati i nerekursivnim linearnim algoritmima najmanjih kvadrata na paketu merenja odgovarajuće dužine. Uz navedeno, može se pokazati da matrica $P(N)$ predstavlja ocenu matrice kovarijanse greške estimacije parametara u koraku N , tako da trag ove matrice određuje ukupnu srednju kvadratnu grešku estimacije po svim parametrima u datom koraku [23,63,68] Na taj način, ukoliko je početna procena $\hat{\theta}_o$ tačnije određena faktor c^2 u izrazu za $P(0)$ treba smanjiti i usvojiti $c^2 \approx 1$.

Primetimo, na kraju, da kod FIR i ARX reprezentacije modela regresioni vektor $\varphi(k)$ sadrži samo zakašnjene ulazne (FIR) ili zakašnjene ulazne i izlazne signale (ARX), odnosno

$$\varphi_{FIR}^T(k) = \{u(k-1) \dots u(k-n_b)\} \quad (4.140)$$

$$\varphi_{ARX}^T(k) = \{-y(k-1) \dots -y(k-n_a)u(k-1) \dots u(k-n_b)\} \quad (4.141)$$

Uz navedeno, zbog osobine kauzalnosti svi signali biće jednaki nuli za negativne indekse vremena, odnosno $u(i) = y(i) = 0$ za $i < 0$, pri čemu je zbog osobine striktnosti kauzalnosti $y(0) = 0$.

4.2.1.5. Pseudolinearni rekurzivni metod najmanjih kvadrata

U slučaju da usvojena reprezentacija modela odgovara jednačini pseudolinearne regresije (ARMAX struktura modela), tada je na osnovu (4.45) i (4.48)

$$y(k) = \varphi^T(k, \theta) \theta + e(k), \quad k = 1, 2, \dots, N \quad (4.142)$$

te je za numeričku minimizaciju srednje-kvadratnog optimizacionog kriterijuma u više koraka

$$J_N(\theta) = \frac{1}{2N} E^T(N, \theta) E(N, \theta) = \frac{1}{2N} \sum_{i=1}^N \varepsilon^2(i, \hat{\theta}_N(k)); \quad E(N, \theta) = Y(N) - \phi(N, \theta) \theta \quad (4.143)$$

gde je $\theta = \hat{\theta}_N(k)$ procena vektora parametara, na bazi mernog uzorka dužine N , u k -tom koraku (iteraciji), pri čemu je uvedena vektorsko-matrična notacija

$$Y(N) = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}; \quad \phi(N, \theta) = \begin{bmatrix} \varphi^T(1, \theta) \\ \vdots \\ \varphi^T(N, \theta) \end{bmatrix} \quad (4.144)$$

potrebno koristiti nerekurzivni deterministički Newtonov, tj. Newton-Raphsonov algoritam (4.97),

$$\hat{\theta}_N(k+1) = \hat{\theta}_N(k) - H^{-1}(\hat{\theta}_N(k)) g(\hat{\theta}_N(k)) \quad (4.145)$$

gde je gradijent kriterijuma dat sa

$$g(\hat{\theta}_N(k)) = \nabla_{\theta} J_N(\hat{\theta}_N(k)) = \sum_{i=1}^N \varepsilon(i, \hat{\theta}_N(k)) \nabla_{\theta} \varepsilon(i, \hat{\theta}_N(k)) \quad (4.146)$$

pri čemu je greška predikcije

$$\varepsilon(i, \hat{\theta}_N(k)) = y(i) - \hat{y}(i, \hat{\theta}_N(k)) = y(i) - \varphi^T(i, \hat{\theta}_N(k)) \hat{\theta}_N(k), \quad i = 1, \dots, N \quad (4.147)$$

Na taj način se dobija

$$\nabla_{\theta} \varepsilon(i, \hat{\theta}_N(k)) = -\nabla_{\theta} \hat{y}(i, \hat{\theta}_N(k)) = -\psi(i, \hat{\theta}_N(k)) \quad (4.148)$$

odakle sledi

$$\begin{aligned} \nabla_{\theta} J_N(\hat{\theta}_N(k)) &= -\sum_{i=1}^N \varepsilon(i, \hat{\theta}_N(k)) \psi(i, \hat{\theta}_N(k)) \\ &= -\sum_{i=1}^{N-1} \varepsilon(i, \hat{\theta}_N(k)) \psi(i, \hat{\theta}_N(k)) - \varepsilon(N, \hat{\theta}_N(k)) \psi(N, \hat{\theta}_N(k)) \\ &= \nabla J_{N-1}(i, \hat{\theta}_N(k)) - \varepsilon(N, \hat{\theta}_N(k)) \psi(N, \hat{\theta}_N(k)) \end{aligned} \quad (4.149)$$

Pretpostavljajući, dalje, da je procena $\hat{\theta}_N(k)$ bliska minimumu θ^* usvojenog kriterijuma, može se pretpostaviti da je

$$\nabla J_{N-1}(\hat{\theta}_N(k)) \approx 0 \quad (4.150)$$

odakle sledi

$$g(\hat{\theta}_N(k)) = \nabla J_N(\hat{\theta}_N(k)) \approx -\varepsilon(N, \hat{\theta}_N(k)) \psi(N, \hat{\theta}_N(k)) \quad (4.151)$$

Izvođenje rekurzivne verzije algoritma zahteva da se uvedu dodatne aproksimacije. Naime, ako se zanemari zavisnost regresionog vektora $\varphi(N, \hat{\theta}_N(k))$ od $\hat{\theta}_N(k)$, to će rezultovati u aproksimaciju greške predikcije

$$\varepsilon(N, \hat{\theta}_N(k)) = y(N) - \hat{y}(N, \hat{\theta}_N(k)) = y(N) - \varphi^T(N, \hat{\theta}_N(k)) \hat{\theta}_N(k) \quad (4.152)$$

sa izrazom

$$\varepsilon(N, \hat{\theta}_N(k)) \approx \varepsilon(N) = y(N) - \hat{y}(N) = y(N) - \varphi^T(N) \hat{\theta}_N(k) \quad (4.153)$$

Na taj način, može se pisati

$$\psi(N, \hat{\theta}_N(k)) = \nabla_{\theta} \hat{y}(N, \hat{\theta}_N(k)) \approx \nabla_{\theta} \hat{y}(N) = \varphi(N) \quad (4.154)$$

odakle sledi

$$g(\hat{\theta}_N(k)) \approx \varphi(N) \varepsilon(N) \quad (4.155)$$

Na sličan način se za Gauss-Newtonovu aproksimaciju Hessiana može usvojiti

$$\begin{aligned}
 H(\hat{\theta}_N(k)) &= \nabla^2 J_N(\hat{\theta}_N(k)) \approx \sum_{i=1}^N \nabla \varepsilon(i, \hat{\theta}_N(k)) \nabla^T \varepsilon(i, \hat{\theta}_N(k)) \\
 &= \sum_{i=1}^N \psi(i, \hat{\theta}_N(k)) \psi^T(i, \hat{\theta}_N(k))
 \end{aligned} \tag{4.156}$$

odakle sledi

$$\begin{aligned}
 \nabla^2 J_N(\hat{\theta}_N(k)) &= \sum_{i=1}^{N-1} \psi(i, \hat{\theta}_N(k)) \psi^T(i, \hat{\theta}_N(k)) + \psi(N, \hat{\theta}_N(k)) \psi^T(N, \hat{\theta}_N(k)) \\
 &= \nabla^2 J_{N-1}(\hat{\theta}_N(k)) + \psi(N, \hat{\theta}_N(k)) \psi^T(N, \hat{\theta}_N(k))
 \end{aligned} \tag{4.157}$$

Dobijanje rekurzivne varijante ovog izraza zahteva odgovarajuće aproksimacije. Naime, pretpostavljajući da su procene parametara blizu traženog minimuma kriterijuma, može se dodatno usvojiti aproksimacija da su procene parametara zasnovane na mernom skupu dužine N približno jednake procenama parametara zasnovanim na mernoj populaciji dužine $N-1$, odakle sledi

$$\hat{\theta}_N(k) \approx \hat{\theta}_{N-1}(k) \tag{4.158}$$

na osnovu čega se zaključuje da je

$$\nabla^2 J_{N-1}(\hat{\theta}_N(k)) \approx \nabla^2 J_{N-1}(\hat{\theta}_{N-1}(k)) \tag{4.159}$$

tako da se Hessijan matrica, na osnovu (4.154), (4.157) i (4.159), može aproksimirati sa izrazom

$$\begin{aligned}
 H(\hat{\theta}_N(k)) &= \nabla^2 J_N(\hat{\theta}_N(k)) \approx \nabla^2 J_{N-1}(\hat{\theta}_{N-1}(k)) + \varphi(N) \varphi^T(N) \\
 &= H(\hat{\theta}_{N-1}(k)) + \varphi(N) \varphi^T(N)
 \end{aligned} \tag{4.160}$$

Usvajajući dalje oznake

$$\begin{aligned}
 \hat{\theta}(N) &= \hat{\theta}_N(k), & \hat{\theta}(N-1) &= \hat{\theta}_{N-1}(k), & H(\hat{\theta}_N(k)) &= \bar{P}(N), \\
 H(\hat{\theta}_{N-1}(k)) &= \bar{P}(N-1)
 \end{aligned}$$

dolazi se do rekurzivne verzije više-koračne iterativne nerekurzivne Gauss-Newtonove metode

$$\begin{aligned}
 \hat{\theta}(N) &= \hat{\theta}(N-1) + \bar{P}^{-1}(N) \varphi(N) \varepsilon(N), \quad N=1,2,\dots \\
 \varepsilon(N) &= y(N) - \varphi^T(N) \hat{\theta}(N-1) \\
 \bar{P}(N) &= \bar{P}(N-1) + \varphi(N) \varphi^T(N)
 \end{aligned} \tag{4.161}$$

ili uvodeći smenu

$$\bar{P}^{-1}(N) = P(N) \quad (4.162)$$

i koristeći lemu o inverziji matrice

$$\begin{aligned} \hat{\theta}(N) &= \hat{\theta}(N-1) + K(N)\varepsilon(N), \quad N = 1, 2, \dots \\ K(N) &= P(N)\varphi(N) \\ \varepsilon(N) &= y(N) - \varphi^T(N)\hat{\theta}(N-1) \end{aligned} \quad (4.163)$$

$$P(N) = P(N-1) - P(N-1)\varphi(N)\left[\varphi^T(N)P(N-1)\varphi(N) + 1\right]^{-1}\varphi^T(N)P(N-1)$$

Izvedeni algoritam predstavlja pseudolinearni rekurzivni metod najmanjih kvadrata (engleski Pseudo Linear Recursive Least Square, ili skraćeno PLRLS). [35,36,82] Slično kao i rekurzivni linearni metod najmanjih kvadrata RLLS i ovaj algoritam je izveden iz odgovarajuće nerekurzivne varijante na osnovu izvesnih aproksimacija koje su uobičajene za takva izvođenja. Kao i kod RLLS algoritma, startovanje algoritma zahteva da se usvoji početna pogađanja $\hat{\theta}(0)$ i $P(0)$, na identičan način kao kod RLLS metode. Posebno, za ARMAX reprezentaciju modela, linearna aproksimacija pseudo-linearnog regresionog vektora data je sa

$$\varphi^T(N) = \{-y(N-1) \dots -y(N-n_a) u(N-1) \dots u(N-n_b) \varepsilon(N-1) \dots \varepsilon(N-n_c)\} \quad (4.164)$$

pri čemu je zbog uslova kauzalnosti $u(i) = y(i) = \varepsilon(i) = 0$ za $i < 0$, kao i zbog uslova striktnosti kauzalnosti $y(0) = \varepsilon(0) = 0$, dok vrednost $u(0)$ može postojati. Sofisticiranije varijante PLRLS algoritma mogu se naći u literaturi posvećenoj parametarskoj identifikaciji [22,36,63,82,84].

4.2.2 Nelinearni algoritmi estimacije parametara

4.2.2.1 Rekurzivni stohastički Newton-Raphsonov algoritam parametarske identifikacije

U cilju izvođenja nelinearne rekurzivne forme procedure koja minimizira nelinearni usrednjeni indeks performanse, posmatrajmo, umesto kriterijuma optimalnosti, empirijski kriterijum (indeks performanse) (4.5b), odnosno

$$J_i(\hat{\theta}) = i^{-1} \sum_{k=1}^i F[v(k, \hat{\theta})]. \quad (4.165)$$

Podrazumeva se da je izabrano dovoljno veliko i koje obezbeđuje da $J_i(\cdot)$ konvergira, u skladu sa zakonom velikih brojeva, ka $J(\cdot)$ datom u (4.5a). Za rešavanje sistema nelinearnih jednačina koje se javljaju kao rezultat uslova optimalnosti $\partial J_i(\theta)/\partial \theta = 0$, gde $\partial J(\cdot)/\partial \theta$ označava operator parcijalnog izvoda, može se primeniti Newton-Raphsonov metod. Slično kao u prethodnom poglavlju,

linearizujući $J_i(\theta)$ u okolini prethodne procene $\hat{\theta}(i-1)$ i primenjujući razvoj u Taylorov red, dobija se

$$J_i(\theta) = J_i[\hat{\theta}(i-1)] + \frac{\partial J_i[\hat{\theta}(i-1)]}{\partial \theta} [\theta - \hat{\theta}(i-1)] + \frac{1}{2} [\theta - \hat{\theta}(i-1)]^T \cdot \frac{\partial^2 J_i[\hat{\theta}(i-1)]}{\partial \theta^2} [\theta - \hat{\theta}(i-1)] + O(\|\theta - \hat{\theta}(i-1)\|^2) \quad (4.166)$$

gde je

$$\lim_{\|x\| \rightarrow \infty} \frac{O(\|x\|)}{\|x\|} = 0 \quad (4.167)$$

a $\|\cdot\|$ označava Euklidsku normu. Željena vrednost $\hat{\theta}(i)$ dobija se rešavanjem vektorske jednačine

$$\partial J_i(\hat{\theta}(i)) / \partial \theta = 0 \quad (4.168)$$

iz čega se dobija, nakon diferenciranja relacije (4.168) i korišćenja pravila (4.60) i (4.69) za diferenciranje skalara po vektoru

$$\hat{\theta}(i) = \hat{\theta}(i-1) - \left\{ i \frac{\partial^2 J_i[\hat{\theta}(i-1)]}{\partial \theta^2} \right\}^{-1} \left\{ i \frac{\partial J_i[\hat{\theta}(i-1)]}{\partial \theta} \right\} + O(\|\hat{\theta}(i) - \hat{\theta}(i-1)\|) \quad (4.169)$$

Pretpostavimo, dalje, da su zadovoljene sledeće hipoteze:

H1. Procena $\hat{\theta}(i)$ je u okolini procene $\hat{\theta}(i-1)$, što prouzrokuje:

- a) $O(\|\hat{\theta}(i) - \hat{\theta}(i-1)\|) = 0$
- b) $\frac{\partial^2 J_i[\hat{\theta}(i)]}{\partial \theta^2} \approx \frac{\partial^2 J_i[\hat{\theta}(i-1)]}{\partial \theta^2}$ za i dovoljno veliko.

H2. Procena $\hat{\theta}(i-1)$ je optimalna na koraku $i-1$, što daje $\partial J_i[\hat{\theta}(i-1)] / \partial \theta = 0$. Nadalje, iz (4.165) sledi

$$i J_i(\hat{\theta}) = (i-1) J_{i-1}(\hat{\theta}) + F[v(i, \hat{\theta})] \quad (4.170)$$

Uzimajući $\hat{\theta} = \hat{\theta}(i-1)$ i dva puta diferencirajući (4.170), uz korišćenje hipoteze H1(b), kao i činjenice da je zbog (4.32) ili (4.38) (podprimeri sa FIR ili IIR strukture u obliku ARX forme) $\partial v(i, \theta) / \partial \theta = -Z(i)$, gde je oznaka $\varphi(i)$ zamenjena sa $Z(i)$, dobija se

$$i \frac{\partial^2 J_i [\hat{\theta}(i-1)]}{\partial \theta^2} = (i-1) \frac{\partial^2 J_{i-1} [\hat{\theta}(i-1)]}{\partial \theta^2} + \psi' [v(i, \hat{\theta}(i-1))] Z(i) Z^T(i) \quad (4.171)$$

pri čemu je $\psi(\cdot) = F(\cdot)$. Pored toga, uzimajući u obzir i hipotezu H2, dobija se iz (4.170), nakon diferenciranja i zamene θ sa $\hat{\theta}(i-1)$,

$$i \frac{\partial^2 J_i [\hat{\theta}(i-1)]}{\partial \theta} = -Z(i) \psi [v(i, \hat{\theta}(i-1))]. \quad (4.172)$$

Uvodeći notaciju

$$R(i) = i \frac{\partial^2 J_i [\hat{\theta}(i-1)]}{\partial \theta^2} \quad (4.173)$$

i koristeći (4.171) i (4.172), relacija (4.169) dobija formu

$$\hat{\theta}(i) = \hat{\theta}(i-1) + R(i)^{-1} Z(i) \psi [v(i, \hat{\theta}(i-1))] \quad (4.174)$$

$$v(i, \hat{\theta}) = y(i) - Z^T(i) \hat{\theta} \quad (4.175)$$

$$R(i) = R(i-1) + \psi' [v(i, \hat{\theta}(i-1))] Z(i) Z^T(i). \quad (4.176)$$

Veoma često je u upotrebi matrica pojačanja $\Gamma(i) = R^{-1}(i)$. Korišćenjem leme o inverziji matrica (4.131) i (4.132), iz (4.174) i (4.175) se dobija

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \Gamma(i) Z(i) \psi [v(i, \hat{\theta}(i-1))] \quad (4.177)$$

$$v(i, \hat{\theta}) = y(i) - Z^T(i) \hat{\theta} \quad (4.178)$$

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1) Z(i) Z^T(i) \Gamma(i-1) \psi' [v(i, \hat{\theta}(i-1))]}{1 + \psi' [v(i, \hat{\theta}(i-1))] Z^T(i) \Gamma(i-1) Z(i)}. \quad (4.179)$$

Relacije (4.176) - (4.179) definišu rekursivni nelinearni metod tipa Newton-Raphsonove stohastičke procedure, gde je $\psi(\cdot)$ nelinearna transformaciona greška predikcije (reziduala merenja), koja se u statističkoj literaturi naziva i funkcija uticaja (engleski influence function) [33,39,40].

Pošto funkcija $\psi'(\cdot)$ u opštem slučaju nije merljiva u okolini tačke minimuma kriterijuma θ^* , kada je $v(i, \theta) \approx v(i, \theta^*) = \xi(i)$, gde je $\xi(i)$ beli stohastički proces ($\xi(i)$ odgovara $e(i)$ u poglavlju 4.1), posmatraćemo dalje nešto izmenjenu formu algoritma parametarske identifikacije (4.176) - (4.179). U ovako modifikovanoj proceduri drugi izvod empirijskog funkcionala (4.165) aproksimira se na sledeći način. Naime, kriterijumska funkcija (funkcija gubitaka ili rizika), koja radi kao generator stohastičke procedure, ima oblik usrednjenog indeksa performanse (4.5b), odnosno

$$J(\theta) = E\{F(v(i, \theta))\} \quad (4.180)$$

gde je greška predikcije $v(i, \theta) = y(i) - Z^T(i)\theta$. (Oznaka $v(\cdot)$ odgovara oznaci $\varepsilon(i)$ u poglavlju 4.1). Imajući u vidu poslednju relaciju, dalje se dobija

$$\frac{\partial J(\theta)}{\partial \theta} = E\left\{\psi(v(i, \theta)) \frac{\partial v(i, \theta)}{\partial \theta}\right\} = -E\{\psi(v(i, \theta))Z(i)\} \quad (4.181)$$

$$\begin{aligned} \frac{\partial^2 J(\theta)}{\partial \theta^2} &= E\left\{\psi'(v(i, \theta)) \frac{\partial v(i, \theta)}{\partial \theta} \frac{\partial v(i, \theta)^T}{\partial \theta}\right\} + E\left\{\psi(v(i, \theta)) \frac{\partial^2 v(i, \theta)}{\partial \theta^2}\right\} \\ &\approx E\{\psi'(v(i, \theta))Z(i)Z^T(i)\} \end{aligned} \quad (4.182)$$

gde je $\psi(\cdot) = F'(\cdot)$. Prilikom izvođenja relacije (4.182) uzeto je u obzir da je $\partial^2 v(i, \theta)/\partial \theta^2 = 0$. Ukoliko algoritam konvergira ka minimumu funkcionala (4.180), za dovoljno veliko i je $v(i, \theta) \approx \xi(i)$, gde je $\xi(i)$ beli diskretni slučajni proces, te se može pisati

$$\frac{\partial^2 J(\theta)}{\partial \theta^2} \approx E\{\psi'(v(i, \theta))\}E\{Z(i)Z^T(i)\} = \alpha E\{Z(i)Z^T(i)\}; \quad \alpha = E\{\psi'(\xi(i))\} \quad (4.183)$$

Na osnovu ove činjenice i aproksimacije matematičkog očekivanja u (4.183) aritmetičkom sredinom, dalje se dobija

$$\frac{\partial^2 J(\theta)}{\partial \theta^2} = \frac{\alpha}{i} \sum_{k=1}^i Z(k)Z^T(k) \quad (4.184)$$

odakle sledi, s obzirom da je $\partial^2 J(\theta)/\partial \theta^2 \approx \partial^2 J_i(\theta)/\partial \theta^2$, gde je $J_i(\cdot)$ definisano izrazom (4.5b), modifikovana relacija (4.177)

$$i \frac{\partial^2 J_i(\theta)}{\partial \theta^2} = (i-1) \frac{\partial^2 J_i(\theta)}{\partial \theta^2} + \alpha Z(i)Z^T(i) \quad (4.185)$$

Ako se u relaciji (4.185) usvoji $\theta = \hat{\theta}(i-1)$ i uvede smena (4.173), dalje sledi modifikovana relacija (4.175)

$$R(i) = R(i-1) + \alpha Z(i)Z^T(i) \quad (4.186)$$

koja se nakon uvođenja smene

$$\Gamma(i) = R^{-1}(i) \quad (4.187)$$

i primene leme o inverziji matrice (4.131) i (4.132), svodi na oblik

$$\Gamma^{-1}(i) = \Gamma^{-1}(i-1) + \alpha Z(i)Z^T(i) \quad (4.188a)$$

odnosno

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1)Z(i)Z^T(i)\Gamma(i-1)}{\alpha^{-1} + Z^T(i)\Gamma(i-1)Z(i)} \quad (4.188b)$$

Na taj način, modifikovani algoritam parametarske identifikacije dobija konačnu formu

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \Gamma(i)Z(i)\nu(i, \hat{\theta}(i-1)) \quad (4.189)$$

$$\nu(i, \hat{\theta}(i-1)) = y(i) - \hat{\theta}^T(i-1)Z(i) = y(i) - Z^T(i)\hat{\theta}(i-1) \quad (4.190)$$

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1)Z(i)Z^T(i)\Gamma(i-1)}{\alpha^{-1} + Z^T(i)\Gamma(i-1)Z(i)} \quad (4.191)$$

gde je vektor observacija, tj. ulazno-izlaznih merenja $Z^T(i) = \varphi^T(i)$ definisan relacijom (4.54) za FIR strukturu modela, odnosno relacijom (4.55) za IIR strukturu u obliku ARX reprezentacije, respektivno, dok je vektor parametara dat sa relacijom (4.28) za FIR, odnosno relacijom (4.34) za ARX strukturu modela, respektivno.

4.2.2.1 Rekurzivni nelinearni algoritam parametarske identifikacije tipa stohastičkog gradijenta

Tipičan problem stohastičke aproksimacije može se formulisati na sledeći način. Neka je $\{e(i)\}$ niz identički raspodeljenih slučajnih promenljivih. Neka je, dalje, data funkcija $Q(e(i), \chi)$, te se traži rešenje jednačine

$$E\{Q(\chi, e(i))\} = f(\chi) = 0 \quad (4.192)$$

gde $E\{\cdot\}$ označava matematičko očekivanje u odnosu na slučajne varijable $e(i)$. U opštem slučaju funkcija raspodele verovatnoće slučajnih varijabli $e(i)$ nije poznata, kao ni oblik funkcije $Q(\chi, e)$, ali se vrednosti funkcije $Q(\chi, e)$ mogu odrediti ili izmeriti za svako izabrano χ . Na taj način, u diskretnom vremenskom trenutku čiji je indeks i korisnik bira vrednost χ i dobija vrednost $Q(\chi, e(i))$. Primenjujući ovako formulisani zadatak na razmatrni problem parametarske identifikacije, koji predstavlja proceduru minimizacije kriterijumske funkcije (4.5a), odnosno

$$J(\theta) = E\{F(\nu(i, \theta))\}; \quad \nu(i, \theta) = y(i) - Z^T(i)\theta \quad (4.193)$$

koja se svodi na rešavanje sistema jednačina (4.6), tj.

$$\nabla_{\theta} J(\theta) = \frac{\partial J(\theta)}{\partial \theta} = E\{\psi(\nu(i, \theta))\nabla_{\theta} \nu(i, \theta)\} = E\{\psi(\nu(i, \theta))Z(i)\} = 0 \quad (4.194)$$

gde je $\psi(\cdot) = F'(\cdot)$, zaključuje se da je (4.193) specijalan slučaj (4.191) za koji je

$$\chi = \theta, \quad e^T(i) = \{y(i)Z^T(i)\} \text{ i } Q(\chi, e(i)) = Z(i)\nu(i, \theta).$$

U razmatranom slučaju $e(i)$ je merljivo, a $Q(\chi, e)$ je poznata funkcija argumenata χ i e , dok je funkcija raspodele verovatnoće vektorske slučajne veličine $e(i)$ nepoznata.

Postavljeni problem se sada svodi na određivanje niza vrednosti $\chi(i)$, $i=1,2,\dots$, observacije odgovarajuće vrednosti $Q(\chi(i), e(i))$ i izvođenje zaključka o rešenju razmatrane jednačine (4.192). Trivijalno rešenje sastoji se u fiksiranju vrednosti χ , određivanje velikog broja vrednosti $Q(\chi, e(i))$ za dato χ , čime se dobija adekvatna procena funkcije $f(\chi)$, kao i ponavljanju postupka za različite vrednosti varijable χ dok se ne odredi rešenje posmatrane jednačine. Međutim, mnogo je efikasnije da se vrednost varijable χ menja sa svakom novom observacijom, kako se ne bi ulagao nepotreban napor da se tačno odredi funkcija $f(\chi)$ i za one vrednosti argumenta χ koje su daleko od rešenja posmatrane jednačine (4.192), tj. $f(\chi)=0$. Robins i Monro su predložili sledeći postupak za rekurzivno traženje rešenja jednačine kako vreme evoluira [63]

$$\hat{\chi}(i) = \hat{\chi}(i-1) + \gamma(i)Q(\hat{\chi}(i-1), e(i)) \quad (4.195)$$

gde je $\{\gamma(i)\}$ sekvence pozitivnih skalarnih promenljivih koje konvergiraju ka nuli. Konvergenijska svojstva ovog algoritma razmatrana su od strane Robbins-a i Monro-a (1951), Bluma (1951) i Dvoretzky-a (1956) [63]. Može se pokazati da $\hat{\chi}(i)$ konvergira ka rešenju jednačine $f(\chi)=0$ pod izvesnim pretpostavkama, od kojih je tipična pretpostavka da je $\{e(i)\}$ niz nezavisnih slučajnih vektora. Primenjujući ovaj algoritam na razmatrani problem, dobija se algoritam parametarske identifikacije

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \gamma(i)Z(i)\nu(v(i, \hat{\theta}(i-1))) \quad (4.196)$$

Ovaj algoritam poznat je u literaturi kao metod stohastičke aproksimacije ili metod stohastičkog gradijenta [63]. U literaturi posvećenoj adaptivnoj obradi signala ovaj algoritam se naziva i metod minimalne srednjekvadratne greške, a skraćeno se označava sa LMS (po odgovarajućem anglosaksonskom terminu Least Mean-Square Error) [38,48,94]. Niz brojeva $\{\gamma(i)\}$ naziva se i sekvenca pojačanja, a uobičajeni izbori za ovu sekvencu su [3]

- a) $\gamma(i) = \gamma_0$ (konstanta)
- b) $\gamma(i) = \gamma_0 / i$ (opadajuća sekvenca)
- c) $\gamma(i) = \gamma_0 / \|Z(i)\|^2$; (normalizovana sekvenca)
- d) $\gamma(i) = \left[\sum_{k=1}^i \|Z(k)\|^2 \right]^{-1}$; (normalizovana opadajuća sekvenca)

Prednost normalizovane sekvence pojačanja je što je takav algoritam invarijantan na skaliranje izlaznog signala $y(i)$ i regresionog vektora $Z(i)$. U slučaju opadajuće normalizovane sekvence pojačanja algoritam parametarske identifikacije ima oblik

$$\hat{\theta}(i) = \hat{\theta}(i-1) + r^{-1}(i)Z(i)v(i, \hat{\theta}(i-1)) \quad (4.197)$$

$$v(i, \hat{\theta}(i-1)) = y(i) - Z^T(i)\hat{\theta}(i-1) \quad (4.198)$$

$$Z^T(i) = [-u(i-1) \dots -y(i-n), u(i-1) \dots u(i-m)] \quad (4.199)$$

$$r(i) = r(i-1) + \|Z(i)\|^2 \quad (4.200)$$

Karakteristično za stohastičke rekurzivne procedure sa skalarnim faktorom pojačanja, kao što je algoritam (4.196) – (4.199) je njihova jednostavna računarska implementacija, ali i relativno mala brzina konvergencije. Zato je od izuzetne važnosti da se izvrši sinteza takvih rekurentnih stohastičkih procedura koje poseduju znatno veću brzinu konvergencije od algoritma stohastičke aproksimacije (stohastičkog gradijenta). Rešenje se nalazi u uvođenju matrične sekvence pojačanja, što će rezultovati u algoritme tipa stohastičke Newton-Raphsonove procedure, čija je jedna varijanta već razmatrana u prethodnom poglavlju. Do ovakvih algoritama može se doći i pogodnom modifikacijom bazičnog algoritma stohastičkog gradijenta. Da bismo došli do ovakvog rezultata, podsetimo se da je naš razmatrani problem minimizacioni zadatak (4.4) i (4.5), odnosno

$$\min_{\theta} J(\theta); \quad J(\theta) = E\{F(\theta, e(i))\} \quad (4.201)$$

gde je $e^T(i) = [y(i)Z^T(i)]$. Ako je

$$-\frac{\partial J(\theta)}{\partial \theta} = f(\theta) \quad (4.202)$$

i ako se gradijent

$$-\frac{\partial}{\partial \theta} F(\theta, e(i)) = Q(\theta, e(i)) \quad (4.203)$$

može odrediti ili izmeriti za bilo koju vrednost θ , tada se postavljeni minimizacioni problem svodi na rešavanje jednačine oblika (4.192), odnosno

$$-\frac{\partial J(\theta)}{\partial \theta} = f(\theta) = E\{Q(\theta, e(i))\} = 0 \quad (4.204)$$

gde je pretpostavljeno da se redosled operacija usrednjavanja (nalaženja matematičkog očekivanja) i diferenciranja može zameniti. Na taj način, algoritam stohastičke aproksimacije može se posmatrati kao algoritam minimizacije kriterijumske funkcije (4.200), a podešavanje parametara vrši se u pravcu izračunatog

negativnog gradijenta kriterijuma funkcije $F(\theta, e(i))$. Na taj način, pravac podešavanja procena parametara odgovaraće u srednjem negativnom gradijentu indeksa performanse $J(\theta)$ u (4.201). Iz navedenog razloga algoritam se naziva metod stohastičkog gradijenta [63]. Ovaj algoritam predstavlja stohastičku verziju gradijentnog metoda najbržeg spusta (odgovarajući anglosaksonski termin je steepest descent gradient method) koji se koristi za minimizaciju determinističkih funkcija i ima oblik (4.101), odnosno

$$\theta_{i+1} = \theta_i - \gamma_i \frac{\partial J(\chi_i)}{\partial \chi} \quad (4.205)$$

gde je γ_i pozitivan skalar, koji je odabran na pogodan način, a indeks i označava i -tu iteraciju, a ne vremenski trenutak. Glavni nedostatak ovakvog metoda je da sporo konvergira, pogotovo u blizini tačke minimuma, te Newton-Raphsonov algoritam (4.96) umesto podešavanja parametara u smeru negativnog gradijenta $-\partial J(\theta)/\partial \theta$ koristi mnogo efikasniji modifikovani pravac podešavanja parametara $-H^{-1}g$, gde je $g = \partial J(\theta)/\partial \theta$ gradijentni vektor, a $H = \partial^2 J(\theta)/\partial \theta^2$ je Hessian (matrica drugih izvoda). Newton-Raphsonova iterativna procedura ima oblik (4.96), odnosno

$$\theta_{iH} = \theta_i - \left[\frac{\partial^2 J(\theta_i)}{\partial \theta^2} \right]^{-1} \frac{\partial J(\theta_i)}{\partial \theta} \quad (4.206)$$

i konvergiraće u jednom koraku ka minimumu kriterijuma $J(\theta)$ ukoliko je to kvadratna funkcija argumenta θ . Pošto se u okolini tačke minimuma kriterijumska funkcija može aproksimirati dovoljno tačno kvadratnom funkcijom, ovaj algoritam radiće vrlo efikasno u okolini minimuma kriterijuma. Daleko od tačke minimuma algoritam može da bude neefikasan ili čak da divergira. Da bi se ovo izbeglo, Hessian se uobičajeno aproksimira nekom matricom koja će garantovano biti pozitivno-definitna. Slično poboljšanje može se očekivati i u slučaju algoritma stohastičke aproksimacije. Pretpostavimo da se za svako θ može odrediti aproksimacija Hessiana $\bar{J}''(\theta, e^i)$, gde e^i označava da ova aproksimacija može zavisiti od svih prethodnih realizacija $e^i = \{e(i), e(i-1), \dots, e(1)\}$. Tada stohastički Newton-Raphsonov algoritam predstavlja prirodnu modifikaciju algoritma stohastičkog gradijenta (4.194) i (4.195) i ima oblik

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \gamma(i) \left[\bar{J}''(\hat{\theta}(i-1), e^i) \right]^{-1} Q(\hat{\theta}(i-1), e(i)) \quad (4.207)$$

U našem razmatranom problemu parametarske identifikacije gradijent je dat relacijom

$$\frac{\partial J(\theta)}{\partial \theta} = E\{\psi(v(i, \theta))Z(i)\} \quad (4.208)$$

dok je Hessian definisan izrazom

$$\frac{\partial^2 J(\theta)}{\partial \theta^2} = E\{\psi'(v(i, \theta))Z(i)Z^T(i)\} \quad (4.209)$$

Na osnovu poslednjeg izraza, Hessian se može odrediti kao rešenje jednačine

$$E\{\psi'(v(i, \theta))Z(i)Z^T(i) - \bar{R}\} = 0 \quad (4.210)$$

a za iterativno rešavanje ove jednačine može se dalje koristiti algoritam stohastičke aproksimacije (4.194-4.196).

Ako se u originalnom algoritmu stohastičkog gradijenta (4.192) usvoji

$$\hat{\chi} = \bar{R}, \quad e^T(i) = [v(i, \hat{\theta}(i-1)), Z^T(i)], \quad Q(\hat{\chi}, e(i)) = \psi'(v(i, \hat{\theta}(i-1)))Z(i)Z^T(i) - \bar{R}$$

ovaj algoritam dobija oblik

$$\bar{R}(i) = \bar{R}(i-1) + \gamma(i) [\psi'(v(i, \hat{\theta}(i-1)))Z(i)Z^T(i) - \bar{R}(i-1)] \quad (4.211a)$$

odnosno

$$\bar{R}(i) = [1 - \gamma(i)]\bar{R}(i-1) + \gamma(i) \psi'(v(i, \hat{\theta}(i-1)))Z(i)Z^T(i) \quad (4.211b)$$

Na taj način, našu procenu Hessiana $\partial^2 J(\theta)/\partial \theta^2$ u trenutku i predstavlja matricu $R(i)$. Posebno za $\gamma(i) = 1/i$ poslednja relacija odgovara izrazu (4.176), s obzirom da je na osnovu (4.173) $R(i) = i\bar{R}(i)$. Koristeći dobijenu procenu gradijenta, nepoznati vektor parametara u trenutku i može se proceniti na osnovu stohastičkog Newton-Raphsonovog algoritma

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \gamma(i) \bar{R}^{-1}(i) Z(i) \psi(v(i, \hat{\theta}(i-1))) \quad (4.212)$$

Kao što je već istaknuto, ovaj algoritam je detaljno razmatran u prethodnom poglavlju (izrazi (4.174) – (4.176)) za opadajuću skalarnu sekvencu $\gamma(i) = 1/i$, kada uzima oblik

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \Gamma(i) Z(i) \psi(v(i, \hat{\theta}(i-1))) \quad (4.213)$$

gde je

$$\Gamma^{-1}(i) = i\bar{R}(i) = R(i) \quad (4.214)$$

Matrica pojačanja algoritma definisana je rekursivno izrazom (4.211), odnosno

$$R(i) = R(i-1) + \psi'(v(i, \hat{\theta}(i-1)))Z(i)Z^T(i) \quad (4.215)$$

ili

$$\Gamma^{-1}(i) = \Gamma^{-1}(i-1) + \psi'(v(i, \hat{\theta}(i-1)))Z(i)Z^T(i) \quad (4.216)$$

odakle se primenom leme o inverziji matrice (4.131) i (4.132) dobija

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1)Z(i)Z^T(i)\Gamma(i-1)\psi'(\nu(i, \hat{\theta}(i-1)))}{1 + Z^T(i)\Gamma(i-1)Z(i)\psi'(\nu(i, \hat{\theta}(i-1)))} \quad (4.217)$$

Algoritam (4.213), (4.217) identičan je rekurzivnoj stohastičkoj proceduri (4.176) - (4.179), koja je izvedena polazeći od metodologije izvođenja Newton-Raphsonove stohastičke metode. Slično kao prilikom izvođenja Newton-Raphsonove rekurzivne stohastičke procedure, u okolini tačke minimuma θ^* kriterijuma $J(\theta)$ u (4.200) može se pretpostaviti da je greška predikcije ili rezidual merenja $\nu(i, \theta) \approx \nu(i, \theta^*) = \xi(i)$, gde je $\xi(i)$ beli diskretni slučajni proces nulte srednje vrednosti, tako da se Hessian (4.208) može aproksimirati izrazom

$$\frac{\partial^2 J(\theta)}{\partial \theta^2} \approx \alpha E\{Z(i)Z^T(i)\}; \quad \alpha = E\{\psi'(\xi(i))\} \quad (4.218)$$

te se na osnovu (4.209) Hessian može odrediti kao rešenje jednačine

$$E\{\alpha Z(i)Z^T(i) - \bar{R}\} = 0 \quad (4.219)$$

Kao i u slučaju relacije (4.209), za iterativno rešavanje jednačine (4.219) može se koristiti algoritam stohastičke aproksimacije (4.192). Ako se u ovom algoritmu usvoji

$$\hat{x} = \bar{R}, \quad e^T(i) = [\nu(i, \hat{\theta}(i-1))Z^T(i)], \quad Q(\hat{x}, e(i)) = \alpha Z(i)Z^T(i) - \bar{R}$$

ovaj algoritam dobija oblik modifikovane jednačine (4.211)

$$\bar{R}(i) = \bar{R}(i-1) + \gamma(i)[\alpha Z(i)Z^T(i) - \bar{R}(i-1)] \quad (4.220)$$

koja se može napisati u ekvivalentnoj formi

$$\bar{R}(i) = [1 - \gamma(i)]\bar{R}(i-1) + \gamma(i)\alpha Z(i)Z^T(i) \quad (4.221a)$$

Uvodeći smenu (4.214), relacija (4.221) dobija oblik:

$$R(i) = R(i-1) + \alpha Z(i)Z^T(i) \quad (4.221b)$$

odnosno nakon primene leme o inverziji matrice (4.131), (4.132)

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1)Z(i)Z^T(i)\Gamma(i-1)}{\alpha^{-1} + Z^T(i)\Gamma(i-1)Z(i)} \quad (4.222)$$

Algoritam dinamičke stohastičke aproksimacije (4.213), (4.222) identičan je rekurzivnom stohastičkom Newton-Raphsonovom algoritmu (4.188) – (4.190). Prilikom izvođenja ovih algoritama pretpostavljeno je da je regresioni vektor $Z(k)$, odnosno $\varphi(i)$ u prethodnom poglavlju, nezavisan od nepoznatog vektora parametara θ koji se estimira. Ovakva pretpostavka podrazumeva FIR reprezentaciju modela sistema ili ARX strukturu modela, koja predstavlja specifičnu IIR reprezentaciju kod koje je aditivni šum merenja $\xi(i)$, odnosno $e(i)$ u prethodnom poglavlju, diskretan

beli slučajni proces nulte srednje vrednosti. U slučaju da se usvoji ARMAX reprezentacija modela sistema, ekvivalentni šum izlaza $v(i)$ u poglavlju 4.1 je korelisani ili obojeni šum, koji se dobija propuštanjem belog šuma $e(i) = \xi(i)$ kroz odgovarajući linearni filter $C(q)/A(q)$ (videti sliku 4.5), dok regresioni vektor $\varphi(i) = Z(i)$ u (4.53) postaje funkcije vektora parametara θ , tj. $\varphi(i, \theta) = Z(i, \theta)$. Slično kao u slučaju algoritma pseudolinearne regresije u poglavlju 4.2.1.5, ovaj vektor se može odrediti na osnovu jednačine (4.164), te se direktno mogu primeniti prethodno izvedeni algoritmi za ARX strukturu modela. Međutim, brojne teorijski i eksperimentalne analize su pokazale da će modelovanje ARX strukturom podataka koji nisu egzaktno generisani ARX procesom, već su generisani, na primer, ARMAX procesom, rezultovati u nezadovoljavajuće procene parametara modela, odnosno tako generisane procene biće pomerene (engleski termin je “biased”) [7,45]. Sa druge strane, glavna prednost ARX strukture je da je PEM estimator, bilo da je zasnovan na kvadratnom ili nelinearnom kriterijumu greške predikcije, relativno jednostavan, s obzirom da regresioni vektor ne zavisi od nepoznatih parametara modela koji se estimiraju. Stoga se postavlja prirodno pitanje kako poboljšati procene parametara izvedene na bazi ARX reprezentacije modela, koji pretpostavlja da je ekvivalentni aditivni šum izlaza diskretan beli stohastički proces, da bi se one efikasno mogle primeniti i u slučaju kada su merni podaci egzaktno generisani ARMAX procesom, kod koga je ekvivalentni šum izlaza korelisani (obojeni) diskretan slučajni proces. Odgovor se nalazi u filtriranju ulazno-izlaznih signala i greške predikcije, o čemu će više reči biti u nastavku [7,45,46,63].

4.2.2.3 Filtriranje greške predikcije

Na osnovu jednačine ARMAX modela (4.40) greška jednačine ili rezidual, zasnovan na poznavanju procene θ nepoznatih koeficijenata polinoma $A(\cdot)$, $B(\cdot)$ i $C(\cdot)$, definisan je izrazima (4.45) – (4.48). Ako se izraz za grešku predikcije izlaza (4.46) zameni u (4.45) i uvede smena $\varepsilon(k|k-1, \theta) = v(k, \theta)$ dobija se sledeći izraz za grešku predikcije (rezidual merenja)

$$C(q)v(i, \theta) = A(q)y(i) - B(q)u(i) \quad (4.223)$$

Ako se, dalje, desna strana ovog izraza linearizuje oko prave vrednosti θ^* vektora θ i primeni metod konačnih razlika (numeričko diferenciranje) dobija se

$$\begin{aligned} C^*(q)v(i, \theta^*) + [C(q) - C^*(q)]v(i, \theta^*) + C^*(q)[v(i, \theta) - v(i, \theta^*)] = \\ = A(q)y(i) - B(q)u(i). \end{aligned} \quad (4.224)$$

gde je $v(i, \theta^*) = \xi(i)$. Na osnovu (4.223), relacija (4.224) može se napisati u obliku

$$v(i, \theta) = A(q)\tilde{y}^*(i) - B(q)\tilde{u}^*(i) - C(q)\tilde{\xi}^*(i) + \xi(i) \quad (4.225)$$

gde su $\tilde{y}^*(i)$, $\tilde{u}^*(i)$ i $\tilde{\xi}^*(i)$ filtrirane vrednosti izlaza, ulaza i poremećaja respektivno, dobijene propuštanjem signala $y(i)$, $u(i)$ i $\xi(i)$ kroz filter $1/C^*(z^{-1}) = D^*(z^{-1})$, tj.

$$\tilde{y}^*(i) = \frac{y(i)}{C^*(q)} = D^*(q)y(i) \quad (4.226a)$$

$$\tilde{u}^*(i) = \frac{u(i)}{C^*(q)} = D^*(q)u(i) \quad (4.226b)$$

$$\tilde{\xi}^*(i) = \frac{\xi(i)}{C^*(q)} = D^*(q)\xi(i) \quad (4.226c)$$

Veza između koeficijenata polinoma $C^*(q)$ i $D^*(q)$ određuje se na osnovu identiteta $D^*(q)C^*(q)=1$, odakle se dobija

$$d_j^* = -\sum_{k=1}^j c_k^* d_{j-k}^* \quad j=1,2,\dots,n, \quad d_0^* = 1. \quad (4.227)$$

Ako se, dalje, iskoriste definicioni izrazi za operator kašnjenja q^{-1} i polinome $A(\cdot)$, $B(\cdot)$, $C(\cdot)$, tj.

$$A(q^{-1}) = 1 + \sum_{i=1}^n a_i q^{-i}, \quad B(q^{-1}) = \sum_{i=1}^n b_i q^{-i}, \quad C(q^{-1}) = 1 + \sum_{j=1}^n c_j q^{-j} \quad (4.228)$$

gde je $q^{-1}\chi(i) = \chi(i-1)$, na osnovu (4.225) sledi

$$v(i, \theta) = y_p^*(i) - Z_f^T(i)\theta \quad (4.229)$$

gde je

$$\theta^T = \{a_1 \dots a_n b_1 \dots b_n c_1 \dots c_n\} \quad (4.230)$$

dok je

$$y_p^*(i) = \tilde{y}^*(i) - \tilde{\xi}^*(i) + \xi(i) \quad (4.231)$$

pseudo-opservacija izlaza, a vektor

$$Z_f^{*T}(i) = \{-\tilde{y}^*(i-1), \dots, -\tilde{y}^*(i-n), \tilde{u}^*(i-1), \dots, \tilde{u}^*(i-n), \tilde{\xi}^*(i-1), \dots, \tilde{\xi}^*(i-n)\}, \quad (4.232)$$

predstavlja vektor filtriranih vrednosti izlaza, ulaza i poremećaja, respektivno. Pošto, na osnovu jednačine modela (4.40), takođe sledi

$$A^*(q)\tilde{y}^*(i) = B^*(q)\tilde{u}^*(i) + C^*(q)\tilde{\xi}^*(i) \quad (4.233)$$

gde simbol zvezdice označava tačne vrednosti parametara $\theta = \theta^*$, odnosno tačne vrednosti koeficijenata polinoma $A(\cdot)$, $B(\cdot)$ i $C(\cdot)$, to dalje implicira

$$\tilde{y}^*(i) = Z_f^{*T}(i)\theta^* + \xi(i) \quad (4.234)$$

Na taj način, greška predikcije ili rezidual merenja (4.229) može se napisati i u obliku

$$v(i, \theta) = Z_f^{*T}(i)(-\theta + \theta^*) + \xi(i). \quad (4.235)$$

Izrazi (4.229) – (4.231) ili (4.235) definišu linearizovanu grešku predikcije ili rezidual, koja predstavlja polaznu osnovu za sintezu rekurzivne stohastičke procedure parametarske identifikacije. Njihovo određivanje, međutim, zahteva poznavanje prave vrednosti θ^* nepoznatog vektora parametara. U praksi navedeni izrazi mogu se odrediti samo približno, polazeći od pretpostavke da su procene nepoznatog vektora parametara bliske njihovoj pravnoj vrednosti, koja rezultuje u sledeći aproksimativni izraz za jednačinu pseudo-opservacija

$$y_p(i) = \tilde{y}(i) - \tilde{v}(i) + v(i) \quad (4.236)$$

gde se, na osnovu jednačina modela (4.30), greška predikcije može aproksimirati izrazom

$$v(i) = y(i) - Z^T(i)\hat{\theta}(i-1) \quad (4.237)$$

pri čemu je regresioni vektor $Z(\cdot)$ definisan relacijom (4.164), odnosno

$$Z^T(i) = \{-y(i-1), \dots, -y(i-n), u(i-1), \dots, u(i-n), v(i-1), \dots, v(i-n)\} \quad (4.238)$$

Na isti način, vektor filtriranih signala definisan je aproksimativno izrazom

$$Z_f^T(i) = \{-\tilde{y}(i-1), \dots, -\tilde{y}(i-n), \tilde{u}(i-1), \dots, \tilde{u}(i-n), \tilde{v}(i-1), \dots, \tilde{v}(i-n)\}. \quad (4.239)$$

Filtrirane vrednosti $\tilde{y}(i)$, $\tilde{u}(i)$ i $\tilde{v}(i)$ izlaza, ulaza i reziduala, respektivno, dobijene su propuštanjem signala $y(i)$, $u(i)$, $v(i)$ kroz filter $\frac{1}{C^{i-1}}(q) = D^{i-1}(q)$, gde indeks $(i-1)$ označava procenu koeficijenata polinoma $C(\cdot)$ u $(i-1)$ -vom koraku (relacije (4.226)). Nepoznati koeficijenti digitalnog filtra $D(\cdot)$ određuju se u svakom koraku iz relacija (4.227), polazeći od poznatih koeficijenata polinoma $C(\cdot)$. Na taj način, linearizovana greška jednačine ili rezidual približno je definisana izrazom

$$v(i, \theta) = y_p(i) - Z_f^T(i)\theta = Z_f^T(i)(\theta^* - \theta) + \xi(i) \quad (4.240)$$

Kao kriterijum-generator procedure parametarske identifikacije može se, dalje, usvojiti indeks performanse (4.5), a za njegovu rekurzivnu minimizaciju može se primeniti Newton-Raphsonova rekurzivna stohastička procedura, odnosno procedura dinamičke stohastičke aproksimacije. Na taj način, rekurzivna stohastička procedura za estimaciju parametara ARMAX strukture modela sistema (IIR struktura sa korelisanim ekvivalentnim aditivnim šumom na izlazu sistema) dobija oblik algoritma za identifikaciju parametara ARX strukture u kome se regresioni vektor zamenjuje sa odgovarajućim regresionim vektorom filtriranih podataka. Polazeći od ove činjenice i primenjujući izraze (4.213) i (4.222) ili (4.188) - (4.190), dalje se može pisati

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \Gamma(i) Z_f(i) \nu(i) \quad (4.241)$$

$$\nu(i) = y(i) - Z^T(i) \hat{\theta}(i-1) \quad (4.242)$$

$$Z^T(i) = \{-y(i-1) \dots -y(i-n) u(i-1) \dots u(i-n) \nu(i-1) \dots \nu(i-n)\} \quad (4.243)$$

$$Z_f^T(i) = \{-\tilde{y}(i-1) \dots -\tilde{y}(i-n) \tilde{u}(i-1) \dots \tilde{u}(i-n) \tilde{\nu}(i-1) \dots \tilde{\nu}(i-n)\} \quad (4.244)$$

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1) Z_f(i) Z_f^T(i) \Gamma(i-1) \alpha}{1 + \alpha Z_f^T(i) \Gamma(i-1) Z_f(i)} \quad (4.245)$$

Filtrirane vrednosti signala u relaciji (4.244) dobijaju se na osnovu relacije (4.226), odnosno

$$Z_f(i) = \frac{1}{C^{i-1}(q^{-1})} Z(i) = D^{i-1}(q^{-1}) Z(i) \quad (4.246)$$

gde je veza između koeficijenta polinoma $C(\cdot)$ i $D(\cdot)$ u koraku $(i-1)$ data sa (4.227), odnosno

$$d_k^{i-1} = -\sum_{j=1}^k c_j^{i-1} d_{k-j}^{i-1}; \quad k=1,2,\dots,n; \quad d_0^{i-1} = 1 \quad (4.247)$$

Relacija (4.246) može se napisati u sledećoj formi koja je pogodna za numeričko izračunavanje filtriranih amplituda relevantnih signala

$$\tilde{y}(k) = y(k) + Y^T(k) d^{i-1} \quad (4.248)$$

$$Y^T(k) = \{y(k-1) \dots y(k-1+n)\} \quad (4.249)$$

$$d^{i-1} = \{d_1^{i-1} d_2^{i-1} \dots d_n^{i-1}\}^T \quad (4.250)$$

$$\tilde{u}(k) = u(k) + U^T(k) d^{i-1} \quad (4.251)$$

$$U^T(k) = \{u(k-1) \dots u(k-1+n)\} \quad (4.252)$$

$$\tilde{\nu}(k) = \nu(k) + E^T(k) d^{i-1} \quad (4.253)$$

$$E^T(k) = \{\nu(k-1) \dots \nu(k-1+n)\} \quad (4.254)$$

gde indeks k uzima vrednosti $k = i-1, \dots, i-n$, a brojač iteracija i se inkrementira u svakom koraku, tj. $i = 1, 2, \dots$.

Ukoliko algoritam konvergira i procene parametara se nalaze u blizini njihovih tačnih vrednosti θ^* , odnosno u blizini minimuma kriterijuma za ocenu kvaliteta ponašanja (indeksa performanse), postupak filtracije signala (4.248) – (4.254) može se pojednostaviti, tako što će se u svakom koraku i usvojiti samo $k = i-1$, odnosno izračunavaće se samo najsvežije filtrirani podaci $\tilde{y}(i-1)$, $\tilde{u}(i-1)$ i $\tilde{v}(i-1)$, umesto celih sekvenci opservacija $\tilde{y}(k)$, $\tilde{u}(k)$ i $\tilde{v}(k)$, $k = i-1, \dots, i-n$.

Alternativa predloženom pristupu je da se ARMAX proces aproksimira ARX procesom, te da se primene algoritmi parametarske identifikacije izvedeni za ARX proces (estimira se samo koeficijenti polinoma $A(\cdot)$ i $B(\cdot)$ a ne i polinoma $C(\cdot)$), bilo linearni ili nelinearni, ali nad filtriranim podacima. Opravdanje za ovakav pristup nalazi se u činjenici da je kod SISO sistema (sistemi sa jednim ulazom i jednim izlazom ili na engleskom Single Input-Single Output ili skraćeno SISO) filtriranje greške predikcije ekvivalentno filtriranju samih ulazno-izlaznih podataka. Zaista, pošto je greška predikcije za opštu strukturu modela definisana izrazom (4.26b), odnosno

$$v(k, \theta) = H^{-1}(q, \theta)[y(k) - G(q, \theta)u(k)] \quad (4.255)$$

to se propuštanjem ovog signala kroz neki filter $F(q)$ dobija filtrirana greška predikcije

$$\tilde{v}(k, \theta) = F(q)v(k, \theta) = H^{-1}(q)[\tilde{y}(k) - G(q, \theta)\tilde{u}(k)] \quad (4.256)$$

gde su filtrirani ulazno-izlazni signali

$$\tilde{y}(k) = F(q)y(k); \quad \tilde{u}(k) = F(q)u(k) \quad (4.257)$$

Ukoliko su opservacije egzaktno generisane ARX modelom, tada filtracija ulazno-izlaznih merenja nije potrebna, što se svodi na uslov

$$F(q)H^{-1}(q, \theta) = 1 \quad (4.258)$$

gde je θ nepoznati vektor parametara koji se estimira. Takođe, pošto je za ARX strukturu na osnovu (4.33)

$$H^{-1}(q, \theta) = A(q, \theta) = 1 + \sum_{i=1}^n a_i q^{-i} \quad (4.259)$$

gde koeficijenti a_i , $i = 1, \dots, n$ predstavljaju elemente vektora θ , iz relacija (4.258) i (4.259) se dobija

$$F(q) = 1 + \sum_{i=1}^n f_i q^{-i} = \frac{1}{A(q, \theta - 1)} = \frac{1}{1 + \sum_{i=1}^n a_i q^{-i}} \quad (4.260)$$

Na taj način, veza između parametara koeficijenata polinoma u (4.260) data je relacijom (4.247), odnosno

$$f_k = -\sum_{j=1}^k a_j f_{k-j}, \quad k=1,2,\dots,n; \quad f_0 = 1 \quad (4.261)$$

Sam algoritam parametarske identifikacije ima oblik rekurzivne stohastičke procedure (4.241) – (4.254) i dat je relacijama

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \Gamma(i)Z(i)v(i) \quad (4.262)$$

$$v(i) = y(i) - Z^T(i)\hat{\theta}(i-1) \quad (4.263)$$

$$Z^T(i) = \{-y(i-1)\dots -y(i-n)u(i-1)\dots u(i-n)\} \quad (4.264)$$

$$\hat{\theta}^T(i-1) = \{a_1^{i-1} \dots a_n^{i-1} b_1^{i-1} \dots b_n^{i-1}\} \quad (4.265)$$

$$f_k^{i-1} = -\sum_{j=1}^k a_j^{i-1} f_{k-j}^{i-1}, \quad k=1,2,\dots,n, \quad f_0^{i-1} = 1, \quad F^{i-1}(q) = \sum_{j=1}^n b^{i-1} q^{-j} \quad (4.266)$$

$$Z_f(i) = F^{i-1}(q)Z(i) = \{-\tilde{y}(i-1)\dots -\tilde{y}(i-1)\tilde{u}(i-1)\dots \tilde{u}(i-r)\} \quad (4.267)$$

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1)Z_f(i)Z_f^T(i)\Gamma(i-1)}{\alpha^{-1} + Z_f^T(i)\Gamma(i-1)Z_f(i)} \quad (4.268)$$

Relacija (4.267) može se dalje napisati u formi relacija (4.248) – (4.252).

Za razliku od algoritma (4.242) – (4.254) gde je estimirani vektor parametara

$$\hat{\theta} = \{\hat{a}_1 \hat{a}_2 \dots \hat{a}_n \hat{b}_1 \hat{b}_2 \dots \hat{b}_n \hat{c}_1 \hat{c}_2 \dots \hat{c}_n\}^T \quad (4.269)$$

algoritmom (4.262) – (4.269) estimiraju se samo koeficijenti polinoma $A(\cdot)$ i $B(\cdot)$ u ARMAX strukturi, isto kao u slučaju ARX strukture, a ne estimiraju se koeficijenti polinoma šuma $C(\cdot)$, odnosno

$$\hat{\theta} = \{\hat{a}_1 \hat{a}_2 \dots \hat{a}_n \hat{b}_1 \hat{b}_2 \dots \hat{b}_n\}^T \quad (4.270)$$

Međutim, pošto polinom $A(\cdot)$ u (4.260) nije poznat, moguć pristup sastoji se u tome da se u prvom koraku proceni ovaj polinom, odnosno odredi početno pogađanje vektora parametara $\hat{\theta}(\cdot)$ u (4.270), koristeći na primer nerekurzivni linearni metod najmanjih kvadrata. Ovakav postupak može da se primeni iterativno, u nekoliko koraka, kako bi se dobio zadovoljavajući kvalitet početne procene vektora parametara.

Inicijalizacija algoritma takođe zahteva i inicijalnu vrednost težinske matrice $\Gamma(\cdot)$, a uobičajeno je da se ova matrica izabere kao $\Gamma(\cdot) = cI$, gde je I jedinična matrica odgovarajućih dimenzija, a c je pogodno odabrana pozitivna konstanta.

U narednom poglavlju biće proučavana svojstva predloženih algoritama kroz teorijsku analizu njihove konvergencije, kao i na bazi simulacija na računaru opšte namene.

5 ANALIZA REKURZIVNIH STOHAISTIČKIH ALGORITMA PARAMETARSKE IDENTIFIKACIJE: KONVERGENCIJA I KONZISTENCIJA PROCENA PARAMETARA I SIMULACIJA

U ovom poglavlju biće analizirana konvergencija rekurzivnih stohastičkih algoritama parametarske estimacije koji su izvedeni u prethodnom poglavlju. Prvo će biti prikazani osnovni teorijski rezultati vezani za konvergenciju procena parametara generisanih linearnim algoritmima. U nastavku će biti dokazana konvergencija nelinearnih rekurzivnih algoritama za estimaciju parametara linearnih diskretnih vremenski invarijantnih stohastičkih sistema. Ova analiza zasnovana je na primeni osnovnih teorijskih rezultata vezanih za konvergenciju specifičnih diskretnih stohastičkih procesa koji se nazivaju martingale [32,62,70,83]. Takođe, biće prikazan i alternativan metod za analizu konvergencije, koji se zasniva na pridruživanju odgovarajuće obične diferencijalne jednačine razmatranoj rekurzivnoj stohastičkoj proceduri i primeni teorije stabilnosti rešenja diferencijalnih jednačina [63,93]. S obzirom da teorijska analiza ima asimptotski karakter, što podrazumeva merni uzorak neograničene dužine, praktične osobine predloženih algoritama na uzorku konačne, čak i male, dužine biće analizirane putem računarske simulacije, uz primenu Matlab programskog paketa [54,57,65,74,92].

5.1 ANALIZA KONVERGENCIJE LINEARNOG ALGORITMA NAJMANJIH KVADRATA

Pretpostavimo, kao u trećem poglavlju, da se realan sistem može egzaktno opisati linearnim diskretnim vremenski-invarijantnim modelom, koji je opisan jednačinom linearne regresije

$$y(k) = \varphi^T(k) \theta_0 + v_0(k) \quad (5.1)$$

gde je θ_0 tačan vektor parametara, $y(k)$ izlazni signal u k -tom diskretnom trenutku, $v_0(k)$ realizacija aditivnog šuma na izlazu u k -tom diskretnom trenutku, a $\varphi(k)$ je regresioni vektor koji sadrži zakašnjene izlaze i ulaze (ARX reprezentacija modela) ili samo zakašnjene ulazne signale (FIR struktura modela). Tada linearni metod najmanjih kvadrata LLS daje sledeću procenu $\hat{\theta}_N$ nepoznatog vektora θ_0 na bazi skupa ulazno-izlaznih podataka $\{u(k), y(k), k = 1, \dots, N\} = \{u(k), y(k)\}_N$ (poglavlje 4)

$$\hat{\theta}_N = \left(\frac{1}{N} \phi^T \phi \right)^{-1} \frac{1}{N} \phi^T Y \quad (5.2)$$

gde je kovarijaciona matrica regresionog vektora dimenzije $p \times 1$

$$R_\phi^p(N) = \frac{1}{N} \sum_{k=1}^N \phi(k) \phi^T(k) = \{\phi(1) \dots \phi(N)\} \begin{Bmatrix} \phi^T(1) \\ \vdots \\ \phi^T(N) \end{Bmatrix} = \phi^T \phi \quad (5.3)$$

dok je vektor

$$\frac{1}{N} \phi^T Y = \frac{1}{N} \{\phi(1) \dots \phi(N)\} \begin{Bmatrix} y(1) \\ \vdots \\ y(N) \end{Bmatrix} = \frac{1}{N} \sum_{k=1}^N \phi(k) y(k) \quad (5.4)$$

Na taj način, LLS procena $\hat{\theta}_N$ vektora parametara θ_0 postaje

$$\hat{\theta}_N = [R_\phi^p(N)]^{-1} \frac{1}{N} \sum_{k=1}^N \phi(k) y(k) \quad (5.5)$$

odakle se, nakon zamene izraza za $y(k)$, dobija

$$\begin{aligned} \hat{\theta}_N &= [R_\phi^p(N)]^{-1} \frac{1}{N} \sum_{k=1}^N \phi(k) [\phi^T(k) \theta_0 + v_0(k)] \\ &= [R_\phi^p(N)]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \phi(k) \phi^T(k) \right] \theta_0 + [R_\phi^p(N)]^{-1} \frac{1}{N} \sum_{k=1}^N \phi(k) v_0(k) \\ &= [R_\phi^p(N)]^{-1} R_\phi^p(N) \theta_0 + [R_\phi^p(N)]^{-1} \frac{1}{N} \sum_{k=1}^N \phi(k) v_0(k) \\ &= \theta_0 + [R_\phi^p(N)]^{-1} \frac{1}{N} \sum_{k=1}^N \phi(k) v_0(k) \end{aligned} \quad (5.6)$$

Željene osobine procene $\hat{\theta}_N$ su sledeće:

- 1) Da je $\hat{\theta}_N$ blisko tačnom θ_0 , odnosno da $\hat{\theta}_N$ konvergira nekom fiksnom vektoru.
- 2) Da $\hat{\theta}_N$ konvergira ka tačnom vektoru parametara θ_0 kada broj podataka N neograničeno raste ($N \rightarrow \infty$), odnosno da je procena $\hat{\theta}_N$ konzistentna.

Primetimo prvo da će član

$$[R_\varphi^p(N)]^{-1} \frac{1}{N} \sum_{k=1}^N \varphi(k) v_0(k) \quad (5.7)$$

u (5.6), koji definiše grešku ili odstupanje procene $\hat{\theta}_N$ od tačne vrednosti θ_0 i naziva se pomeraj procene, biti mali ukoliko je $v_0(k)$ malo u poređenju sa $\varphi(k)$, tako da će pod navedenim uslovima $\hat{\theta}_N$ biti blisko θ_0 .

Da bismo ispitali šta se dešava sa procenom $\hat{\theta}_N$ kada N neograničeno raste, tj. asimptotske osobine procene, uobičajeno je da se $v_0(k)$ posmatra kao realizacija široko (slabo) stacionarnog slučajnog procesa, a da se regresioni vektor $\varphi(k)$ pripiše ARX ili FIR strukturi modela. Ako se dalje pretpostavi da je pobudni signal $u(k)$ kvazi-stacionaran, tada će s obzirom da je sistem vremenski-invarijantan i stabilan i izlazni signal $y(k)$ biti kvazi-stacionaran, tako da će procene auto- i kros-kovarijacionih funkcija, koje predstavljaju elemente kovarijacione matrice $R_\varphi^p(N)$, konvergirati u određenom smislu (sa verovatnoćom 1), na osnovu zakona velikih brojeva, sa porastom broja podataka N ka njihovim tačnim vrednostima, odnosno

$$\hat{R}_u^N(\tau) = \frac{1}{N} \sum_{k=1}^N u(k)u(k-\tau) \xrightarrow[N \rightarrow \infty]{} R_u(\tau) = \bar{E}\{u(k)u(k-\tau)\} \quad (5.8)$$

$$\hat{R}_y(\tau) = \frac{1}{N} \sum_{k=1}^N y(k)y(k-\tau) \xrightarrow[N \rightarrow \infty]{} R_y(\tau) = \bar{E}\{y(k)y(k-\tau)\} \quad (5.9)$$

$$\hat{R}_{yu}(\tau) = \frac{1}{N} \sum_{k=1}^N y(k)u(k-\tau) \xrightarrow[N \rightarrow \infty]{} R_{yu}(\tau) = \bar{E}\{y(k)u(k-\tau)\} \quad (5.10)$$

gde je $\bar{E}\{\cdot\}$ operator generalisanog matematičkog očekivanja. Na taj način, i sama procena kovarijacione matrice regresionog vektora na bazi N izlazno-ulaznih podataka $R_\varphi^p(N)$ konvergiraće (sa verovatnoćom 1) ka konstantnoj matrici R^* , tj.

$$R_\varphi^p(N) = \frac{1}{N} \sum_{k=1}^N \varphi(k)\varphi^T(k) \xrightarrow[N \rightarrow \infty]{} R^* = \bar{E}\{\varphi(k)\varphi^T(k)\} \quad (5.11)$$

Analogno se može zaključiti da će i član koji definiše pomeraj LLS procena u (5.6) konvergirati ka fiksnom vektoru, tj.

$$\frac{1}{N} \sum_{k=1}^N \varphi(k)v_0(k) \xrightarrow[N \rightarrow \infty]{} h^* = \bar{E}\{\varphi(k)v_0(k)\} \quad (5.12)$$

Uzimajući u obzir poslednje dve relacije, dalje se zaključuje da LLS procena $\hat{\theta}_N$ konvergira kada $N \rightarrow \infty$ (asimptotski) ka fiksnoj vrednosti, odnosno

$$\hat{\theta}_N \xrightarrow[N \rightarrow \infty]{} \theta_0 + (R^*)^{-1} h^* \quad (5.13)$$

pretpostavljajući da je R^* regularna matrica. Poslednji uslov naziva se i uslov perzistentne eksitacije, koji u suštini označava da je pobudni signal tako izabran da je dovoljno sadržajan ili bogat, u smislu da će pobuditi sve relevantne modove sistema, kako bi se identifikovali svi nepoznati parametri u modelu sistema [23,30,63,67]. Drugim rečima, takav signal mora da generiše potpunu informaciju o ulazno-izlaznim osobinama posmatranog sistema. Bogastvo pobude ogleda se u njenom frekvencijskom liku, koji je opisan odgovarajućim spektrom snage signala [35,41,42,66]. Pobudni signal poseduje osobinu perzistentne eksitacije reda n ukoliko je njegov periodični spektar snage različit od nule za najmanje n kružnih učestanosti u intervalu osnovne periode. Kovarijaciona matrica ovakvog signala R_u^n , koja je definisana u drugom poglavlju, biće regularna, a ovakav signal se ne može kompletno potisnuti nekim MA filtrom reda $n-1$. Dobijeni rezultat (5.13) ukazuje da će procena $\hat{\theta}_N$ konvergirati ka θ_0 , tj. da će biti konzistentna, pod sledećim uslovima:

1) $\hat{R}_\varphi^p(N)$, odnosno R^* , je regularna matrica, što će biti ispunjeno ukoliko je pobudni signal $u(k)$ perzistentne eksitacije reda p (p je dimenzija vektora parametara koji se procenjuje).

2) $h^* = \bar{E}\{\varphi(k)v_0(k)\} = 0$, što će biti ispunjeno u bilo kome od sledeća dva slučaja:

2a) $v_0(k)$ je beli šum nulte srednje vrednosti $\bar{E}\{v_0(k)\} = 0$; u navedenom slučaju realizacija u k -tom diskretnom trenutku ne zavisi od toga šta se dogodilo do trenutka $(k-1)$, pri čemu regresioni vektor $\varphi^T(k) = \{-y(k-1) \dots -y(k-n_a) u(k-1) \dots u(k-n_b)\}$ sadrži samo ulazno-izlazne signale do trenutka $(k-1)$, tako da se može pisati $\bar{E}\{\varphi(k)v_0(k)\} = \bar{E}\{\varphi(k)\}\bar{E}\{v_0(k)\} = 0$.

2b) Istovremeno su ispunjene sledeće tri pretpostavke:

2b.1) pobudni signal $u(k)$ je nezavisan od šuma $v_0(k)$;

2b.2) bar jedan od signala $u(k)$ i $v_0(k)$ ima nultu srednju vrednost;

2b.3) $n_a = 0$, odnosno $\varphi(k)$ sadrži samo odbirke pobude $u(i)$, $k-1 \leq i \leq k-n_b$, $p = n_b$, što odgovara FIR strukturi modela.

Ukoliko je $n_a > 0$ a $v_0(k)$ nije beli šum, tada regresioni vektor $\varphi(k)$ sadrži i zakašnjene odbirke izlaza $y(i)$, $k-1 \leq i \leq k-n_a$, pa pošto $y(i)$ sadrži $v_0(i)$, koje je korelisano sa $v_0(k)$, to će u opštem slučaju važiti $\bar{E}\{\varphi(k)v_0(k)\} \neq 0$, odnosno $\hat{\theta}_N$ neće asimptotski konvergirati ($N \rightarrow \infty$) ka θ_0 , te procena LLS neće biti asimptotski

Analiza rekurzivnih stohastičkih algoritma parametarske identifikacije: konvergencija i konzistencija procena parametara i simulacija

konzistentna. Dakle, jedino u slučajevima (1), (2a) i (1), (2b) može se očekivati asimptotski konzistentna LLS procena vektora parametara.

Uz navedeno, u slučajevima (1) i (2a) može se takođe dokazati da greška procene konvergira po raspodeli ka normalnoj (gausovskoj) raspodeli nulte srednje vrednosti i matrice kovarijanse $\sigma_0^2(R^*)^{-1}$, gde je σ_0^2 varijansa belog šuma $v_0(k)$, odnosno [63,64]

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow{N \rightarrow \infty} N(0, \sigma_0^2(R^*)^{-1}) \quad (5.14)$$

Na taj način, zadatak planiranja eksperimenta, odnosno selekcije pobudnog signala, svodi se na zadatak da se R^* učini što je moguće većim pod zadatim realnim ograničenjima koja se postavljaju na sistem [29,30,43,67].

Teorija statističkog odlučivanja, kao i identifikacije sistema i estimacije parametara, bavi se problemom izdvajanja informacije iz merenja (observacija), koja sama po sebi mogu biti nepouzdana usled prisustva mernog šuma [25,35,60,73]. Stoga se observacije posmatraju kao slučajne varijable. U navedenom kontekstu pretpostavimo da su ulazno-izlazne observacije $\{u(i), y(i), i=1, \dots, N\} = \{u(i), y(i)\}_N$ predstavljene multidimenzionom slučajnom varijablom, odnosno slučajnim vektorom

$$z^N = \{z(1), \dots, z(N)\}; \quad z(i) = \{u(i), y(i)\}, \quad i=1, \dots, N \quad (5.15)$$

čije komponente $z(i) = \{u(i), y(i)\}$, $i=1, \dots, N$ uzimaju vrednosti iz realnog skupa brojeva R^{2N} . Pretpostavimo, takođe, da postoji združena (zajednička) funkcija gustine verovatnoće slučajnih varijabli $z(i)$, $i=1, \dots, N$

$$f(z_1, z_2, \dots, z_N; \theta) = f_z(z^N; \theta) \quad (5.16)$$

tako da je verovatnoća događanja $\{z^N \in A\}$ da realizacija slučajnog vektora z^N pripada nekom skupu (podprostoru) A skupa (N -dimenzionog Euklidovog prostora) R^{2N} data sa

$$P(z^N \in A) = \int_{z^N \in A} f_z(z^N; \theta) dz^N \quad (5.17)$$

U navedenim relacijama θ predstavlja p -dimenzioni vektor koji opisuje osobine razmatranih slučajnih varijabli. Pretpostavljajući da su izabrane slučajne varijable kao komponente slučajnog vektora merljive, cilj merenja ovih varijabli je da se estimira (proceni) vektor θ na bazi raspoložive informacije koja je dostupna kroz merenja. Ovaj zadatak se ostvaruje kroz estimator [63,64,88]

$$\hat{\theta}_N = \hat{\theta}(z^N) \quad (5.18)$$

koji na bazi izmerenih komponenti $N \times 1$ vektora z^N proračunava (procenjuje) vrednost $\hat{\theta}_N$ nepoznatog $p \times 1$ vektora parametara θ , odnosno predstavlja funkciju ili

preslikavanje iz skupa realnih brojeva R^{2N} na skup realnih brojeva R^p [61,79]. Na taj način, ako je izmereni vektor z^N označen sa z_N^* , tada je rezultujuća procena $\hat{\theta}_N$ data sa $\hat{\theta}_N = \hat{\theta}_N(z_N^*)$. Očigledno je da postoji vrlo veliki broj funkcija ili preslikavanja navedenog tipa koje se mogu uzeti za estimacioni algoritam, pri čemu se kao prirodni kriterijum za ocenu kvaliteta estimacije uzima matrica kovarijanse greške estimacije [63,64,88]

$$P = E \left\{ \left[\hat{\theta}(z^N) - \theta_0 \right] \left[\hat{\theta}(z^N) - \theta_0 \right]^T \right\} = E \left\{ \tilde{\theta}_N \tilde{\theta}_N^T \right\} \quad (5.19)$$

gde $E\{\cdot\}$ označava matematičko očekivanje, a

$$\tilde{\theta}_N = \hat{\theta}(z^N) - \theta_0 = \hat{\theta}_N - \theta_0 \quad (5.20)$$

je vektor grešaka estimacija, čije komponente predstavljaju greške estimacije pojedinih komponenti vektora parametara θ , čija je tačna vrednost označena sa θ_0 . Pošto dijagonalni elementi matrice P predstavljaju varijanse grešaka estimacija pojedinih komponenti vektora parametara, to je ukupna (totalna) greška estimacije definisana sa tragom matrice P (predstavlja sumu dijagonalnih elemenata ove matrice)

$$\text{Trag}P = \text{Trag}E\left\{ \tilde{\theta}_N \tilde{\theta}_N^T \right\} = \tilde{\theta}_N^T \tilde{\theta}_N; \quad \tilde{\theta}_N = \hat{\theta}_N - \theta_0 \quad (5.21)$$

Cramer-Raova teorema (nejednakost) određuje donju granicu matrice P za sve estimatore koji zadovoljavaju uslov nepomerivosti procene (eng. unbiased estimates)

$$E\left\{ \hat{\theta}_N \right\} = \theta_0; \quad E\left\{ \hat{\theta}_N \right\} = 0 \quad (5.22)$$

i opisana je matematički sa relacijom [60,63,64,88]

$$P = E \left\{ \left[\hat{\theta}_N - \theta_0 \right] \left[\hat{\theta}_N - \theta_0 \right]^T \right\} = E \left\{ \tilde{\theta}_N \tilde{\theta}_N^T \right\} \geq L^{-1} \quad (5.23)$$

gde je Cramer-Raova donja međa

$$L = E \left\{ \left[\frac{d}{d\theta} \log f_z(z^N; \theta) \right] \left[\frac{d}{d\theta} \log f_z(z^N; \theta) \right]^T \right\} \Big|_{\theta=\theta_0} = -E \left\{ \frac{d^2}{d\theta^2} \log f_z(z^N; \theta) \right\} \Big|_{\theta=\theta_0} \quad (5.24)$$

pretpostavljajući da funkcija gustine verovatnoće $f_z(z^N; \theta_0)$ postoji za bilo koje θ_0 , kao i da z^N uzima vrednosti iz nekog podskupa skupa realnih brojeva R^{2N} čije granice ne zavise od θ . Pošto je θ vektor kolona dimenzije $p \times 1$, to je gradijent $\left(\frac{d}{d\theta} \right) \log f_z(z^N; \theta)$ skalarne funkcije $\log f_z(z^N; \theta)$ $p \times 1$ dimenzioni vektor kolona, dok je Hessian $\left(\frac{d^2}{d\theta^2} \right) \log f_z(z^N; \theta)$ matrica dimenzije $p \times p$. Matrica L naziva se Fisherova

informaciona matrica. Pošto određivanje ove matrice zahteva tačnu vrednost θ_0 nepoznatog vektora parametara θ , to u praksi ova matrica obično nije poznata. Estimator koji dostiže Cramer-Raovu donju među L^{-1} je optimalan estimator, u smislu minimalno moguće ukupne (totalne) varijanse greške estimacije. U petom poglavlju biće razmatran problem optimalnog izbora indeksa performanse u zadacima parametarske identifikacije, koji se zasniva na primeni relacije (5.23)

Izvedimo sada izraz za matricu P za razmatrani slučaj modela linearne regresije (5.1), koji se može napisati u vektorsko-matričnom obliku

$$Y = \phi\theta_0 + V_0 \quad (5.25)$$

gde je

$$Y = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}; \quad \phi = \begin{bmatrix} \phi^T(1) \\ \vdots \\ \phi^T(N) \end{bmatrix}; \quad V_0 = \begin{bmatrix} v_0(1) \\ \vdots \\ v_0(N) \end{bmatrix} \quad (5.26)$$

pri čemu je, kao što je ranije pokazano u (5.6), linearni estimator najmanjih kvadrata LLS definisan sa

$$\hat{\theta}_N = (\phi^T \phi)^{-1} \phi^T Y = (\phi^T \phi)^{-1} \phi^T (\phi\theta_0 + V_0) = (\phi^T \phi)^{-1} \phi^T \phi \theta_0 + (\phi^T \phi)^{-1} \phi^T V_0$$

odnosno

$$\hat{\theta}_N = \theta_0 + (\phi^T \phi)^{-1} \phi^T V_0 \quad (5.27)$$

gde je kovarijaciona matrica regresionog vektora ϕ

$$R_\phi^p(N) = \phi^T \phi \quad (5.28)$$

Pretpostavljajući dalje da je ispunjen uslov perzistentne eksitacije, odnosno da je R_ϕ^p regularna matrica dok je V_0 slučajni vektor čije su komponente $v_0(k)$, $k=1, \dots, N$ realizacije belog šuma nulte srednje vrednosti koje su nekorelisane sa komponentama regresionog vektora $\phi(k)$, $k=1, \dots, N$, koji sadrži zakašnjene ulazno-izlazne signale $u(i)$, $k-1 \leq i \leq k-n_b$ i $y(j)$, $k-1 \leq j \leq k-n_a$. Tada se može pisati

$$\bar{E}\{\hat{\theta}_N\} = \theta_0 + \bar{E}\{(\phi^T \phi)^{-1} \phi^T V_0\} = \theta_0 + \bar{E}\{(\phi^T \phi)^{-1} \phi^T\} \bar{E}\{V_0\} \quad (5.29)$$

ili pošto je $\bar{E}\{V_0\} = 0$

$$\bar{E}\{\hat{\theta}_N\} = \theta_0 \quad (5.30)$$

odakle se zaključuje da je LLS procene nepomerena (srednja vrednost procene jednaka je tačnoj vrednosti vektora parametara). Dalje je greška estimacije

$$\tilde{\theta}_N = \hat{\theta}_N - \theta_0 = (\phi^T \phi)^{-1} \phi^T V_0 = (R_\phi^p)^{-1} \phi^T V_0 \quad (5.31)$$

pa je matrica kovarijanse greške estimacije

$$P_N = \bar{E} \{ \tilde{\theta}_N \tilde{\theta}_N^T \} = \bar{E} \{ (\phi^T \phi)^{-1} \phi^T V_0 V_0^T \phi (\phi^T \phi)^{-1} \} \quad (5.32)$$

Ukoliko su ispunjene pretpostavke (1) i (2) dalje se može pisati, pošto je regresioni vektor nezavisan od šuma

$$P_{N/N-1} = \bar{E} \{ \hat{\theta}_N \tilde{\theta}_N^T \big|_{z_{N-1}} \} = (\phi^T \phi)^{-1} \phi^T \bar{E} \{ V_0 V_0^T \big|_{z_{N-1}} \} \phi (\phi^T \phi)^{-1} \quad (5.33)$$

gde je kovarijaciona matrica šuma

$$R_0 = \bar{E} \{ v_0 v_0^T \} = \bar{E} \left\{ \begin{bmatrix} v_0(1) \\ \vdots \\ v_0(N) \end{bmatrix} \begin{bmatrix} v_0(1) \dots v_0(N) \end{bmatrix} \right\} \\ = \begin{bmatrix} \bar{E} \{ v_0^2(1) \} & \bar{E} \{ v_0(1) v_0(2) \} & \dots & \bar{E} \{ v_0(1) v_0(N) \} \\ \bar{E} \{ v_0(2) v_0(1) \} & \bar{E} \{ v_0^2(2) \} & \dots & \bar{E} \{ v_0(2) v_0(N) \} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{E} \{ v_0(N) v_0(1) \} & \bar{E} \{ v_0(N) v_0(2) \} & \dots & \bar{E} \{ v_0^2(N) \} \end{bmatrix} \quad (5.34)$$

Pošto je $v_0(k)$ realizacija belog šuma nulte srednje vrednosti, to je

$$\bar{E} \{ v_0(i) v_0(j) \} = 0, \quad i \neq j; \quad \bar{E} \{ v_0^2(i) \} = \sigma_0^2, \quad \forall i = 1, 2, \dots, N \quad (5.35)$$

gde je pretpostavljeno da sve realizacije belog šuma $v_0(i)$ imaju identičnu funkciju raspodele sa istom varijansom, tako da se može pisati

$$R_0 = \sigma_0^2 I_N \quad (5.36)$$

pri čemu je I_N jedinična matrica dimenzija $N \times N$. Na taj način se dobija

$$P_N = \sigma_0^2 \bar{E} \{ (\phi^T \phi)^{-1} \} = \sigma_0^2 [R_\phi^p(N)]^{-1} \quad (5.37)$$

Prethodno je takođe napomenuto da asimptotski, kada $N \rightarrow \infty$,

$$R_\phi^p(N) = \bar{E} \left\{ \frac{1}{N} \sum_{k=1}^N \phi(k) \phi^T(k) \right\} \xrightarrow{N \rightarrow \infty} R^* \quad (5.38)$$

gde je R^* konstantna matrica, tako da važi

$$P_N \xrightarrow[N \rightarrow \infty]{} P = \sigma_0^2 (R^*)^{-1} \quad (5.39)$$

Rekurzivni linearni algoritam najmanjih kvadrata izveden je u trećem poglavlju direktno iz nerekurzivnog linearnog algoritma najmanjih kvadrata, te poseduje ista asimptotska svojstva kao poslednji algoritam, a izvršena analiza odnosi se na oba pomenuta algoritma.

5.2 ANALIZA KONVERGENCIJE PSEUDO-LINEARNOG METODA NAJMANJIH KVADRATA (PLLS):

Kao što je istaknuto u trećem poglavlju, u opštem slučaju modela pseudo-linearne regresije, kod koga i regresioni vektor $\varphi(k)$ zavisi od nepoznatog vektora parametara θ , tj. $\varphi(k) = \varphi(k, \theta)$ kriterijum koji se minimizira

$$J_N(\theta) = \frac{1}{N} \sum_{k=1}^N \varepsilon^2(k, \theta) = \frac{1}{N} \sum_{k=1}^N [y(k) - \varphi^T(k, \theta)\theta]^2 \quad (5.40)$$

ne predstavlja konveksnu funkciju argumenta θ , te ne predstavlja ni unimodalnu funkciju koja ima jedinstven minimum, tako da će procena pseudo-linearne regresije koja minimizira usvojeni kriterijum

$$\hat{\theta}_N = \arg \min_{\theta} J_N(\theta) \quad (5.41)$$

u najboljem slučaju konvergirati ka jednom od postojećih lokalnih minimuma kriterijuma. Može se pokazati, na osnovu zakona velikih brojeva, da pod određenim i ne tako strogim uslovima kriterijum $J_N(\theta)$ konvergira u određenom smislu (sa verovatnoćom 1) ka graničnoj vrednosti $J_{\infty}(\theta)$ kada broj podataka $N \rightarrow \infty$ [87,88]

$$J_N(\theta) \xrightarrow[N \rightarrow \infty]{} J_{\infty}(\theta) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \varepsilon^2(k, \theta) = \overline{E} \{ \varepsilon^2(k, \theta) \} \quad (5.42)$$

Ako je D_M skup svih mogućih vrednosti vektora parametara θ , poslednja relacija označava da supremum apsolutne vrednosti odstupanja $J_N(\theta)$ od $J_{\infty}(\theta)$ (maksimum od svih odstupanja za $\theta \in D_M$) konvergira (sa verovatnoćom 1) ka nuli, odnosno

$$\sup_{\theta \in D_M} |J_N(\theta) - J_{\infty}(\theta)| \xrightarrow[N \rightarrow \infty]{} 0 \quad (5.43)$$

Takođe, ukoliko kriterijum $J_N(\theta)$ asimptotski konvergira ($N \rightarrow \infty$) ka $J_\infty(\theta)$ tada će i argument $\hat{\theta}_N$ koji minimizira $J_N(\theta)$ konvergirati (sa verovatnoćom 1) ka vrednosti θ_∞ koja minimizira graničnu vrednost kriterijuma $J_\infty(\theta)$. Pri tome, vrednost θ_∞ u opštem slučaju nije jedinstvena, već pripada nekom skupu D_∞ definisanom sa

$$D_\infty = \arg \min_{\theta \in D_M} J_\infty(\theta) = \left\{ \theta \mid \theta \in D_M, J_\infty(\theta) = \min_{\eta \in D_M} J_\infty(\eta) \right\} \quad (5.44)$$

odnosno

$$\hat{\theta}_N \xrightarrow[N \rightarrow \infty]{} D_\infty \quad (5.45)$$

Konvergencija u okviru skupa označava da infimum apsolutnog odstupanja $\hat{\theta}_N$ od θ_∞ (minimum od svih mogućih odstupanja) konvergira asimptotski (kada $N \rightarrow \infty$) ka nultoj vrednosti [1,79], tj.

$$\inf_{\theta_\infty \in D_\infty} \left| \hat{\theta}_N - \theta_\infty \right| \xrightarrow[N \rightarrow \infty]{} 0 \quad (5.46)$$

Imajući u vidu sve rečeno, za neki estimator $\hat{\theta}_N$ se kaže da je konvergentan ukoliko procena $\hat{\theta}_N$ konvergira (sa verovatnoćom 1) u okviru skupa D_∞ , odnosno da važi relacija (5.45), gde je D_∞ skup koji kao elemente sadrži sve minimume graničnog kriterijuma $J_\infty(\theta)$, kao što je definisano u (5.42) i (5.44).

Ukoliko skup D_∞ sadrži samo jednu vrednost θ_0

$$D_\infty = \{\theta_0\} \quad (5.47)$$

tada je estimator $\hat{\theta}_N$ i konzistentan.

Očigledno je da je konzistentnu procenu parametara sistema moguće dobiti jedino ukoliko usvojena parametrizacija, odnosno izabrani vektor parametara θ egzaktno opisuje realan sistem, kada će tačna vrednost vektora parametara θ_0 pripadati skupu mogućih vrednosti vektora parametara D_M . Ukoliko stvarni vektor parametara θ_0 pripada skupu D_M , tada će funkcije prenosa $G(q, \hat{\theta}_N)$ i $H(q, \hat{\theta}_N)$ u standardnoj reprezentaciji modela sistema na slici (2.6) konvergirati kada $N \rightarrow \infty$ ka njihovim stvarnim vrednostima $G_0(q)$ i $H_0(q)$, respektivno. Pitanje da li se stvarna funkcija prenosa $G_0(q)$ i $H_0(q)$, kojima je opisan realan sistem, mogu odrediti na osnovu planiranog eksperimenta naziva se problem identifikabilnosti. Ako realan sistem ne pripada usvojenom skupu modela. D_M tada se može govoriti o aproksimativnoj (približnoj) identifikaciji, o čemu će biti više reči u nastavku.

5.2.1 Aproksimativna identifikacija

Kada se koristi metod greške predikcije (PEM) u (5.40), (5.41) za parametarsku identifikaciju sistema, tada se vrednost θ_∞ u (5.46) kojoj asimptotski konvergira PEM procena $\hat{\theta}_N$ kada $N \rightarrow \infty$ naziva granična vrednost i njoj odgovara granični model. Analizirajmo u nastavku dobijeni granični model.

Granična vrednost PEM kriterijuma $J_N(\theta)$ je data sa (5.42). Sekvenci (vremenskoj seriji) greške predikcije $\varepsilon(k, \theta)$, $k=1, \dots, N$ odgovara u frekvencijskom domenu spektar snage $S_\varepsilon(\omega, \theta)$ koji je definisan kao diskretna Fourierova transformacija auto-kovarijacione funkcije (ovaj rezultat je poznat kao Viner-Hinčinova teorema) [35,48,71]

$$R_\varepsilon(\tau, \theta) = \bar{E}\{\varepsilon(k, \theta)\varepsilon(k - \tau, \theta)\} \quad (5.48)$$

pri čemu je srednja vrednost $\bar{E}\{\varepsilon(k, \theta)\} = 0$, odnosno

$$S_\varepsilon(\omega, \theta) = \sum_{\tau=-\infty}^{\infty} R_\varepsilon(\tau, \theta) e^{-i\omega\tau} \quad (5.49)$$

Auto-kovarijaciona funkcija se može jedinstveno odrediti na osnovu spektra snage primenom inverzne vremenski diskretne Fourierove transformacije

$$R_\varepsilon(\tau, \theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_\varepsilon(\omega, \theta) e^{i\omega\tau} d\omega \quad (5.50)$$

Posebno, za $\tau = 0$ sledi

$$R_\varepsilon(0, \theta) = \bar{E}\{\varepsilon^2(k, \theta)\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_\varepsilon(\omega, \theta) d\omega \quad (5.51)$$

Imajući u vidu da je na osnovu (5.42)

$$J_\infty(\theta) = \bar{E}\{\varepsilon^2(k, \theta)\} \quad (5.52a)$$

zaključuje se da se granična vrednost PEM kriterijuma u (5.42) može izraziti u frekvencijskom domenu kao

$$J_\infty(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_\varepsilon(\omega, \theta) d\omega \quad (5.52b)$$

Pretpostavimo dalje da je realni sistem opisan egzaktno (tačno) sa funkcijom prenosa $G_0(q)$, tako da je izlaz sistema $y(k)$ dat sa

$$y(k) = G_0(q)u(k) + v_0(k) \quad (5.53)$$

gde je $u(k)$ pobudni signal a $v_0(k)$ poremećaj (šum) nulte srednje vrednosti na izlazu, koji se egzaktno može modelovati propuštanjem belog šuma nulte srednje vrednosti $e(k)$ kroz funkciju prenosa $H_0(q)$, tj.

$$v_0(k) = H_0(q)e(k) \quad (5.54)$$

Sa druge strane, estimirane (procenjene) funkcije prenosa realnog sistema i aditivnog šuma su $G(q, \theta)$ i $H(q, \theta)$, respektivno, gde je θ procenjena vrednost vektora parametara. Greške predikcije, kao procena belog šuma $e(k)$, izvedena je u trećem poglavlju i zadovoljava relaciju

$$\varepsilon(k, \theta) = H^{-1}(q, \theta)[y(k) - G(q, \theta)u(k)] \quad (5.55)$$

ili kada se zameni izraz za $y(k)$ iz (5.53) u (5.55)

$$\varepsilon(k, \theta) = H^{-1}(q, \theta)[G_0(q) - G(q, \theta)]u(k) + v_0(k) = G_e(q, \theta)u(k) + H_e(q, \theta)v_0(k) \quad (5.56)$$

gde je

$$G_e(q, \theta) = \frac{G_0(q) - G(q, \theta)}{H(q, \theta)}; \quad H_e(q, \theta) = \frac{1}{H(q, \theta)} \quad (5.57)$$

pri čemu je $v_0(k)$ dato sa (5.54).

Pretpostavljajući da su pobuda $u(k)$ i šum $v_0(k)$ nekorelisani kvazi-stacionarni slučajni procesi i koristeći ranije izvedene izraze za transformaciju spektra u trećem poglavlju, zaključuje se da je spektar snage greške predikcije

$$S_\varepsilon(\omega, \theta) = |G_e(e^{i\omega}, \theta)|^2 S_u(\omega) + |H_e(e^{i\omega}, \theta)|^2 S_{v_0}(\omega) \quad (5.58)$$

gde je

$$|G_e(e^{i\omega}, \theta)|^2 = \frac{|G_0(e^{i\omega}) - G(e^{i\omega}, \theta)|^2}{|H(e^{i\omega}, \theta)|^2} \quad (5.59)$$

$$|H_e(e^{i\omega}, \theta)|^2 = \frac{1}{|H(e^{i\omega}, \theta)|^2} \quad (5.60)$$

a vrednost kriterijuma (5.52) u frekvencijskom području je specificirana sa

$$J_{\infty}(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G_0(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 S_u(\omega) + S_{v_0}(\omega)}{|H(e^{i\omega}, \theta)|^2} d\omega \quad (5.61)$$

gde je spektar snage aditivnog šuma

$$S_{v_0}(\omega) = |H_0(e^{i\omega})|^2 S_e(\omega) = |H_0(e^{i\omega})|^2 \sigma_e^2 \quad (5.62)$$

U navedenim izrazima σ_e^2 označava varijansu belog šuma $e(k)$ čijim se filtriranjem u (5.54) dobija šum $v_0(k)$, a $S_u(\omega)$ je spektar snage pobudnog signala. Tada je skup D_{∞} kome pripada granična vrednost PEM procene parametara definisan sa (5.44). Dobijeni rezultat ukazuje na sledeće zaključke. Ukoliko je stvarni sistem $G_0(q) = G(q, \theta_0)$ u usvojenom skupu modela D_M tada se može postići da model konvergira ka stvarnom sistemu, odnosno da

$$|G_0(q) - G(q, \theta)| \rightarrow 0 \quad \text{za} \quad \theta \rightarrow \theta_0 \quad (5.63)$$

gde je θ_0 pravi vektor parametara, tako da će se minimum kriterijumske funkcije dobiti ako model šuma konvergira ka stvarnom procesu kojim je generisan aditivni šum, tj. ako

$$H(q, \theta) \rightarrow H_0(q) = H(q, \theta_0) \quad \text{za} \quad \theta \rightarrow \theta_0 \quad (5.64)$$

kada se za minimum kriterijuma dobija

$$\min_{\theta \in D_M} J_{\infty}(\theta) = \sigma_e^2 \quad (5.65)$$

Tada skup D_{∞} sadrži samo jednu graničnu vrednost θ_0 i PEM procena parametara θ je konzistentna.

Međutim, ako sistem ne pripada usvojenom skupu modela D_M u (5.44) ne može se izvesti neki generalni teorijski rezultat, ali se intuitivno mogu definisati sledeća dva slučaja:

- 1) Fiksni model šuma;
- 2) Generalni model šuma.

U fiksnom modelu šuma se usvaja

$$H(q, \theta) = H^*(q) \quad \text{za} \quad \forall \theta \in D_M \quad (5.66)$$

tako da je granična vrednost PEM kriterijuma

$$J_{\infty}(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |G_0(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 Q^*(\omega) d\omega + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{S_{v_0}(\omega)}{|H^*(e^{i\omega})|^2} d\omega \quad (5.67)$$

gde je

$$Q^*(\omega) = \frac{S_u(\omega)}{|H^*(e^{i\omega})|^2} \quad (5.68)$$

tako da je skup D_{∞} kome pripada granična vrednost PEM procene θ_{∞} određen samo minimumom prvog člana u $J_{\infty}(\theta)$ (drugi član ne zavisi od θ), odnosno

$$D_{\infty} = \arg \min_{\theta \in D_M} \int_{-\pi}^{\pi} |G_0(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 Q^*(\omega) d\omega \quad (5.69)$$

Na taj način, za bilo koje $\theta^* \in D_{\infty}$ procena funkcije prenosa sistema $G(e^{i\omega}, \theta^*)$ je optimalna u smislu minimuma srednje-kvadratnog kriterijuma greške procene $G_0(e^{i\omega}) - G(e^{i\omega}, \theta)$, pri čemu je kvadrat greške procene ponderisan (otežinjen) u frekvencijskom domenu sa faktorom $Q^*(\omega)$ koji zavisi od spektra snage pobude $S_u(\omega)$ i usvojenog fiksnog modela šuma $H^*(q)$. Posebno za OE reprezentaciju modela (D_M je skup OE modela) fiksni model šuma $H^*(q) = 1$, tako da je na osnovu (5.68)

$$Q^*(\omega) = S_u(\omega) \quad (5.70)$$

Na taj način, selekcijom pobudnog signala, odnosno izborom njegove spektralne karakteristike, može se uticati na rezultat estimacije i posebnu težinu prilikom minimizacije dati ovom frekvencijskom području u kome se nalazi koristan pobudni signal koji se prenosi i obrađuje posmatranim sistemom (u tom području će biti najbolje slaganje stvarne i estimirane funkcije prenosa).

U slučaju generalnog parametarskog modela sistema nije moguće egzaktno teorijski okarakterisati graničnu vrednost rezultujuće PEM procene. U literaturi postoji teorijski rezultat koji se odnosi na BJ strukturu modela i pokazuje da je kod ovakve strukture PEM procena θ tako određena da estimirani model spektra šuma $|H(e^{i\omega}, \theta)|^2$ što je moguće bolje aproksimira spektar $S_{\varepsilon}(\omega, \theta)$ greške predikcije $\varepsilon(k, \theta) = y(k) - G(q, \theta)u(k)$. Međutim, ovakav rezultat nije u opštem slučaju moguće izvesti za ostale parametarske reprezentacije, kao što su ARX i ARMAX. Kod ovakvih struktura granična vrednost θ^* PEM procene određena je u opštem slučaju kao kompromis između što bolje aproksimacije stvarnog sistema $G_0(e^{i\omega})$ procenjenim modelom $G(e^{i\omega}, \theta)$, u smislu minimalno mogućeg ponderisanog srednje-kvadratnog kriterijuma greške u frekvencijskom domenu, gde je težinski faktor u frekvencijskom domenu dat sa (5.68), odnosno

$$Q(\omega, \theta^*) = \frac{S_u(\omega)}{|H(e^{i\omega}, \theta^*)|^2} \quad (5.71)$$

i što je moguće većeg približavanja (fitovanja) modela spektra šuma $|H(e^{i\omega}, \theta^*)|^2$ spektru $S_\varepsilon(\omega, \theta^*)$ greške predikcije $\varepsilon(k, \theta^*) = y(k) - G(q, \theta^*)u(k)$.

5.2.2 Filtriranje grešaka predikcije

Prethodne analize su pokazale da će modelovanje ARX strukturom podataka koji nisu egzaktno generisani ARX procesom, već na primer ARMAX procesom, rezultovati u pomerene PEM procene parametara. Međutim, glavna prednost ARX strukture je da je PEM estimator linearan, tako da se procene parametara dobijaju u jednom koraku, rešavanjem linearnog sistema jednačine po nepoznatim elementima vektora parametara θ . Dakle, postavlja se pitanje kako poboljšati linearnu procenu najmanjih kvadrata LLS, da bi se koristila zbog svoje računske jednostavnosti i u slučajevima kada podaci nisu egzaktno generisani modelom linearne regresije. Na osnovu analize u prethodnom poglavlju zaključuje se da je asimptotska (granična) vrednost θ_∞ PEM procene u slučaju ARX strukture određene kao

$$\theta_\infty = \arg \min_{\theta} \int_{-\pi}^{\pi} \left\{ \left| G_0(e^{i\omega}) - \frac{B(e^{i\omega}, \theta)}{A(e^{i\omega}, \theta)} \right|^2 Q_{ARX}(\omega, \theta) + |A(e^{i\omega}, \theta)|^2 S_{v_0}(\omega) \right\} d\omega \quad (5.72)$$

gde je

$$Q_{ARX}(\omega, \theta) = |A(e^{i\omega}, \theta)|^2 S_u(\omega) \quad (5.73)$$

Primetimo da je za ARX strukturu $G = B/A$ i $H = 1/A$. U opštem slučaju $1/A(q)$ odgovara funkciji prenosa nisko-frekvencijskog (NF) filtra, tako da inverzna funkcija prenosa $A(q)$ odgovara funkciji prenosa visoko-frekvencijskog (VF) filtra. To znači da će se u otežinjenom srednje-kvadratnom kriterijumu greške procene u frekvencijskom domenu veće težine dodeliti visokim frekvencijama, što će imati za posledicu da će PEM procene $G(e^{i\omega}, \theta) = B(e^{i\omega}, \theta)/A(e^{i\omega}, \theta)$ bolje aproksimirati stvarnu funkciju prenosa $G_0(e^{i\omega})$ na višim nego na nižim frekvencijama. Ovo predstavlja neželjeni efekat, pošto je u opštem slučaju područje niskih frekvencija od većeg značaja u mnogim primenama. Da bi se poboljšala PEM procena na nižim frekvencijama, po cenu smanjenja tačnosti procene funkcije prenosa sistema na višim frekvencijama, koriste se filtri. Postupak se sastoji u tome da se umesto greške predikcije $\varepsilon(k, \theta)$ koristi filtrirana greškapredikcije

$$\varepsilon_F(k, \theta) = F(q)\varepsilon(k, \theta) \quad (5.74)$$

gde je $F(q)$ linearan, diskretan, vremenski-invarijantan inverzno-stabilan filter konačnog reda. Na osnovu rezultata o transformaciji spektra, spektar filtrirane greške predikcije dat je sa

$$S_{\varepsilon_F}(\omega) = |F(e^{i\omega})|^2 S_{\varepsilon}(\omega) \quad (5.75)$$

gde je ranije navedeni izraz za spektar snage greške predikcije dat sa (5.58)-(5.60), odnosno

$$S_{\varepsilon}(\omega) = \frac{|G_0(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 S_u(\omega) + S_{v_0}(\omega)}{|H(e^{i\omega}, \theta)|^2} \quad (5.76)$$

Tada je granična (asimptotska) vrednost PEM kriterijuma data sa (5.52), t.j.

$$J_{\infty}(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{\varepsilon_F}(\omega) d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{i\omega})|^2 S_{\varepsilon}(\omega) d\omega \quad (5.77)$$

Na taj način, za ARX model frekvencijska težinska funkcija data je sa

$$Q_{ARX}(\omega) = |F(e^{i\omega})|^2 |A(e^{i\omega}, \theta)|^2 S_u(\omega) \quad (5.78)$$

Postoji više načina da se izabere filter $F(q)$. Posebno, ako se izabere kao u četvrtom poglavlju

$$F(q) = \frac{1}{A(q)} \quad (5.79)$$

dostiće se ARX strukturom kvalitet procene parametara kao kod OE reprezentacije modela, t.j. fiksnog modela šuma u (5.70). Međutim, pošto polinom $A(q)$ nije poznat, mogući pristup sastoji se da se u prvom koraku proceni ovaj polinom, pa da se tako estimirani A -polinom koristi kao filter u drugom koraku. Ovakav postupak može se koristiti iterativno, u nekoliko koraka, kako bi se dobio zadovoljavajući kvalitet procena.

Međutim, ideja da se filtrira greška predikcije može se koristiti kod bilo koje parametarske reprezentacije modela. Izborom filtra se dodeljuje veća težina u ponderisanom kriterijumu srednje-kvadratne greške estimacije u frekvencijskom domenu onim frekvencijama koje su od interesa, čime se u ovom području postiže bolja aproksimacija stvarne funkcije prenosa procesa. Ovo je posebno važno ukoliko nije moguće da se eksperiment redizajnira. Primetimo, takođe, da je filtriranje greške predikcije sa filtrom $F(q)$ ekvivalentno kod sistema sa jednim ulazom i jednim izlazom filtriranju samih ulazno-izlaznih podataka, odnosno isti rezultat se dobija kada se umesto podataka $\{u(k), y(k)\}_N$ koriste filtrirani podaci $\{u_F(k), y_F(k)\}_N$, o čemu je već bilo reči u četvrtom poglavlju.

5.3 ANALIZA KONVERGENCIJE NELINEARNIH REKURZIVNIH STOHAISTIČKIH ALGORITAMA PARAMETARSKE IDENTIFIKACIJE

5.3.1 Analiza konvergencije nelinearnih rekurzivnih stohastičkih algoritama za estimaciju parametara u dinamičkim modelima sistema tipa linearne regresije (ARX i FIR modelima) primenom teorije martingala

U ovom poglavlju razmatrana je klasa modela sistema koji su opisani diferencnom jednačinom sa fiksnim koeficijentima

$$y(i) = -\sum_{k=1}^n a_k y(i-k) + \sum_{k=1}^m b_k u(i-k) + \xi(i) \quad (5.80)$$

gde su $y(i) \in R^1$, $u(i) \in R^1$, $\xi(i) \in R^1$ izlaz sistema, merljivi pobudni signal ili ulaz i nemerljivi stohastički ulaz ili poremećaj, respektivno, dok su $a_i, i=1, \dots, n$ i $b_j, j=1, \dots, m$ konstantni parametri u modelu. Dalje se pretpostavlja da je diskretan niz $\{\xi(i)\}$ stohastički proces, koji generiše rastući niz sub-sigma algebri $\{F_i\}$, kao i da su poznate konstante m i n ($0 < m \leq n$). Takodje, funkcija gustine verovatnoće (fgv) slučajne varijable $\xi(i)$ nije u opštem slučaju poznata, ali je uobičajeno da se na osnovu izvesnog apriornog znanja o statističkim karakteristikama poremećaja definiše određena klasa funkcija kojoj će i realna fgv poremećaja pripadati.

Ako se dalje uvede operator kašnjenja $q^{-k}y(i) = y(i-k)$, model (5.80) može se zapisati u polinomijalnom obliku

$$A(q^{-1})y(i) = B(q^{-1})u(i) + \xi(i), \quad (5.81)$$

gde se polinomi

$$A(q^{-1}) = 1 + \sum_{k=1}^n a_k q^{-k}, \quad B(q^{-1}) = \sum_{k=1}^m b_k q^{-k}, \quad (5.82)$$

nazivaju karakterističnim i kontrolnim, respektivno. Model (5.80) može se takodje zapisati i u formi linearnr regresije

$$y(i) = Z^T(i)\Theta + \xi(i) \quad (5.83)$$

gde je regresioni vektor

$$Z^T(i) = [-y(i-1), \dots, -y(i-n), u(i-1), \dots, u(i-m)] \quad (5.84)$$

dok je $\Theta^T = [a_1 \dots a_n b_1 \dots b_m]$ vektor konstantnih parametra čije vrednosti treba proceniti na osnovu raspoložive merne sekvence ulaznih i izlaznih signala. U prethodnom poglavlju je razmatrana klasa rekurzivnih stohastičkih algoritama za estimaciju nepoznatog vektora parametara

$$\hat{\Theta}(i) = \hat{\Theta}(i-1) + \Gamma(i) Z(i) \psi(v(i, \hat{\Theta}(i-1))) \quad (5.85)$$

gde je rezidual merenja ili is greška predikcije

$$v(i, \hat{\Theta}) = y(i) - \hat{\Theta}^T Z(i) \quad (5.86)$$

dok je težinska matrica algoritma određena rekurzivnom relacijom

$$\begin{aligned} \Gamma(i) &= \Gamma(i-1) - \Gamma(i-1) Z(i) \\ &\times [\alpha^{-1} + Z^T(i) \Gamma(i-1) Z(i)]^{-1} Z^T(i) \Gamma(i) \end{aligned} \quad (5.87)$$

pri čemu je njena inicijalna vrednost $\Gamma(0) = \rho I$, $\rho > 0$. Koeficijent α predstavlja slobodan parametar čijim se izborom može uticati na konvergencijska svojstva algoritma. Takodje, pogodnim izborom ovog parametra može se uspostaviti veza između nelinearne transformacije reziduala $\psi(\cdot)$ i matrice pojačanja algoritma $\Gamma(\cdot)$. Naime, ako se izabere da je α jednak matematičkom očekivanju, u odnosu na fgv poremećaja, prvog izvoda nelinearne transformacije reziduala, tj. $\alpha = E\{\psi'(v(i, \hat{\Theta}))\}$, tada se ovaj parametar može posmatrati kao koeficijent statističke linearizacije $\psi(v(i, \hat{\Theta})) \approx \alpha v(i, \hat{\Theta})$, a algoritam (5.85)-(5.87) se svodi na robusni Kalmanov filter, odnosno rekurzivni robusni algoritam najmanjih kvadrata, koji je izveden u literaturi u sasvim drugačijem kontekstu i na različit način od predloženog [45,84]. Sa druge strane, ako se matematičko očekivanje u statističkoj linearizaciji aproksimira samo sa jednom realizacijom, tj. $\alpha = \alpha(i) = \psi'(v(i, \hat{\Theta}))$, statistička linearizacija se svodi na klasičnu, dok se razmatrani algoritam parametarske estimacije svodi na varijantu rekurzivnog robusnog algoritma najmanjih kvadrata koji je izveden u literaturi polazeći od odgovarajućeg nerekurzivnog algoritma, uz uvođenje pogodnih aproksimacija koje su uobičajene kod ovakvih izvodjenja [72,89]. Konačno, ako se usvoji $\alpha = 1$ relacija (5.87) svodi se na diskretnu Rikatiјevu jednačinu, a algoritam identifikacije postaje jedna varijanta rekurzivnog robusnog algoritma najmanjih kvadrata, koji se od osnovnog algoritma razlikuje samo po nelinearnoj transformaciji reziduala.

Konvergencija razmatranog algoritma, pored faktora α u izrazu (5.87) za matricu pojačanja $\Gamma(\cdot)$ i nelinearnosti $\psi(\cdot)$ u izrazu (5.85) za estimaciju parametara modela, zavisi i od karakteristika determinističkog i slučajnog ulaza u reprezentaciji sistema (5.81), kao i od osobina usvojenog modela sistema (5.81). Pogodan alat za analizu konvergencije rekurzivnih stohastičkih algoritama razmatranog tipa predstavlja teorija martingala [37,62,70]. Martingal je specifičan stohastički proces koji ispunjava sledeće zahteve. Neka je sa (Ω, F, P) označen prostor verovatnoće, gde je Ω siguran događaj, F polje događaja a P verovatnoća, pri čemu je $F_1 \subset F_2 \subset \dots$ rastići niz sub- σ -algebri generisanih sa F , dok je $x(t)$ niz diskretnih slučajnih varijabli definisanih na polju događaja F . Tada je $\{x(t), F_t\}$ martingal ako su ispunjeni uslovi $E\{|x(t)|\} < \infty$ skoro sigurno (s.s), odnosno sa verovatnoćom 1 (s.v.1), i $E\{x(t)/F_{t-1}\} = x(t-1)$ s.v.1. Takodje, ukoliko je $E\{x(t)/F_{t-1}\} \leq x(t-1)$ s.v.1 tada je

$\{x(t), F_t\}$ supermartingal. Osnovni rezultat vezan za konvergenciju martingala formulisao je Neveu[78] i ovaj rezultat je prikazan u okviru sledeće leme 1.

Lema 5.1: Neka je $\{z_n\}$ niz nenegativnih diskretnih slučajnih varijabli i $\{F_n\}$ odgovarajući niz rastućih sigma algebra, takvih da je $z_n \in F_n$. Pretpostavimo da su ispunjeni zahtevi

$$E\{z_n / F_{n-1}\} \leq z_{n-1} + \alpha_n$$

i $\sum_{n=1}^{\infty} \alpha_n < \infty$ s.v.1. Tada diskretni niz $\{z_n\}$ konvergira s.v.1 ka konačnoj, nenegativnoj slučajnoj varijabli z^* kada $n \rightarrow \infty$, t.j. $\lim_{n \rightarrow \infty} z_n = z^*$ s.v.1

Ovaj rezultat preformulisan je u literature u nekoliko alternativnih formi, koje su pogodnije za primenu u određenim situacijama. Robbins i Siegmund [77] predložili su unificirani pristup za proučavanje većeg broja teorema o skoro sigurnoj konvergenciji slučajnih procesa koji poseduju skoro uvek osobine supermartingala. Ovaj rezultat prikazan je u okviru sledeće leme.

Lema 5.2: Neka za svako n postoje nenegativne i F_n -merljive diskretne slučajne varijable z_n, β_n, ξ_n i ζ_n takve da je

$$E\{z_{n+1} | F_n\} \leq z_n(1 + \beta_n) + \xi_n - \zeta_n$$

Tada postoji konačna granična vrednost $\lim_{n \rightarrow \infty} z_n$, t.j. $\lim_{n \rightarrow \infty} z_n = z^*$ s.v.1, pri čemu je ispunjeno

$$\sum_{n=1}^{\infty} \zeta_n < \infty \text{ s.v.1, na skupu } \left\{ \sum_{n=1}^{\infty} \beta_n < \infty, \sum_{n=1}^{\infty} \xi_n < \infty \right\}.$$

Prilikom dokaza teorema o konvergenciji martingala često se koriste i sledeće osobine realnih nizova, a ovaj rezultat je poznat u literature kao Kronekerova lema. Ova lema predstavlja specijalan slučaj Toplitzove leme i prikazana je u nastavku [44].

Lemma 5.3: Neka niz $\sum_{k=1}^n x_k$ konvergira, gde je $\{x_k\}$ niz realnih brojeva, i neka je $\{b_n\}$ neopadajući niz realnih brojeva takav da je $\lim_{n \rightarrow \infty} b_n = \infty$. Tada je u važnosti

$$\lim_{n \rightarrow \infty} \frac{1}{b_n} \sum_{k=1}^n b_k x_k = 0$$

Analiza konvergencije rekurzivne stohastičke procedure (5.85)-(5.87) može se zasnovati na osnovnim rezultatima teorije martingala pod uslovom da se pokaže da greška estimacije poseduje svojstvo martingala. Da bi se ovaj zahtev ispunio neophodno je da se nametnu određena ograničenja na nelinearnu transformaciju reziduala $\psi(\cdot)$, faktor konvergencije α , matricu pojačanja $\Gamma(\cdot)$, karakteristike determinističkog i stohastičkog pobudnog signala u (5.81), kao i osobine samog modela sistema (5.81). Jedan mogući skup ovakvih ograničenja koja rezultuju u željena svojstva greške estimacije kao slučajne veličine predložena su u nastavku.

A1: Sve nule karakterističnog polinoma $A(q^{-1})$ leže unutar jediničnog kruga, sa centrom u koordinatnom početku kompleksne ravni, a deterministički diskretni pobudni signal $\{u(i)\}$ je ograničen.

A2: Diskretni slučajni pobudni signal ili poremećaj $\{\xi(i)\}$ predstavlja niz nezavisnih i identički raspodeljenih slučajnih varijabli (odgovarajući engleski termin je: a sequence of independent and identically distributed random variables ili skraćeno i.i.d.) koje poseduju simetričnu funkciju raspodele verovatnoće $P(\cdot)$ i odgovarajuću funkciju gustine verovatnoće (fgv), pri čemu su uslovna srednja vrednost i uslovna varijansa date izrazima $E\{\xi(i)/F_{i-1}\}=0$, $E\{\xi^2(i)/F_{i-1}\}=\sigma^2<\infty$, dok slučajna veličina $v(i)-\xi(i)$ ispunjava uslov $v(i)-\xi(i)\in F_{i-1}$, gde je $v(i)$ rezidual merenja u (5.86) a F_{i-1} je σ -algebra generesina sa slučajnim veličinama $\xi(0), \dots, \xi(i-1), Z(0), \dots, Z(i-1)$.

A3: Nelinearna transformacija reziduala merenja $\psi(\cdot)$ je neparna i skoro svugde neprekidna realna funkcija.

A4: Funkcija $\psi(\cdot)$ raste ne brže od linearne funkcije, odnosno apsolutna vrednost ove funkcije zadovoljava uslov $|\psi(z)|\leq k(1+|z|)$; $k\in(0,\infty)$ ili alternativno $|\psi(z)|\leq k'_1(1+k'_2|z|)$; $k'_1\in(0,\infty), k'_2\in[0,\infty)$.

A5: Koeficijent α in (5.87) je pozitivna i ograničena realna veličina, odnosno $\alpha\in(0,k''), k''<\infty$.

A6: Uslovno matematičko očekivanje $\phi_1(z)=E\{\psi(-\xi(i)+z)|F_{i-1}\}$ zadovoljava uslov $z\phi_1(z)\geq\frac{1}{2}\alpha z^2$ za svako $z\neq 0$, gde je koeficijent α definisan u pretpostavci A5.

A7a: Uslovni broj inverzne matrice $\Gamma^{-1}(i)$ matrice pojačanja u (5.87) je konačan, odnosno $\lambda_{\max}\{\Gamma^{-1}(i)\}/\lambda_{\min}\{\Gamma^{-1}(i)\}\leq\gamma$ za svaki diskretni vremenski trenutak i kao i neko $\gamma>0$, gde su $\lambda_{\min}$ i $\lambda_{\max}$ minimalna i maksimalna pozitivna sopstvena vrednost pozitivno definitne matrice $\Gamma^{-1}(i)$.

A7b: Regresioni vektor u modelu (5.83) zadovoljava uslov rasta $\|Z(i)\|^2\leq M\log^\delta r(i-1)$, $M>0, \delta>0$, gde je sa $r(i)$ označen trag inverzne matrice $\Gamma^{-1}(i)$, odnosno $r(i)=Tr\{\Gamma^{-1}(i)\}$ gde $Tr\{\cdot\}$ označava trag odgovarajuće matrice, a simbol $\|\cdot\|$ predstavlja Euklidovu normu.

A8a : Matrica pojačanja algoritma u (5.87) zadovoljava uslove

$$\lim_{i\rightarrow\infty}\lambda_{\min}\{\Gamma^{-1}(i)\}=\infty$$

$$\frac{\sum_{j=1}^i Z(j)Z^T(j)}{Tr\left\{\sum_{j=1}^i Z(j)Z^T(j)\right\}}\geq \varepsilon I$$

za svaki indeks diskretnog vremena $i \geq i_0$ i neko $\varepsilon > 0$, gde je $Z(\cdot)$ regresioni vektor u (5.84), I predstavlja jediničnu matricu odgovarajućih dimenzija, a $\lambda_{\min}(\cdot)$ označava minimalnu sopstvenu vrednost odgovarajuće matrice, dok Tr označava trag odgovarajuće matrice.

A8b : Matrica pojačanja algoritma u (5.87) zadovoljava uslove

$$\lim_{i \rightarrow \infty} \lambda_{\min} \{ \Gamma^{-1}(i) \} = \infty \text{ s.v.1}$$

$$\lim_{i \rightarrow \infty} \frac{\log^k r(i)}{\lambda_{\min} \{ \Gamma^{-1}(i) \}} = 0 \text{ s.v.1}$$

za neko $k > 1 + \delta$, gde je δ definisano pretpostavkom A7b.

Pretpostavke A1 i A2 su uobičajene u reprezentacijama sistema u formi ARX modela (5.81), kao i u teorijskim analizama konvergencije rekurzivnih stohastičkih algoritama za estimaciju parametara u ovakvim modelima, na bazi teorije martingala. Pretpostavka A1 označava da je usvojeni model sistema stabilan u smislu da će ograničeni deterministički pobudni signal generisati ograničen izlazni signal ili odziv sistema (ovakav vid stabilnosti se skraćeno označava sa BIBO, a skraćenica potiče od engleskog termina Bounded Input-Bounded Output). Pretpostavka A2 se odnosi na stohastički pobudni signal i označava da je ovaj signal diskretni beli stohastički proces, nulte srednje vrednosti i konačne varijanse.

Pretpostavke A3 i A4 definišu klasu nelinearnih funkcija u estimacionom algoritmu (5.85) kojima se garantuje konvergencija procena parametara ka njihovim tačnim vrednostima. Mnogobrojne nelinearnosti koje se sreću u literaturi posvećenoj robusnoj estimaciji, kao što su Huberova, Hampelova, Tukijska i Endrujeva nelinearnost, zadovoljavaju ove pretpostavke [33,39,40,75,76]. Pretpostavka A4 označava da $\psi(\cdot)$ funkcija može biti i neograničena, što je slučaj sa linearnom funkcijom, kada se estimacioni algoritam (5.84) svodi na rekurzivni algoritam najmanjih kvadrata. Međutim, sve navedene nelinearnosti u okviru robusne estimacije su ograničene i neprekidne realne funkcije iz praktičnih razloga [4,33,39,40]. Naime, u robusnoj estimaciji osnovni cilj je da se eliminiše uticaj retkih sporadičnih realizacija poremećaja velikog intenziteta, koje su nesaglasne sa glavninom merne populacije, a nazivaju se u engleskoj literaturi outlieri. Izborom ograničene $\psi(\cdot)$ funkcije obezbeđuje se da jedna nesaglasna opservacija neće imati značajniji uticaj na estimacionu proceduru, dok se istovremenim izborom i neprekidne $\psi(\cdot)$ funkcije obezbeđuje da grupisane nesaglasne opservacije neće imati dominantan uticaj na kvalitet estimacije. Sa druge strane, ukoliko je glavnina mernog uzorka raspodeljena po Gausu, tada $\psi(\cdot)$ funkcija za uzorke male i srednje veličine treba da bude linearna, pošto na regularnim merenjima estimaciona procedura treba da se ponaša kao linearni algoritam najmanjih kvadrata. Polazeći od navedenih činjenica, Huber je predložio sledeću nelinearnu transformaciju reziduala merenja sa zasićenjem [40]

$$\psi(z) = \min \left(\frac{z}{\sigma^2}, \frac{k}{\sigma^2} \right) \text{sgn}(z) \quad (5.88)$$

gde je σ standardna devijacija poremećaja $\xi(\cdot)$ u modelu sistema (5.81), a k je podešljivi parametar koji treba da obezbedi efikasan rad estimacionog algoritma kako u slučaju čisto Gausovog poremećaja tako i u prisustvu outliera koji su generisani po nekoj drugoj funkciji raspodele verovatnoće, koja je od interesa u konkretnoj primeni i ima znatno veću varijansu od nominalne normalne funkcije raspodele verovatnoće. S obzirom da varijansa poremećaja obično nije poznata u praksi, uobičajeno je da se ona procenjuje na osnovu medijane apsolutne devijacije, pošto je medijana za razliku od aritmetičke sredine uzorka robusna procena [39,75,94],

$$d = \text{median}(|y(i) - \text{median}(y(i))|) / 0.6745 \quad (5.89)$$

Faktor 0.6745 je uvršten pošto je tada d blisko σ , pod uslovom da uzorci potiču od normalne raspodele i da je dužina uzorka dovoljno velika. Polazeći od ove činjenice, Huber je predložio da se za k usvoji vrednost bliska 1.5, čime se obezbeđuje efikasna estimacija parametara i u slučaju čisto normalno raspodeljenog poremećaja kao i uslovima postojanja malog procenta (do 10%) outliera, t.j. impulsnog šuma.

Pretpostavka A5 se odnosi na vrednost parametra α u izrazu za matricu pojačanja algoritma parametarske estimacije (5.87).

Pretpostavka A6 je nova i karakteristična je za nelinearne rekurzivne estimatore tipa stohastičkog gradijenta, čiji je opšti oblik definisan relacijom (5.83). Da bi ova pretpostavka bila zadovoljena neophodno je da uslovno matematičko očekivanje

$$\phi_1(a) = \int_0^\infty [\psi(u+a) - \psi(u-a)] dP(u)$$

bude monotono ne opadajuća realna funkcija realnog argumenta. U gornjem izrazu $P(\cdot)$ predstavlja funkciju raspodele poremećaja u modelu (5.80). Imajući u vidu pretpostavke A2 i A3, ovaj zahtev će biti ispunjen ako nelinearna $\psi(\cdot)$ funkcija i funkcija raspodele verovatnoće $P(\cdot)$ poremećaja $\xi(\cdot)$ u modelu (5.80)-(5.84) imaju zajedničku tačku rasta, što se svodi na zadovoljenje uslova

$$\psi(z+\varepsilon) > \psi(z-\varepsilon), \quad P(z+\varepsilon) > P(z-\varepsilon) \quad \text{za neko } z \text{ i svako } \varepsilon > 0,$$

na osnovu koga dalje sledi

$$a\phi_1(a) > 0 \quad \text{za } a \neq 0 \text{ i } \phi_1(0) = 0.$$

Na taj način, pretpostavka A6 će biti ispunjena ako je nelinearna $\psi(\cdot)$ funkcija neparna, skoro svuda neprekidna, monotono rastuća i u segmentima neprekidno diferencijabilna. Kao što je ranije istaknuto, iz praktičnih razloga je poželjno da ova funkcija bude i ograničena ($k_2 = 0$ u A4).

Posebno ako je poremećaj raspodeljen po Gausovom zakonu raspodele verovatnoće, okarakterisanom sa nultom srednjom vrednošću i varijansom σ^2 , a za nelinearnu transformaciju reziduala se usvoji funkcija verodostojnosti

$$\psi(z) = -[\log p(z)]' = z/\sigma^2, \quad p(z) = dP(z)/dz$$

i usvoji $\alpha = E\{\psi'(\xi(i))\} = \sigma^{-2}$, algoritam (5.85)-(5.87) se svodi na rekurzivni algoritam najmanjih kvadrata. Dalje se dobija da je $\phi_1(z) = z/\sigma^2$, što znači da je uslov A6 $a\phi_1(a) > 0$ za $a \neq 0$ i $\phi_1(0) = 0$, ispunjen i da on važi i za linearne parametarske estimatore razmatranog tipa.

Pretpostavka A7a je prilično stroga i odnosi se na uslovni broj inverzne matrice pojačanja algoritma (5.87). Ovakav uslov može da prouzrokuje određene probleme prilikom implementacije algoritma, pošto se mogu pojaviti određene poteškoće numeričke prirode, prouzrokovane relativno velikom vrednošću uslovnog broja ove matrice. Stoga, praktična realizacija algoritma može zahtevati da se vrši neprekidna inspekcija navedenog uslovnog broja. Sa druge strane, intuitivno je jasno da pobudni deterministički signal mora posedovati određene karakteristike da bi proces identifikacije parametara bio uspešan. Očigledno je da pobudni signal koji je sve vreme jednak nuli neće moći da generiše relevantne informacije o ulazno-izlaznim osobinama razmatranog sistema, što će rezultovati u loš kvalitet estimacije parametara u usvojenom modelu sistema. Kao što je istaknuto u prethodnom poglavlju, izbor pobudnog signala je vezan se zahtevom perzistentne eksitacije, tako da pobudni signal treba da bude dovoljno bogat, odnosno da poseduje zadovoljavajuće dinamičke karakteristike, kako bi mogao da pobudi sve relevantne modove sistema i kako bi se svi parametri odslikavali u izlaznom signalu, odnosno odzivu sistema. Uslov A8a predstavlja matematički iskaz zahteva o perzistentnoj eksitaciji, čime je garantovana identifikabilnost parametara u usvojenoj strukturi modela.

Navedene teškoće mogu se prevazići ukoliko se uslov perzistentne eksitacije matematički oslabi. Koristeći poznat rezultat iz literature, u kome je pokazano da greška estimacije rekurzivnih algoritama tipa stohastičkog gradijenta pod određenim uslovima zadovoljava uslov [82]

$$\|\hat{\theta}(i) - \theta\| = O\left(\left(\frac{\log(i)}{\lambda_i}\right)^{\frac{1}{2}}\right) \text{ s.v.1, } \lim_{x \rightarrow \infty} \frac{O(\|x\|)}{\|x\|} = 0$$

pri čemu simbol $\|\cdot\|$ označava Euklidovu normu vektora, a λ_i je minimalna pozitivna sopstvena vrednost matrice

$$\Gamma^{-1}(i) = \sum_{j=1}^i Z(j)Z^T(j)$$

Uslovi A7a i A8a zamenjeni su uslovima A7b i A8b, pri čemu uslov A7b nije restriktivan, kao uslov A7a, pošto je M proizvoljna pozitivna konstanta. Na taj način, uslov A8b predstavlja oslabljenu verziju zahteva perzistentne eksitacije.

Rezultati lema 1, 2, 3 i 4 i navedene pretpostavke A1-A8 mogu se dalje primeniti u cilju rigoroznog dokazivanja konvergencije rekurzivnog stohastičkog algoritma (5.85)-(5.87). Rezultati su sistematizovani u okviru sledeće teoreme.

Teorema 5.1: Pretpostavimo da model sistema (5.80)-(5.84) i rekurzivni stohastički algoritam za estimaciju nepoznatih parametara modela (5.85)-(5.87) ispunjavaju hipoteze A1-A6, A7a i A8a. Tada vektor procena parametara

$\hat{\Theta}(i)$ konvergira ka vektoru tačnih vrednosti parametara Θ koji se estimiraju sa verovatnoćom 1 (s.v.1), odnosno

$$P\left\{\lim_{i \rightarrow \infty} \hat{\theta}(i) = \theta\right\} = 1.$$

Dokaz: Ako označimo grešku estimacije u i -tom diskretnom trenutku sa $\tilde{\Theta}(i) = \hat{\Theta}(i) - \Theta$ i uvedemo težinsku sumu kvadrata pojedinih komponenti vektora greške u obliku kvadratne forme, koja se naziva i stohastička funkcija Ljapunova, $V(i) = \tilde{\Theta}^T(i) \Gamma^{-1}(i) \tilde{\Theta}(i)$, na osnovu (5.85) se dobija

$$\begin{aligned} V(i) = & \tilde{\Theta}^T(i-1) \Gamma^{-1}(i) \tilde{\Theta}(i-1) + 2\tilde{\Theta}^T(i-1) Z(i) \psi(v(i, \hat{\Theta}(i-1))) \\ & + Z^T(i) \Gamma(i) Z(i) \psi^2(v(i, \hat{\Theta}(i-1))) \end{aligned} \quad (5.90)$$

Primenjujući dalje lemu o inverziji matrica, relacija (5.87) može se zapisati u alternativnom obliku

$$\Gamma^{-1}(i) = \Gamma^{-1}(i-1) + \alpha Z(i) Z^T(i) \quad (5.91)$$

Dodajući i oduzimajući $V(i-1)$ desnoj strani relacije (5.90) i koristeći (5.91), dalje se može pisati

$$\begin{aligned} V(i) = & V(i-1) + 2\tilde{\Theta}^T(i-1) Z(i) \psi(v(i, \hat{\Theta}(i-1))) \\ & + \alpha (\tilde{\Theta}^T(i-1) Z(i))^2 + Z^T(i) \Gamma(i) Z(i) \psi^2(v(i, \hat{\Theta}(i-1))) \end{aligned} \quad (5.92)$$

Definišimo uslovno matematičko očekivanje slučajne varijable koja se dobija propuštanjem reziduala merenja (5.86) kroz kvadrat nelinearne transformacije

$$\phi_2(\tilde{\Theta}^T(i-1) Z(i)) = E\left\{\psi^2(-v(i, \hat{\Theta}(i-1))) \middle| F_{i-1}\right\} \quad (5.93)$$

gde se rezidual merenja može zapisati i u alternativnom obliku

$$v(i, \hat{\Theta}(i-1)) = -\tilde{\Theta}^T(i-1) Z(i) + \xi(i) \quad (5.94)$$

Na taj način, razmatrano matematičko očekivanje je definisano u odnosu na f.g.v slučajne varijable $\xi(i)$. Na osnovu hipoteze A4 se zaključuje da realna funkcija realnog argumenta (5.93) zadovoljava uslov

$$\phi_2(\tilde{\Theta}^T(i-1) Z(i)) \leq k_1 \left[1 + (\tilde{\Theta}^T(i-1) Z(i))^2\right] \quad (5.95)$$

gde je k_1 konačna i pozitivna konstanta. Ako se relacija (5.92) podeli sa tragom inverzne matrice pojačanja algoritma $r(i-1) = \text{Tr}\{\Gamma^{-1}(i-1)\}$ i primeni operator uslovnog matematičkog očekivanja $E\{./F_{i-1}\}$, kao i hipoteze A2, A3 i relacije (5.94)-(5.95), dobija se

$$E \left\{ \frac{V(i)}{r(i)} | F_{i-1} \right\} \leq \frac{V(i-1)}{r(i-1)} - \frac{2\tilde{\Theta}^T(i-1)Z(i)}{r(i-1)} \left[\phi_1(\tilde{\Theta}^T(i-1)Z(i)) - \frac{\alpha}{2} \tilde{\Theta}^T(i-1)Z(i) \right] \\ + k_1 (\tilde{\Theta}^T(i-1)Z(i))^2 \frac{Z^T(i)\Gamma(i)Z(i)}{r(i-1)} + k_1 \frac{Z^T(i)\Gamma(i)Z(i)}{r(i-1)} \quad (5.96)$$

Prilikom izvodjenja relacije (5.96) iskorišćena je činjenica da je $r(i) \geq r(i-1)$, na osnovu čega se zaključuje da je u važnosti $V(i)/r(i) \leq V(i)/r(i-1)$. Takođe, imajući u vidu da je

$$\lambda_{\max} \{ \Gamma(i) \} = \lambda_{\min}^{-1} \{ \Gamma^{-1}(i) \}; \lambda_{\min} \{ \Gamma(i) \} = \lambda_{\max}^{-1} \{ \Gamma^{-1}(i) \} \quad (5.97)$$

na osnovu hipoteze A7a sledi

$$\frac{\lambda_{\max} \{ \Gamma(i) \}}{\lambda_{\min} \{ \Gamma(i) \}} \leq \gamma \quad (5.98)$$

Sa druge strane, pošto se može pisati

$$r(i) = \text{Tr} \{ \Gamma^{-1}(i) \} = \sum_{j=1}^{n+m} \lambda_j \{ \Gamma^{-1}(i) \} \quad (5.99)$$

gde je $\lambda_j \{ \cdot \}$ sopstvena vrednost matrice $\Gamma^{-1}(i)$, zaključuje se da je u važnosti

$$\lambda_{\max} \{ \Gamma^{-1}(i-1) \} \geq \frac{r(i-1)}{n+m}. \quad (5.100)$$

Dokažimo dalje da je regresioni vektor (5.84) ograničen, odnosno da važi

$$\sup_i \|Z(i)\|^2 \leq k_2 \quad \text{s.v. } 1 \quad (5.101)$$

gde je k_2 pozitivna, konačna konstanta. Dokaz zahteva da se ulazno-izlazna reprezentacija modela sistema (5.80) transformiše u ekvivalentan model u prostoru stanja [6,24,48]. Kao što je dobro poznato u teoriji sistema, diferencna jednačina (5.80) može se predstaviti vektorskom diferencnom jednačinom prvog reda, koja se naziva vektorskom jednačinom stanja

$$X(i+1) = EX(i) + FU(i) + W(i) \quad (5.102)$$

gde je vektor stanja

$$X^T(i) = [y(i) y(i-1) \cdots y(i-n-1)]$$

vektor upravljanja

$$U^T(i) = [u(i) u(i-1) \cdots u(i-m+1)]$$

vektor poremećaja

$$W^T(i) = [\xi(i+1) \ 0 \ \cdots \ 0]$$

a odgovarajuće matrice u modelu stanja definisane su izrazima .

$$E = \begin{bmatrix} -a_1 & -a_2 & -a_3 & \cdots & -a_{n-1} & -a_n \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad F = \begin{bmatrix} b_1 & b_2 & \cdots & b_m \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Rešenje vektorske jednačine stanja (5.102) dato je sledećim izrazom [16,24,48,49]

$$\begin{aligned} X(i) &= E^i X(0) + \sum_{j=0}^{i-1} E^j F U(i-j+1) + \sum_{j=0}^{i-1} E^j W(i-j+1) \\ &= E^i X(0) + \sum_{j=0}^{i-1} E^{i-j-1} [F U(j) + W(j)] \end{aligned} \quad (5.103)$$

Na osnovu hipoteze A1 se zaključuje da je Euklidova norma $\|E^i\| \leq \lambda^i$, gde je $\lambda < 1$ i $i \geq 0$, te se na osnovu (5.103) dobija

$$\begin{aligned} \|X(i)\| &= \|E^i\| \|X(0)\| + \sum_{r=0}^{i-1} \|E^{i-r-1}\| [\|F\| \|U(r)\| + \|W(r)\|] \\ &\leq C_1 \lambda^i + (C_2 + C_3) \sum_{r=0}^{i-1} \lambda^{i-r-1} < \infty \end{aligned} \quad (5.104)$$

gde su C_1 , C_2 i C_3 neke pozitivne konstante. Prilikom izvodjenja (5.104) uzeto je u obzir da je na osnovu hipoteza A1 i A2 $\|F\| \|U(r)\| \leq C_2$ i $\|W(r)\| \leq C_3$. Medjutim, pošto je na osnovu (5.84) regresioni vektor $Z(i)$ u (5.101) sastavljen od diskretnih merenja izlaza $y(\cdot)$ i ulaza $u(\cdot)$, relacija (5.101) je direktna posledica izraza (5.104).

Analizirajmo dalje treći član u izrazu (5.96)

$$\begin{aligned} k_1 k_2 (\tilde{\Theta}(i-1) Z(i))^T \frac{Z^T(i) \Gamma(i) Z(i)}{r(i-1)} = \\ k_1 k_2 \tilde{\Theta}^T(i-1) \Gamma^{-\frac{1}{2}}(i-1) \Gamma^{\frac{1}{2}}(i-1) Z(i) Z^T(i) \Gamma^{\frac{1}{2}}(i-1) \Gamma^{-\frac{1}{2}}(i-1) \tilde{\Theta}(i-1) \frac{Z^T(i) \Gamma(i) Z(i)}{r(i-1)} \end{aligned} \quad (5.105)$$

U tom cilju primenićemo dobro poznate rezultate iz matrične analize, na osnovu kojih se za kvadratnu matricu A i vektor kolonu b može pisati [5,92]

$$\begin{aligned} |\lambda_{\min}| \|b\| &\leq \|Ab\| \leq |\lambda_{\max}| \|b\| \\ |\lambda_{\min}| \|b\|^2 &\leq |b^T Ab| \leq |\lambda_{\max}| \|b\|^2 \end{aligned} \quad (5.106)$$

gde su $\lambda_{\min}$ i $\lambda_{\max}$ sopstvene vrednosti matrice A koje imaju najmanju i najveću apsolutnu vrednost, respektivno, a kvadrat norme $\|b\|^2 = b^T b$. Na taj način, ako se usvoji $b = \Gamma^{-1/2}(i-1) \tilde{\Theta}(i-1)$, $A = \Gamma^{1/2}(i-1) Z(i) Z^T(i) \Gamma^{1/2}(i-1)$ na osnovu druge relacije u (5.106) se dobija

$$\begin{aligned}
 & \tilde{\Theta}^T(i-1)\Gamma^{-\frac{1}{2}}(i-1)\Gamma^{\frac{1}{2}}(i-1)Z(i)Z^T(i)\Gamma^{\frac{1}{2}}(i-1)\Gamma^{-\frac{1}{2}}(i-1)\tilde{\Theta}(i-1) \leq \\
 & \lambda_{\max} \left\{ \Gamma^{\frac{1}{2}}(i-1)Z(i)Z^T(i)\Gamma^{\frac{1}{2}}(i-1) \right\} \left\| \Gamma^{-\frac{1}{2}}(i-1)\tilde{\Theta}(i-1) \right\|^2 = \\
 & \lambda_{\max} \left\{ \Gamma^{\frac{1}{2}}(i-1)Z(i)Z^T(i)\Gamma^{1/2}(i-1) \right\} \tilde{\Theta}^T(i-1)\Gamma^{-\frac{1}{2}}(i-1)\Gamma^{-\frac{1}{2}}(i-1)\tilde{\Theta}(i-1) \leq \quad (5.107) \\
 & \text{Tr} \left\{ \Gamma^{\frac{1}{2}}(i-1)Z(i)Z^T(i)\Gamma^{\frac{1}{2}}(i-1) \right\} \tilde{\Theta}^T(i-1)\Gamma^{-1}(i-1)\tilde{\Theta}(i-1) = \\
 & \text{Tr} \left\{ \Gamma^{\frac{1}{2}}(i-1)Z(i)Z^T(i)\Gamma^{\frac{1}{2}}(i-1) \right\} V(i-1)
 \end{aligned}$$

Imajući u vidu poznat rezultat iz matrične algebre

$$\text{Tr}\{AB\} = \text{Tr}\{BA\}$$

i usvajajući

$$B = \Gamma^{1/2}(i-1), \quad A = \Gamma^{1/2}(i-1)Z(i-1)Z^T(i-1)$$

dalje se može pisati

$$\text{Tr}\{\Gamma^{1/2}(i-1)Z(i)Z^T(i)\Gamma^{1/2}(i-1)\} = \text{Tr}\{\Gamma(i-1)Z(i)Z^T(i)\} \quad (5.108)$$

Iz matrične analize je takodje poznato da se za kvadratnu matricu A i vektor kolonu b može pisati

$$\text{Tr}\{Abb^T\} = b^T Ab$$

,pa se usvajajući

$$A = \Gamma(i-1), \quad b = Z(i)$$

i primenjujući (5.106), dobija iz (5.108)

$$\text{Tr}\{\Gamma(i-1)Z(i)Z^T(i)\} \leq \lambda_{\max}\{\Gamma(i-1)\}\|Z(i)\|^2 \quad (5.109)$$

Zamenom relacije (5.109) u izraz (5.107), uz primenu (5.98), dobija se

$$\begin{aligned}
 & \tilde{\Theta}^T(i-1)\Gamma^{-\frac{1}{2}}(i-1)\Gamma^{\frac{1}{2}}(i-1)Z(i)Z^T(i)\Gamma^{\frac{1}{2}}(i-1)\Gamma^{-\frac{1}{2}}(i-1)\tilde{\Theta}(i-1) \leq \\
 & \lambda_{\max}\{\Gamma(i-1)\}\|Z(i)\|^2 V(i-1) \leq \gamma \lambda_{\min}\{\Gamma(i-1)\}\|Z(i)\|^2 V(i-1) = \quad (5.110) \\
 & = \gamma \lambda_{\max}^{-1}\{\Gamma^{-1}(i-1)\}\|Z(i)\|^2 V(i-1)
 \end{aligned}$$

Na taj način, zamenom (5.110) u (5.105), uz primenu relacija (5.99) i (5.100), dalje sledi

$$\begin{aligned}
 & k_1 k_2 \left(\tilde{\Theta}^T(i-1)Z(i) \right)^2 \frac{Z^T(i)\Gamma(i)Z(i)}{r(i-1)} \leq \gamma k_1 k_2 \frac{V(i-1)}{r(i-1)} \frac{1}{\lambda_{\max}\{\Gamma^{-1}(i-1)\}} \|Z(i)\|^2 Z^T(i)\Gamma(i)Z(i) \\
 & \leq k_3 \frac{V(i-1)}{r(i-1)} \frac{Z^T(i)\Gamma(i)Z(i)}{r(i-1)} \quad (5.111)
 \end{aligned}$$

gde je $k_3 = \gamma k_1 k_2 (n+m)$. Polazeći od (5.111), relacija (5.96) može se alternativno napisati

$$\begin{aligned}
 E \left\{ \frac{V(i)}{r(i)} \middle| F_{i-1} \right\} &\leq \frac{V(i-1)}{r(i-1)} \left[1 + k_3 \frac{Z^T(i) \Gamma(i) Z(i)}{r(i-1)} \right] \\
 &- \frac{2\tilde{\Theta}^T(i-1) Z(i)}{r(i-1)} \left[\phi_1(\tilde{\Theta}^T(i-1) Z(i)) - \frac{\alpha}{2} \tilde{\Theta}^T(i-1) Z(i) \right] + \\
 &+ k_1 \frac{Z^T(i) \Gamma(i) Z(i)}{r(i-1)}
 \end{aligned} \tag{5.112}$$

Takodje, i relacija (5.87) se može pisati u alternativnom obliku

$$\Gamma(i) = \Gamma(i-1) \left[1 + \alpha Z^T(i) \Gamma(i-1) Z(i) \right]^{-1} \tag{5.113}$$

Uz navedeno, pošto je

$$\lambda_{\min} \{ \Gamma(i) \} = \lambda_{\max}^{-1} \{ \Gamma^{-1}(i) \} \geq r^{-1}(i) \tag{5.114}$$

zaključuje se da je

$$r^{-1}(i) I \leq \lambda_{\min} \{ \Gamma(i) \} I \leq \Gamma(i), \tag{5.115}$$

gde je I jedinična matrica odgovarajućih dimenzija. Na osnovu (5.91) (5.113) (5.115),
takodje, se dobija

$$\frac{Z^T(i) \Gamma(i) Z(i)}{r(i-1)} \leq \frac{Z^T(i) \Gamma^2(i-1) Z(i) \alpha^{-1}}{\alpha^{-1} + Z^T(i) \Gamma(i-1) Z(i)} = Z^T(i) \Gamma(i) \Gamma(i-1) Z(i) \tag{5.116}$$

Polazeći od poznatog rezultata iz matrične analize [2,5]

$$a^T b = b^T a = \text{Tr} \{ a b^T \},$$

i usvajajući

$$a^T = Z^T(i) \Gamma(i), \quad b = \Gamma(i-1) Z(i),$$

na osnovu (5.91) sledi

$$\begin{aligned}
 Z^T(i) \Gamma(i) \Gamma(i-1) Z(i) &= \alpha^{-1} \text{Tr} \{ \Gamma(i) \alpha Z(i) Z^T(i) \Gamma(i-1) \} = \\
 &= \alpha^{-1} \text{Tr} \{ \Gamma(i) [\Gamma^{-1}(i) - \Gamma^{-1}(i-1)] \Gamma(i-1) \} = \alpha^{-1} [\text{Tr} \{ \Gamma(i-1) \} - \text{Tr} \{ \Gamma(i) \}]
 \end{aligned} \tag{5.117}$$

Ako se sumira relacija (5.117) i primeni hipoteza A5, dobija se

$$\sum_{i=1}^{\infty} \frac{Z^T(i) \Gamma(i) Z(i)}{r(i-1)} \leq \alpha^{-1} \text{Tr} \{ \Gamma(0) \} < \infty \tag{5.118}$$

Da bismo primenili leme o konvergenciji martingala uvedimo oznake

$$\begin{aligned}
 z_{i-1} &= V(i-1) / r(i-1); \\
 \beta_{i-1} &= k_3 r^{-1}(i-1) Z^T(i) \Gamma(i) Z(i); \\
 \xi_{i-1} &= k_1 r^{-1}(i-1) Z^T(i) \Gamma(i) Z(i); \\
 \zeta_{i-1} &= 2r^{-1}(i-1) \tilde{\Theta}^T(i-1) Z(i) \left[\phi_1(\tilde{\Theta}^T(i-1) Z(i)) - \frac{\alpha}{2} \tilde{\Theta}^T(i-1) Z(i) \right]
 \end{aligned} \tag{5.119}$$

Na osnovu hipoteze A6 zaključuje se da su $z_{i-1}, \beta_{i-1}, \xi_{i-1}$ i ζ_{i-1} nenegativne i F_{i-1} -
merljive slučajne varijable. Takodje, na osnovu (5.118) sledi da je

$$\sum_{i=1}^{\infty} \beta_i < \infty ; \quad \sum_{i=1}^{\infty} \xi_i < \infty \tag{5.120}$$

Na taj način, A6 i (5.120) imaju za posledicu da (5.119) zadovoljava pretpostavke
leme 5.2, pa se rezultati ove leme mogu primeniti na (5.112), čime se dobija

$$\lim_{i \rightarrow \infty} r^{-1}(i)V(i) = V^* \quad \text{s.v.1} \quad (5.121)$$

$$\sum_{i=1}^{\infty} \frac{\tilde{\Theta}^T(i-1)Z(i)}{r(i-1)} \frac{\zeta_{i-1}r(i-1)}{2\tilde{\Theta}^T(i-1)Z(i)} < \infty \quad \text{s.v.1} \quad (5.122)$$

gde je V^* konačna slučajna konstanta. Na osnovu A6 se takodje zaključuje da postoji realan broj $\rho > 1$ takav da je

$$\inf_i \Theta^T(i-1)Z(i)\phi_1(\tilde{\Theta}^T(i-1)Z(i)) = \rho \frac{\alpha}{2} (\tilde{\Theta}^T(i-1)Z(i))^2 \quad (5.123)$$

Na osnovu (5.122) i (5.123) sledi

$$\frac{(1-\rho)\alpha}{2} \sum_{i=1}^{\infty} \frac{(\tilde{\Theta}^T(i-1)Z(i))^2}{r(i-1)} < \infty \quad \text{s.v.1} \quad (5.124)$$

odakle se zaključuje da je

$$\sum_{i=1}^{\infty} \frac{(\tilde{\Theta}^T(i-1)Z(i))^2}{r(i-1)} < \infty \quad \text{s.v.1} \quad (5.125)$$

Polazeći od A7a, (5.100) and (5.106) takodje sledi

$$\frac{V(i)}{r(i)} \geq \frac{\lambda_{\min}\{\Gamma^{-1}(i)\} \|\tilde{\Theta}(i)\|^2}{(n+m)\lambda_{\max}\{\Gamma^{-1}(i)\}} \geq \frac{\|\tilde{\Theta}(i)\|^2}{(n+m)\gamma} \quad (5.126)$$

odakle se, kao direktna posledica (5.121), dobija

$$\lim_{i \rightarrow \infty} \|\tilde{\Theta}(i)\| = \|\Theta^*\|^2 \quad \text{s.v.1} \quad (5.127)$$

gde je Θ^* konačan i konstantan vektor. Sa druge strane, na osnovu (5.91) se zaključuje da je

$$r(i) = \text{Tr}\{\Gamma^{-1}(i)\} = r(i-1) + \alpha Z^T(i)Z(i) \quad (5.128)$$

niz neopadajućih realnih brojeva, t.j. $r(i+1) \geq r(i)$, i $r(i) \geq \lambda_{\min}\{\Gamma^{-1}(i)\}$, tako da se na osnovu prvog dela hipoteze A8a zaključuje da je $\lim_{i \rightarrow \infty} r(i) = \infty$. Izvedeni zaključak pokazuje da su ispunjeni uslovi leme 5.3 i da se ona može primeniti na (5.125), na osnovu čega se dobija

$$\lim_{i \rightarrow \infty} \frac{\sum_{k=1}^i \tilde{\Theta}^T(k-1)Z(k)Z^T(k)\tilde{\Theta}(k-1)}{r(i)} = 0 \quad \text{s.v.1} \quad (5.129)$$

Uz navedeno, pretpostavka A8a implicira da se za dovoljno veliko i može pisati

$$\begin{aligned} \sum_{k=1}^i \tilde{\Theta}^T(k-1)Z(k)Z^T(k)\tilde{\Theta}(k-1) &\geq \Theta^{*T} \sum_{k=1}^i Z(k)Z^T(k)\Theta^* \\ &\geq \Theta^{*T} \varepsilon \text{ITr} \left\{ \sum_{k=1}^i Z(k)Z^T(k) \right\} \Theta^* = \varepsilon \|\Theta^*\|^2 \sum_{k=1}^i \text{Tr}\{Z(k)Z^T(k)\} \\ &= \varepsilon \|\Theta^*\|^2 \sum_{k=1}^i \|Z(k)\|^2 \end{aligned} \quad (5.130)$$

Konačno, na osnovu (5.129) i (5.130) sledi

$$\lim_{i \rightarrow \infty} \frac{\|\Theta^*\|^2 \sum_{k=1}^{\infty} \|Z(k)\|^2}{r(0) + \alpha \sum_{k=1}^i \|Z(k)\|^2} = 0 \quad \text{s.v.1} \quad (5.131)$$

odakle se zaključuje da je $\Theta^* = 0$, čime je dokaz teoreme završen.

Pretpostavke pod kojima je izvedena teorema 5.1 imaju za posledicu da je regresioni vektor (5.84) ograničen, što je matematički formulisano kroz relaciju (5.101), koja se alternativno može zapisati i u obliku

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N [y^2(k) + u^2(k)] < \infty \quad \text{s.v.1}$$

gde su $y(\cdot)$ i $u(\cdot)$ merljivi izlaz i ulaz sistema, respektivno, kao i da asimptotska autokovarijaciona matrica regresionog vektora

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N Z(k) Z^T(k) = \Gamma > 0 \quad \text{s.v.1}$$

gde simbol >0 označava pozitivno definitnu matricu. Primetimi, takodje, da teorema važi samo za stabilne sisteme, pri čemu je njen osnovni nedostatak, sa stanovišta primene, ograničenost uslovnog broja inverzne matrice pojačanja algoritma, što može dovesti do numeričkih poteškoća u situacijama kada ovaj parametar postane suviše veliki. Uslovi teoreme mogu se oslabiti, tako da se tako modifikovana teorema može primenjivati i za nestabilne sisteme, pri čemu varijansa slučajne pobude ili poremećaja ne mora biti čak ni konačna u slučaju da je nelinearna transformacija reziduala ograničena funkcija. Ovako modifikovana teorema prikazana je u nastavku.

Teorema 5.2: Posmatrajmo model sistema (5.80)-(5.84) i algoritam za estimaciju nepoznatih parametara modela (5.85)-(5.87). Ukoliko model i algoritam estimacije zadovoljavaju pretpostavke A2-A6, A7b i A8b, tada vektor procenjenih parametara $\hat{\Theta}(i)$ konvergira skoro sigurno (sa verovatnoćom 1 ili skraćeno s.v.1) ka vektoru tačnih parametara Θ , odnosno

$$P\left\{\lim_{i \rightarrow \infty} \hat{\Theta}(i) = \Theta\right\} = 1 \quad \text{s.v.1.}$$

Dokaz: Dokaz metodološki sledi postupak koji je primenjen prilikom dokazivanja teoreme 1. Naime, ako se sa $\tilde{\Theta}(i) = \hat{\Theta}(i) - \Theta$ označi greška estimacije i uvede stohastička funkcija Ljapunova $V(i) = \tilde{\Theta}^T(i) \Gamma^{-1}(i) \tilde{\Theta}(i)$, na osnovu (5.85) se dobija (5.90), odnosno

$$V(i) = \tilde{\Theta}^T(i-1) \Gamma^{-1}(i) \tilde{\Theta}(i-1) + 2 \tilde{\Theta}^T(i-1) Z(i) \psi\left(v(i, \hat{\Theta}(i-1))\right) + Z^T(i) \Gamma(i) Z(i) \psi^2\left(v(i, \hat{\Theta}(i-1))\right) \quad (5.132)$$

Analogno sa dokazom teoreme 5.1, ako se desnoj strani relacije (5.132) doda i oduzme $V(i-1)$ i uvrsti izraz (5.91) za inverznu matricu pojačanja estimacione procedure, dobija se (5.92), t.j.

$$V(i) = V(i-1) + 2\theta^T(i-1)Z(i)\psi(v(i, \hat{\Theta}(i-1))) + \alpha(\theta^T(i-1)Z(i))^2 + Z^T(i)\Gamma(i)Z(i)\psi^2(v(i, \hat{\Theta}(i-1))) \quad (5.133)$$

Po analogiji sa teoremom 1, uvedimo funkciju

$$\phi_2(\theta^T(i-1)Z(i)) = E\left\{\psi^2(-v(i, \hat{\Theta}(i-1))) \middle| F_{i-1}\right\} \quad (5.134)$$

gde je rezidul merenja dat relacijom (5.94), odnosno

$$v(i, \hat{\Theta}(i-1)) = -\theta^T(i-1)Z(i) + \xi(i). \quad (5.135)$$

Funkcija $\phi_2(\cdot)$ postoji pod istim uslovima kao i funkcija $\phi_1(\cdot)$ u hipotezi A6. Uzimajući, dalje, u obzir hipotezu A4, zaključuje se da je

$$\phi_2(\theta^T(i-1)Z(i)) \leq k_1 \left[1 + k_2 (\theta^T(i-1)Z(i))^2 \right] \quad (5.136)$$

gde su k_1 i k_2 pozitivna i nenegativna konačna konstanta, respektivno. Imajući u vidu da je za proizvoljan vector kolonu b u važnosti $Tr\{bb^T\} = b^T b$, ako se na izraz (5.91) primeni operacija traga matrice i usvoji $b = Z(i)$, dobija se

$$r(i) = r(i-1) + \alpha Z^T(i)Z(i) \quad (5.137)$$

Slično kao u teoremi 5.1, ako se sada (5.133) podeli sa $\log^k r(i)$, umesto sa $r(i)$, primeni opertor uslovnog matematičko očekivanja $E\{ \cdot / F_{i-1} \}$, kao i primene hipoteze A2, A3 i uvrste relacije (5.134)-(5.136), može se pisati

$$\begin{aligned} E\left\{ \frac{V(i)}{\log^k r(i)} \middle| F_{i-1} \right\} &\leq \frac{V(i-1)}{\log^k r(i)} - \frac{2\theta^T(i-1)Z(i)}{\log^k r(i)} \left[\phi_1(\theta^T(i-1)Z(i)) - \frac{\alpha}{2} \theta^T(i-1)Z(i) \right] \\ &+ k_1 k_2 (\theta^T(i-1)Z(i))^2 \frac{Z^T(i)\Gamma(i)Z(i)}{\log^k r(i)} + k_1 \frac{Z^T(i)\Gamma(i)Z(i)}{\log^k r(i)} \end{aligned} \quad (5.138)$$

Po analogiji sa teoremom 5.1, analizirajmo dalje treći član sa desne strane relacije (5.138)

$$\begin{aligned} k_1 k_2 (\theta^T(i-1)Z(i))^2 \frac{Z^T(i)\Gamma(i)Z(i)}{\log^k r(i)} = \\ k_1 k_2 \theta^T(i-1)\Gamma^{-1/2}(i-1)\Gamma^{1/2}(i-1)Z(i)Z^T(i)\Gamma^{1/2}(i-1)\Gamma^{-1/2}(i-1)\theta(i-1) \frac{Z^T(i)\Gamma(i)Z(i)}{\log^k r(i)} \end{aligned} \quad (5.139)$$

U tom cilju primenimo izraz (5.106), odnosno

$$|\lambda_{\min}| \|b\| \leq \|Ab\| \leq |\lambda_{\max}| \|b\|, \quad |\lambda_{\min}| \|b\|^2 \leq b^T Ab \leq |\lambda_{\max}| \|b\|^2$$

uz izbor

$$b = \Gamma^{-1/2}(i-1) \mathcal{O}(i-1), \quad A = \Gamma^{1/2}(i-1) Z(i) Z^T(i) \Gamma^{1/2}(i-1)$$

na osnovu čega sledi

$$\begin{aligned} & \mathcal{O}^T(i-1) \Gamma^{-\frac{1}{2}}(i-1) \Gamma^{\frac{1}{2}}(i-1) Z(i) Z^T(i) \Gamma^{\frac{1}{2}}(i-1) \Gamma^{-\frac{1}{2}}(i-1) \mathcal{O}(i-1) \leq \\ & \lambda_{\max} \left\{ \Gamma^{\frac{1}{2}}(i-1) Z(i) Z^T(i) \Gamma^{\frac{1}{2}}(i-1) \right\} \left\| \Gamma^{-\frac{1}{2}}(i-1) \mathcal{O}(i-1) \right\|^2 = \\ & \lambda_{\max} \left\{ \Gamma^{\frac{1}{2}}(i-1) Z(i) Z^T(i) \Gamma^{\frac{1}{2}}(i-1) \right\} \mathcal{O}^T(i-1) \Gamma^{-\frac{1}{2}}(i-1) \Gamma^{\frac{1}{2}}(i-1) \mathcal{O}(i-1) \leq \\ & \text{Tr} \left\{ \Gamma^{\frac{1}{2}}(i-1) Z(i) Z^T(i) \Gamma^{\frac{1}{2}}(i-1) \right\} \mathcal{O}^T(i-1) \Gamma^{-1}(i-1) \mathcal{O}(i-1) = \\ & \text{Tr} \left\{ \Gamma^{\frac{1}{2}}(i-1) Z(i) Z^T(i) \Gamma^{\frac{1}{2}}(i-1) \right\} V(i-1) \end{aligned} \quad (5.140)$$

Takodje, pošto je $\text{Tr}\{AB\} = \text{Tr}\{BA\}$, ako se usvoji

$$B = \Gamma^{1/2}(i-1), \quad A = \Gamma^{1/2}(i-1) Z(i) Z^T(i)$$

dobija se

$$\text{Tr} \left\{ \Gamma^{\frac{1}{2}}(i-1) Z(i) Z^T(i) \Gamma^{\frac{1}{2}}(i-1) \right\} = \text{Tr} \{ \Gamma(i-1) Z(i) Z^T(i) \} \quad (5.141)$$

Sa druge strane, za proizvoljnu kvadratnu matricu A i proizvoljan vektor kolonu b može se pisati $\text{Tr}\{Abb^T\} = b^T Ab$, te se izborom $A = \Gamma(i-1)$ i $b = Z(i)$ i primenom relacije (5.106) iz (5.141) takodje dobija

$$\text{Tr} \{ \Gamma(i-1) Z(i) Z^T(i) \} \leq \lambda_{\max} \{ \Gamma(i-1) \} \|Z(i)\|^2 \quad (5.142)$$

Zamenom izraza (5.142) u relaciju (5.140), može se pisati

$$\mathcal{O}^T(i-1) \Gamma^{-\frac{1}{2}}(i-1) \Gamma^{\frac{1}{2}}(i-1) Z(i) Z^T(i) \Gamma^{\frac{1}{2}}(i-1) \Gamma^{-\frac{1}{2}}(i-1) \mathcal{O}(i-1) \leq \lambda_{\max} \{ \Gamma(i-1) \} \|Z(i)\|^2 V(i-1) \quad (5.143)$$

dok se zamenom (5.143) u (5.139) dobija

$$\begin{aligned} & k_1 k_2 \frac{(\mathcal{O}^T(i-1) Z(i))^2}{\log^k r(i)} Z^T(i) \Gamma(i) Z(i) \leq k_1 k_2 \frac{V(i-1)}{\log^k r(i)} \lambda_{\max} \{ \Gamma(i-1) \} P Z(i) P^T Z^T(i) \Gamma(i) Z(i) \\ & \leq k_1 k_2 \frac{M \log^\delta r(i)}{\lambda_{\min} \{ \Gamma^{-1}(i-1) \} \log^k r(i)} \frac{V(i-1)}{\log^k r(i)} Z^T(i) \Gamma(i) Z(i) = k_3 \frac{V(i-1)}{\log^k r(i-1)} \frac{\log^k r(i-1)}{\lambda_{\min} \{ \Gamma^{-1}(i-1) \}} \frac{Z^T(i) \Gamma(i) Z(i)}{\log^c r(i)} \end{aligned} \quad (5.144)$$

gde je $k_3 = k_1 k_2 M$ i $c = k - \delta > 1$. Imajući u vidu (5.144), relacija (5.138) može se zapisati u obliku

$$E \left\{ \frac{V(i)}{\log^k r(i)} \middle| F_{i-1} \right\} \leq \frac{V(i-1)}{\log^k r(i-1)} \left[1 + k_3 \frac{\log^k r(i-1)}{\lambda_{\min} \{ \Gamma^{-1}(i-1) \}} \frac{Z^T(i) \Gamma(i) Z(i)}{\log^c r(i)} \right] - 2 \frac{\theta^{\mathcal{G}}(i-1) Z(i)}{\log^k r(i)} \left[\phi_1(\theta^{\mathcal{G}}(i-1) Z(i)) - \frac{\alpha}{2} \theta^{\mathcal{G}}(i-1) Z(i) \right] + k_1 \frac{Z^T(i) \Gamma(i) Z(i)}{\log^k r(i)} \quad (5.145)$$

Uvedimo dalje oznake

$$\begin{aligned} z_{i-1} &= V(i-1) / \log^k r(i-1); \quad \beta_{i-1} = k_3 \frac{\log^k r(i-1)}{\lambda_{\min} \{ \Gamma^{-1}(i-1) \}} \frac{Z^T(i) \Gamma(i) Z(i)}{\log^c r(i)}, \\ \xi_{i-1} &= k_1 \frac{Z^T(i) \Gamma(i) Z(i)}{\log^k r(i)}; \\ \zeta_{i-1} &= 2 \frac{\theta^{\mathcal{G}}(i-1) Z(i)}{\log^k r(i)} \left[\phi_1(\theta^{\mathcal{G}}(i-1) Z(i)) - \frac{\alpha}{2} \theta^{\mathcal{G}}(i-1) Z(i) \right] \end{aligned} \quad (5.146)$$

Da bismo dalje primenili lemu 5.2 o konvergenciji martingala potrebno je da se dokaže da su $z_{i-1}, \beta_{i-1}, \xi_{i-1}$ i ζ_{i-1} nenegativne i F_{i-1} -merljive slučajne varijable, koje ispunjavaju zahteve $\sum_{i=1}^{\infty} \beta_i < \infty$; $\sum_{i=1}^{\infty} \xi_i < \infty$. Polazeći od (5.137), zaključuje se da je $r(i)$ niz neopadajućih realnih brojeva, t.j. $r(i) \geq r(i-1)$. Sa druge strane, uzimajući u obzir hipotezu A8 i činjenicu da je

$$r(i) = \text{Tr} \{ \Gamma^{-1}(i) \} = \sum_{i=1}^{n+m} \lambda_i \{ \Gamma^{-1}(i) \} \geq \lambda_{\min} \{ \Gamma^{-1}(i) \} \quad (5.147)$$

gde su sa λ_i označene nenegativne sopstvene vrednosti pozitivno-semidefinitne matrice $\Gamma^{-1}(\cdot)$, dok $\lambda_{\min}$ predstavlja minimalnu sopstvenu vrednost, zaključuje se da je $\lim_{i \rightarrow \infty} r(i) = \infty$, odnosno da je za dovoljno veliko i u važnosti $r(i) > 1$ i $\log r(i) > 0$. Uz navedeno, iz hipoteze A6 i činjenica da su kvadratne forme $V(i) = \theta^{\mathcal{G}}(i) \Gamma^{-1}(i) \theta^{\mathcal{G}}(i)$ i $Z^T(i) \Gamma(i) Z(i)$ nenegativne skalarne veličine sledi da su $z_{i-1}, \beta_{i-1}, \xi_{i-1}$ i ζ_{i-1} zaista nenegativne i F_{i-1} -merljive slučajne varijable. Na taj način, još ostaje da se pokaže da su ispunjeni uslovi $\sum_{i=1}^{\infty} \beta_i < \infty$; $\sum_{i=1}^{\infty} \xi_i < \infty$, što je ekvivalentno zahtevu

$$\sum_{i=1}^{\infty} \frac{Z^T(i) \Gamma(i) Z(i)}{\log^c r(i)} < \infty \quad \text{s.v.1} \quad (5.148)$$

za neko $c > 1$. U cilju dokazivanja relacije (5.148), definišimo sledeću matricu u blok formi

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} \alpha^{-1} & Z^T(i) \\ Z(i) & \Gamma^{-1}(i) \end{bmatrix} \quad (5.149)$$

te primenimo Schur-ovu formulu da bismo izrazili determinantu ove matrice preko determinantni matrica koje predstavljaju njene blok-elemente [60,86]

$$\det A = \det A_{11} \det (A_{22} - A_{21} A_{11}^{-1} A_{12}) = \det A_{22} \det (A_{11} - A_{12} A_{11}^{-1} A_{21}) \quad (5.150)$$

Zamenom (5.149) u (5.150) dalje sledi

$$\alpha^{-1} \det (\Gamma^{-1}(i) - \alpha Z(i) Z^T(i)) = \det \Gamma^{-1}(i) (\alpha^{-1} - Z^T(i) \Gamma(i) Z(i)) \quad (5.151)$$

Takodje, zamenjujući relaciju (5.91), odnosno njenu alternativnu formu

$$\Gamma^{-1}(i-1) = \Gamma^{-1}(i) - \alpha Z(i) Z^T(i) \quad (5.152)$$

u izraz (5.151), dobija se

$$\det \Gamma^{-1}(i-1) = \det \Gamma^{-1}(i) (1 - \alpha Z^T(i) \Gamma(i) Z(i)) \quad (5.153)$$

odakle sledi

$$\alpha Z^T(i) \Gamma(i) Z(i) = \frac{\det \Gamma^{-1}(i) - \det \Gamma^{-1}(i-1)}{\det \Gamma^{-1}(i)} \quad (5.154)$$

Takodje, pošto se može pisati

$$\det \Gamma^{-1}(i) = \prod_{j=1}^{n+m} \lambda_j \{ \Gamma^{-1}(i) \} \leq \lambda_{\max}^{n+m} \{ \Gamma^{-1}(i) \} \quad (5.155)$$

i

$$r(i) = \text{Tr} \{ \Gamma^{-1}(i) \} = \sum_{j=1}^{n+m} \lambda_j \{ \Gamma^{-1}(i) \} \geq \lambda_{\max} \{ \Gamma^{-1}(i) \} \quad (5.156)$$

na osnovu (5.155) i (5.156) se zaključuje

$$r(i) \geq \left[\det \Gamma^{-1}(i) \right]^{\frac{1}{n+m}}, \log^c r(i) \geq \frac{\log^c (\det \Gamma^{-1}(i))}{(n+m)^c} \quad (5.157)$$

Na taj način, polazeći od hipoteze A5, relacije (5.154) i druge relacije u (5.157), dobija se

$$\frac{1}{\alpha} \sum_{i=i_0}^{\infty} \frac{\alpha Z^T(i) \Gamma(i) Z(i)}{\log^c r(i)} \leq \frac{(n+m)^c}{\alpha} \sum_{i=i_0}^{\infty} \frac{\det \Gamma^{-1}(i) - \det \Gamma^{-1}(i-1)}{\det \Gamma^{-1}(i) \log^c (\det \Gamma^{-1}(i))} \quad (5.158)$$

Izvedeni izraz (5.158) može se posmatrati kao vremenski diskretna aproksimacija odgovarajućeg vremenski kontinualnog izraza, koji se dobija kada se u posmatranoj relaciji suma zameni sa integralom, t.j.

$$\sum_{i=i_0}^{\infty} \frac{\det \Gamma^{-1}(i) - \det \Gamma^{-1}(i-1)}{\det \Gamma^{-1}(i) \log^c (\det \Gamma^{-1}(i))} \leq \int_{\det \Gamma^{-1}(i_0)}^{\det \Gamma^{-1}(\infty)} \frac{dt}{t \log^c t} = \frac{1}{(c-1) \log^{c-1} (\det \Gamma^{-1}(i_0))} < \infty \text{ s.v.1} \quad (5.159)$$

Ovde je iskorišćena činjenica da je $r(i) = \text{Tr} \{ \Gamma^{-1}(i) \}$ niz neopadajućih realnih brojeva, koji na osnovu relacija (5.137), (5.147) i hipoteze A8 zadovoljava uslov

$\lim_{i \rightarrow \infty} r(i) = \infty$ s.v.1, pri čemu i sekvenca $\det \Gamma^{-1}(i)$ poseduje iste osobine ,na osnovu relacije (5.157). Na taj način, polazeći od nekog indeksa i_0 , može se pisati $\det \Gamma^{-1}(i) > 0$ za $i \geq i_0$ i $\det \Gamma^{-1}(\infty) = \infty$ s.v.1. Konačno, polazeći od (5.148), (5.158), (5.519) i hipoteze A5, zaključuje se da su pretpostavke leme 2 zadovoljene, te se rezultati ove leme mogu primeniti na izraz (5.146), na osnovu čega se dobija da je $\lim_{i \rightarrow \infty} z_i = z^*$ s.v.1, gde je z^* konačna i nenegativna slučajna varijabla. Sa druge strane, polazeći od definicionog izraza za z_i u (5.146) i primanjujući drugu relaciju u (5.146) sa $b = Z(i), A = \Gamma^{-1}(i)$, dobija se

$$z_i \geq \frac{\lambda_{\min} \{ \Gamma^{-1}(i) \} P^{\theta}(i) P^2}{\log^k r(i)} \quad (5.160)$$

Uzimajući u obzir da z_i konvergira ka konačnoj varijabli, kao i hipotezu A8, iz (5.160) se zaključuje da je $\lim_{i \rightarrow \infty} P^{\theta}(i) P = 0$ s.v.1, čime je dokaz završen.

Uz navedeno, na osnovu (5.160) se takodje zaključuje da je norma greške estimacije

$$P^{\theta}(i) P = O \left\{ \frac{\log^k r(i)}{\lambda_{\min} (\Gamma^{-1}(i))} \right\}^{\frac{1}{2}} \quad (5.161)$$

gde je $\lim_{|x| \rightarrow \infty} \frac{O(|x|)}{|x|} = 0$. Očigledno je da ovaj izraz zavisi od parametra α , pošto $r(i)$ u

(5.147) zavisi od α . Na taj način, α implicitno utiče na brzinu konvergencije algoritma i naziva se faktorom konvergencije, mada je izuzetno složeno da se izvede eksplicitni izraz za ovu zavisnost. Više reči o brzini konvergencije biće u petom poglavlju, gde će biti analizirana asimptotska varijansa totalne greške estimacije.

Teorema 5.2 je opštija od teoreme 5.1, sa obzirom da za razliku od teoreme 5.1 ne uključuje hipotezu A1 o stabilnosti sistema, tako da se teorema 5.2 može primeniti kako na stabilne tako i nestabilne linearne diskretne vremenski-invarijantne dinamičke sisteme. Dalje uprošćavanje postiže se izborom klase ograničenih nelinearnih transformacija u hipotezi A4. U navedenom slučaju hipoteza A7b može se izostaviti, tako da se zahtev za konačnom varijansom poremećaja u hipotezi A2 može oslabiti, tako da je u A2 varijansa $\sigma^2 \leq \infty$. Sada se teorema 5.2 može preformulisati na sledeći način.

Teorema 5.3: Posmatrajmo model sistema (5.80)-(5.84) i algoritam za estimaciju nepoznatih parametara modela (5.85)-(5.87). Ukoliko model i algoritam estimacije zadovoljavaju pretpostavke A2, pri čemu varijansa poremećaja u opštem slučaju ne mora biti konačna, A3, A4 uz uslov da je nelinearna funkcija ograničena, t.j. konstanta konstanta $k_2' = 0$, A5, A6 i A8b, tada vektor procenjenih parametara $\hat{\Theta}(i)$ konvergira skoro sigurno (sa verovatnoćom 1 ili skraćeno s.v.1) ka vektoru tačnih parametara Θ , odnosno

$$P \left\{ \lim_{i \rightarrow \infty} \hat{\Theta}(i) = \Theta \right\} = 1 \text{ s.v.1.}$$

Dokaz: Dokaz je na početku identičan dokazu teorema 5.1 i 5.2, tako da relacije (5.132)-(5.135) ostaju u važnosti. Naime, ako se sa $\tilde{\Theta}(i) = \hat{\Theta}(i) - \Theta$ označi greška estimacije i uvede stohastička funkcija Ljapunova $V(i) = \tilde{\Theta}^T(i) \Gamma^{-1}(i) \tilde{\Theta}(i)$, na osnovu (5.85) se dobija (5.132), koju ćemo ponovo zapisati zbog kompletnosti dokaza, t.j.

$$V(i) = \tilde{\Theta}^T(i-1) \Gamma^{-1}(i) \tilde{\Theta}(i-1) + 2\tilde{\Theta}^T(i-1) Z(i) \psi\left(v(i, \hat{\Theta}(i-1))\right) + Z^T(i) \Gamma(i) Z(i) \psi^2\left(v(i, \hat{\Theta}(i-1))\right) \quad (5.162)$$

Analogno sa dokazom teoreme 1, ako se desnoj strani relacije (5.132) doda i oduzme $V(i-1)$ i uvrsti izraz (5.91) za inverznu matricu pojačanja estimacione procedure, dobija se (5.133), odnosno

$$V(i) = V(i-1) + 2\tilde{\Theta}^T(i-1) Z(i) \psi\left(v(i, \hat{\Theta}(i-1))\right) + \alpha \left(\tilde{\Theta}^T(i-1) Z(i)\right)^2 + Z^T(i) \Gamma(i) Z(i) \psi^2\left(v(i, \hat{\Theta}(i-1))\right) \quad (5.163)$$

Po analogiji sa teoremama 5.1 i 5.2, uvedimo kao u (5.134) funkciju

$$\phi_2\left(\tilde{\Theta}^T(i-1) Z(i)\right) = E\left\{\psi^2\left(-v(i, \hat{\Theta}(i-1))\right) \middle| F_{i-1}\right\} \quad (5.164)$$

gde je rezidual merenja dat relacijom (5.135), odnosno

$$v(i, \hat{\Theta}(i-1)) = -\tilde{\Theta}^T(i-1) Z(i) + \xi(i). \quad (5.165)$$

Uzimajući, dalje, u obzir hipotezu A4, uz dodatni uslov da je konstanta $k_2' = 0$, dobija se

$$\phi_2\left(\tilde{\Theta}^T(i-1) Z(i)\right) \leq k_1 \quad (5.166)$$

gde je k_1 pozitivna i konačna konstanta. Imajući u vidu da je za proizvoljan vector kolonu b u važnosti $Tr\{bb^T\} = b^T b$, ako se na izraz (5.91) primeni operacija traga matrice i usvoji $b = Z(i)$, dobija se (5.137), t.j.

$$r(i) = r(i-1) + \alpha Z^T(i) Z(i) \quad (5.167)$$

Slično kao u teoremi 5.2, ako se sada (4,163) podeli sa $\log^k r(i)$, gde je $k > 1$ zbog A8b i primeni opertor uslovnog matematičko očekivanja $E\{./F_{i-1}\}$, kao i primene hipoteze A2, A3 i uvrste relacije (5.164)-(5.166), može se pisati

$$E\left\{\frac{V(i)}{\log^k r(i)} \middle| F_{i-1}\right\} \leq \frac{V(i-1)}{\log^k r(i)} - \frac{2\tilde{\Theta}^T(i-1) Z(i)}{\log^k r(i)} \left[\phi_1\left(\tilde{\Theta}^T(i-1) Z(i)\right) - \frac{\alpha}{2} \tilde{\Theta}^T(i-1) Z(i)\right] + k_1 \frac{Z^T(i) \Gamma(i) Z(i)}{\log^k r(i)}. \quad (5.168)$$

Uvedimo sada nenegativnu veličinu

$$S(i) = 2 \sum_{k=1}^i \tilde{\Theta}^T(k-1) Z(k) \left\{ \phi_1\left(\tilde{\Theta}^T(k-1) Z(k)\right) - \frac{\alpha}{2} \tilde{\Theta}^T(k-1) Z(k) \right\} + k_2 \geq 0 \quad (5.169)$$

gde je k_2 neka konačna i pozitivna konstanta. Nejednakost u (5.169) je direktna posledica hipoteze A6. Poslednja relacija može se napisati i u alternativnom obliku

$$S(i) = S(i-1) + 2\tilde{\theta}^T(i-1)Z(i) \left\{ \phi_1(\tilde{\theta}^T(i-1)Z(i)) - \frac{\alpha}{2} \tilde{\theta}^T(i-1)Z(i) \right\} \quad (5.170)$$

Uz navedeno, uvedimo dodatno i veličine

$$R(i) = \frac{V(i)}{\log^k r(i)}, \quad T(i) = R(i) + \frac{S(i)}{\log^k r(i)} \quad (5.171)$$

Na osnovu (5.170) i (5.171) dalje se može pisati

$$T(i) = \frac{V(i)}{\log^k r(i)} + \frac{1}{\log^k r(i)} \left\{ S(i-1) + 2\tilde{\theta}^T(i-1)Z(i) \left[\phi_1(\tilde{\theta}^T(i-1)Z(i)) - \frac{\alpha}{2} \tilde{\theta}^T(i-1)Z(i) \right] \right\} \quad (5.172)$$

Ako se na relaciju (5.172) primeni operator uslovnog matematičkog očekivanja $E\{\cdot / F_{i-1}\}$ dobija se

$$\begin{aligned} E\{T(i) / F_{i-1}\} &= E\left\{ \frac{V(i)}{\log^k r(i)} / F_{i-1} \right\} + \frac{S(i-1)}{\log^k r(i)} \\ &+ \frac{2\tilde{\theta}^T(i-1)Z(i)}{\log^k r(i)} \left\{ \phi_1(\tilde{\theta}^T(i-1)Z(i)) - \frac{\alpha}{2} \tilde{\theta}^T(i-1)Z(i) \right\} \end{aligned} \quad (5.173)$$

Zamenom (5.168) u (5.173), dalje sledi

$$E\{T(i) / F_{i-1}\} \leq \frac{V(i-1)}{\log^k r(i)} + \frac{S(i-1)}{\log^k r(i)} + k_1 \frac{Z^T(i)\Gamma(i)Z(i)}{\log^k r(i)} \quad (5.174)$$

Pošto je na osnovu (5.167) $r(i)$ neopadajući niz pozitivnih realnih brojeva, to je $\log r(i) \geq \log r(i-1)$, odakle se zaključuje da je za $k > 1$

$$\frac{1}{\log^k r(i)} \leq \frac{1}{\log^k r(i-1)} \quad (5.175)$$

Zamenom (5.175) u (5.174), uz korišćenje definicionog izraza (5.171) za $T(i)$, može se pisati

$$\begin{aligned} E\{T(i) / F_{i-1}\} &\leq \frac{V(i-1)}{\log^k r(i-1)} + \frac{S(i-1)}{\log^k r(i-1)} + k_1 \frac{Z^T(i)\Gamma(i)Z(i)}{\log^k r(i)} \\ &= T(i-1) + k_1 \frac{Z^T(i)\Gamma(i)Z(i)}{\log^k r(i)} \end{aligned} \quad (5.176)$$

Da bi se na relaciju (5.176) primenila lema 1, potrebno je da se dokaže da je

$$\sum_{i=1}^{\infty} \frac{Z^T(i)\Gamma(i)Z(i)}{\log^k r(i)} < \infty \text{ s.v.1} \quad (5.177)$$

Relacija (5.177) dokazuje se na isti način kao relacija (5.148) u okviru dokaza teoreme 5.2, a u suštini ova relacija je direktna posledica pretpostavki A5 i A8b. Na taj način, može se direktno primeniti lema 1 na relaciju (5.176), na osnovu čega se zaključuje

$$\lim_{i \rightarrow \infty} T(i) = T \text{ s.v.1} \quad (5.178)$$

pa na osnovu (5.172) dalje sledi

$$\lim_{i \rightarrow \infty} R(i) = R \text{ s.v.1} \quad (5.179)$$

gde su T i R konačne i nenegativne slučajne veličine. Sa druge strane, za proizvoljnu kvadratnu matricu A i proizvoljan vektor kolonu b može se pisati $Tr\{Abb^T\} = b^T Ab$, te se izborom $A = \Gamma^{-1}(i)$ i $b = \tilde{\theta}(i)$ i primenom relacije (5.106) iz definicionog izraza (5.171) takodje dobija

$$\begin{aligned} R(i) &= \frac{V(i)}{\log^k r(i)} = \frac{\tilde{\theta}^T(i) \Gamma^{-1}(i) \tilde{\theta}(i)}{\log^k r(i)} \\ &= \frac{Tr\{\Gamma^{-1}(i) \tilde{\theta}(i) \tilde{\theta}^T(i)\}}{\log^k r(i)} \leq \frac{\lambda_{\min}\{\Gamma^{-1}(i)\} \|\tilde{\theta}(i)\|^2}{\log^k r(i)} \end{aligned} \quad (5.180)$$

Na osnovu prve relacije u hipotezi A8b i činjenice da je $\lambda_{\min}\{\cdot\} > 0$, pošto je $\Gamma^{-1}(\cdot) > 0$, zaključuje se iz (5.180) da mora biti ispunjeno

$$\lim_{i \rightarrow \infty} \|\tilde{\theta}(i)\|^2 = 0 \text{ s.v.1} \quad (5.181)$$

čime je dokaz teoreme završen.

Primetimo da se teoreme 5.1-5.3 mogu primeniti i na FIR model sistema, koji je po definiciji stabilan i poseduje sve polove u koordinatnom početku kompleksne ravni. U tom smislu, iz pretpostavke A1 treba isključiti uslov o stabilnosti karakterističnog polinoma $A(\cdot)$ (za FIR sisteme polinom $A(\cdot)=1$). Na taj način, iz dokaza teoreme 5.1 treba isključiti relacije (5.103)-(5.104), a relacija (5.101) je direktna posledica ograničenosti pobude na osnovu hipoteze A1 i slučajnog poremećaja na osnovu A2, pri čemu je sada regresioni vektor dat sa

$$Z^T(i) = [u(i-1), \dots, u(i-m)] \quad (5.182)$$

Dokaz i pretpostavke teorema 5.2 i 5.3 ostaju neizmenjeni, pošto ove teoreme ne zahtevaju pretpostavku A1, pri čemu je regresioni vektor umesto relacijom (5.84) definisan izrazom (5.182). U narednom poglavlju biće pokazano da se izvedene teoreme mogu generalizovati i primeniti i na sisteme sa korelisanim ili obojenim slučajnim poremećajem, odnosno i na ARMAX modele sistema.

5.3.2 Analiza konvergencije nelinearnih rekurzivnih stohastičkih algoritama za estimaciju parametara u dinamičkim modelima sistema tipa nelinearne regresije (ARMAX modelima) primenom teorije stabilnosti običnih diferencijalnih jednačina

U ovom poglavlju razmatrana je klasa modela sistema koji su opisani diferencnom jednačinom sa fiksnim koeficijentima

$$y(i) = -\sum_{k=1}^n a_k y(i-k) + \sum_{k=1}^m b_k u(i-k) + \sum_{k=1}^l \xi(i-k) + \xi(i) \quad (5.183)$$

gde su $y(i) \in R^1$, $u(i) \in R^1$, $\xi(i) \in R^1$ izlaz sistema, merljivi pobudni signal ili ulaz i nemerljivi stohastički ulaz ili poremećaj, respektivno, dok su $a_i, i=1, \dots, n$, $b_j, j=1, \dots, m$ i $c_k, k=1, \dots, l$ konstantni parametri u modelu. Dalje se pretpostavlja da je diskretan niz $\{\xi(i)\}$ stohastički proces, koji generiše rastući niz sub-sigma algebri $\{F_i\}$, kao i da su poznate konstante m, n i l . Takodje, funkcija gustine verovatnoće (fgv) slučajne varijable $\xi(i)$ nije u opštem slučaju poznata, ali je uobičajeno da se na osnovu izvesnog apriornog znanja o statističkim karakteristikama poremećaja definiše određena klasa funkcija kojoj će i realna fgv poremećaja pripadati. O ovom problemu biće više reči u petom poglavlju.

Ako se dalje uvede operator kašnjenja $q^{-k}y(i) = y(i-k)$, model (5.181) može se zapisati u polinomijalnom obliku

$$A(q^{-1})y(i) = B(q^{-1})u(i) + C(q^{-1})\xi(i), \quad (5.184)$$

gde se polinomi

$$A(q^{-1}) = 1 + \sum_{k=1}^n a_k q^{-k}, \quad B(q^{-1}) = \sum_{k=1}^m b_k q^{-k}, \quad C(q^{-1}) = 1 + \sum_{k=1}^l c_k q^{-k}, \quad (5.185)$$

nazivaju karakterističnim polinomom, kontrolnim polinomom i polinomom poremećaja, respektivno. Model (5.80) može se takodje zapisati i u formi linearne regresije

$$y(i) = Z^T(i, \theta) \Theta + \xi(i) \quad (5.186)$$

gde je pseudo-linearni regresioni vektor

$$Z^T(i, \theta) = [-y(i-1), \dots, -y(i-n), u(i-1), \dots, u(i-m), v(i-1, \theta), \dots, v(i-l, \theta)] \quad (5.187)$$

dok je $\Theta^T = [a_1 \dots a_n b_1 \dots b_m c_1 \dots c_l]$ vektor konstantnih parametara čije vrednosti treba proceniti na osnovu raspoložive merne sekvence ulaznih $u(\cdot)$ i izlaznih $y(\cdot)$ signala. U prethodnom poglavlju je razmatrana klasa rekurzivnih stohastičkih algoritama za estimaciju nepoznatog vektora parametara u (5.186) u obliku

$$\hat{\Theta}(i) = \hat{\Theta}(i-1) + \Gamma(i) Z(i) \psi(v(i, \hat{\Theta}(i-1))) \quad (5.188)$$

gde je rezidual merenja ili is greška predikcije

$$v(i, \hat{\Theta}(i-1)) = y(i) - \hat{\Theta}^T(i-1) Z(i) \quad (5.189)$$

pri čemu je

$$Z^T(i) = [-y(i-1), \dots, -y(i-n), u(i-1), \dots, u(i-m), v(i-1, \hat{\Theta}(i-2)), \dots, v(i-l, \hat{\Theta}(i-l-1))] \quad (5.190)$$

pogodna rekurzivna aproksimacija pseudo-linearnog regresionog vektora u (5.187), dok je težinska matrica algoritma određena rekurzivnom relacijom

$$\Gamma(i) = \Gamma(i-1) - \Gamma(i-1) Z(i) [\alpha^{-1} + Z^T(i) \Gamma(i-1) Z(i)]^{-1} Z^T(i) \Gamma(i) \quad (5.191)$$

pri čemu je njena inicijalna vrednost $\Gamma(0) = \rho^{-1} I$, $\rho > 0$. Po analogiji sa odgovarajućim algoritmima za estimaciju parametara u modelima linearne regresije, koji su razmatrani u prethodnom odeljku, koeficijent α predstavlja slobodan parametar čijim se izborom može uticati na konvergenijska svojstva algoritma. Takođe, pogodnim izborom ovog parametra može se uspostaviti veza između nelinearne transformacije reziduala $\psi(\cdot)$ i matrice pojačanja algoritma $\Gamma(\cdot)$, kao što je već diskutovano u prethodnom poglavlju.

Konvergencija algoritam (5.188)-(5.191) za estimaciju nepoznatih parametara u modelu pseudolinearne regresije (5.183)-(5.187) može se dokazati sledeći metodologiju analize konvergenije algoritma (5.85)-(5.87) za estimaciju parametara u modelima linearne regresije (5.80)-(5.84). Ovakva analiza zahteva da se uvede sledeće dodatne pretpostavke.

H1: Sve nule polinoma poremećaja $C(q^{-1})$ u ARMAX modelu (5.184), (5.185) leže unutar jediničnog kruga, sa centrom u koordinatnom početku kompleksne ravni.

H6: Postoji uslovno matematičko očekivanje $\phi_1(z) = E\{\psi(-\xi(i) + z) | F_{i-1}\}$ koje

zadovoljava uslov $z \left[\phi_1\left(\frac{z}{C(q^{-1})}\right) - \frac{1}{2} \alpha z \right] > 0$ za svako $z \neq 0$, gde je koeficijent α

definisan u pretpostavci A5, a nelinearna $\psi(\cdot)$ funkcija zadovoljava pretpostavke A3 i A5.

Hipoteza H1 je uobičajena u analizi sistema sa koreliranim ili obojenim ekvivalentnim aditivnim šumom merenja na izlazu sistema, dok hipoteza H6 predstavlja pogodnu modifikaciju ranije uvedene pretpostavke A6. Međutim, obe pomenute hipoteze ispunjene su pod istim uslovima. Naime, obe pomenute hipoteze A6 i H6 biće zadovoljene ukoliko je uvedena funkcija

$$\phi_1(a) = \int_0^\infty [\psi(u+a) - \psi(u-a)] dP(u)$$

monotona, neopadajuća realna funkcija realnog argumenta. Ukoliko su pretpostavke A2 i A3 zadovoljene, ovaj zahtev će biti ispunjen ako nelinearna $\psi(\cdot)$ funkcija i funkcija raspodele verovatnoće $P(\cdot)$ poremećaja $\xi(\cdot)$ u modelima (5.80)-(5.84) i (5.183)-(5.187) imaju zajedničku tačku rasta, što se svodi na zadovoljenje uslova

$\psi(z+\varepsilon) > \psi(z-\varepsilon)$, $P(z+\varepsilon) > P(z-\varepsilon)$ za neko z i svako $\varepsilon > 0$, na osnovu koga dalje sledi $a\phi_1(a) > 0$ za $a \neq 0$ i $\phi_1(0) = 0$.

Na taj način, pretpostavke A6 i H6 će biti ispunjena ako je nelinearna $\psi(\cdot)$ funkcija neparna, skoro svuda neprekidna, monotono rastuća i u segmentima neprekidno diferencijabilna. Kao što je ranije istaknuto, iz praktičnih razloga je poželjno da ova funkcija bude i ograničena ($k_2 = 0$ u A4).

Posebno ako je poremećaj raspodeljen po Gausovom zakonu verovatnoće, okarakterisanom sa funkcijom raspodele verovatnoće $P(\cdot)$ koja ima nultu srednju vrednost i varijansu σ^2 , a za nelinearnu transformaciju reziduala se usvoji funkcija verodostojnosti

$$\psi(z) = -[\log p(z)]' = z/\sigma^2, \quad p(z) = dP(z)/dz$$

i usvoji $\alpha = E\{\psi'(\xi(i))\} = \sigma^{-2}$, algoritam (5.183)-(5.187) se svodi na rekurzivni algoritam najmanjih kvadrata. Dalje se dobija da je

$$a\phi_1(a) = a^2 \left[\frac{1}{C(q^{-1})} - \frac{1}{2} \right],$$

što znači da je ulov A6 $a\phi_1(a) > 0$ za $a \neq 0$ i $\phi_1(0) = 0$, ispunjen ukoliko je zadovoljen uslov pozitivne realnosti

$$\operatorname{Re} \left\{ \frac{1}{C(q^{-1})} - \frac{1}{2} \right\} > 0$$

gde $\operatorname{Re}\{\cdot\}$ označava realni deo odgovarajućeg kompleksnog broja. Dobijeni uslov je dobro poznat u literaturi posvećenoj konvergenciji linearnih rekurzivnih stohastičkih procedura za estimaciju parametara u modelima tipa pseudo-linearne regresije, te hipoteza H6 predstavlja generalizaciju uslova pozitivne realnosti na nelinearne rekurzivne stohastičke procedure za estimaciju parametara u usvojenoj klasi modela sistema [28,63].

Sada se mogu formulisati sledeće teoreme, koje predstavljaju generalizaciju teorema 5.2 i 5.3.

Teorema 5.4: Posmatrajmo model sistema (5.183)-(5.187) i algoritam za estimaciju nepoznatih parametara modela (5.188)-(5.191). Ukoliko model i algoritam estimacije zadovoljavaju pretpostavke H1, A2-A5, H6, A7b i A8b, tada vektor

procenjenih parametara $\hat{\Theta}(i)$ konvergira skoro sigurno (sa verovatnoćom 1 ili skraćeno s.v.1) ka vektoru tačnih parametara Θ , odnosno

$$P\left\{\lim_{i \rightarrow \infty} \hat{\Theta}(i) = \Theta\right\} = 1 \text{ s.v.1.}$$

Dokaz: Dozaz u potpunosti sledi postupak koji je izložen prilikom dokazivanja teoreme 5.2, sa tom razlikom što funkcije

$$\phi_1\left(\tilde{\theta}^T(i-1)Z(i)\right), \phi_2\left(\tilde{\theta}^T(i-1)Z(i)\right)$$

treba zameniti sa

$$\phi_1\left(\tilde{\theta}^T(i-1)Z(i)/C(q^{-1})\right), \phi_2\left(\tilde{\theta}^T(i-1)Z(i)/C(q^{-1})\right)$$

dok izraz (5.135) za rezidual merenja treba modifikovati na sledeći način

$$v(i, \hat{\Theta}(i-1)) = \frac{-\theta^T(i-1)Z(i)}{C(q^{-1})} + \xi(i)$$

Takodje, u relacijam (5.155)-(5.158) dimenziju vektora parametara $(n+m)$ u ARX modelu treba zameniti sa promenjenom dimenzijom vektora parametara $(n+m+l)$ u ARMAX modelu. Tada sve jednačine u dokazu teoreme 5.2 ostaju nepromenjene.

Teorema 5.5: Posmatrajmo model sistema (5.183)-(5.187) i algoritam za estimaciju nepoznatih parametara modela (5.188)-(5.191). Ukoliko model i algoritam estimacije zadovoljavaju pretpostavke H1, A2, pri čemu varijansa poremećaja u opštem slučaju ne mora biti konačna, A3, A4 uz uslov da je nelinearna funkcija ograničena, t.j. konstanta konstanta $k_2' = 0$, A5, H6 i A8b, tada vektor procenjenih parametara $\hat{\Theta}(i)$ konvergira skoro sigurno (sa verovatnoćom 1 ili skraćeno s.v.1) ka vektoru tačnih parametara Θ , odnosno

$$P\left\{\lim_{i \rightarrow \infty} \hat{\Theta}(i) = \Theta\right\} = 1 \text{ s.v.1.}$$

Dokaz: Uvažavajući izmene koje su navedene u okviru dokaza teoreme 5.4, dokaz teoreme 5.5 je identičan dokazu teoreme 5.3.

Na sličan način može se definisati i pandan teoremi 5.1, ali je to ovde izostavljeno, s obzirom da su teoreme 5.2 i 5.3 opštije od teoreme 5.1. Ostalo je još na kraju ovog poglavlja da analiziramo konvergenciju rekurzivnih stohastičkih algoritama parametarske estimacije zasnovanih na filtraciji podataka. Kao što je pokazano u trećem poglavlju, opšta forma ovakvog algoritma definisana je relacijama

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \Gamma(i)Z_f(i)\psi(v(i)) \quad (5.192)$$

$$v(i) = y(i) - Z^T(i)\hat{\theta}(i-1) \quad (5.193)$$

$$Z^T(i) = \{-y(i-1) \dots -y(i-n)u(i-1) \dots u(i-n)v(i-1) \dots v(i-n)\} \quad (5.194)$$

$$Z_f^T(i) = \{-\tilde{y}(i-1) \dots -\tilde{y}(i-n)\tilde{u}(i-1) \dots \tilde{u}(i-n)\tilde{v}(i-1) \dots \tilde{v}(i-n)\} \quad (5.195)$$

Filtrirane vrednosti sinala u (5.195) dobijaju se propuštanjem ovih signala kroz inverzni $C(\cdot)$ polinom, odnosno

$$Z_f(i) = \frac{1}{C^{i-1}(q^{-1})} Z(i) = D^{i-1}(q^{-1}) \quad (5.196)$$

U poslednjoj relaciji gornji indeks $(i-1)$ označava da je $C(\cdot)$ polinom definisan relacijom (5.185), samo su nepoznati koeficijenti zamenjeni njihovim procenama u $(i-1)$ -vom koraku. Kao što je pokazano u trećem poglavlju, veza između koeficijenata $C(\cdot)$ polinoma i $D(\cdot)$ polinoma data je sa

$$d_k = -\sum_{j=1}^k c_j d_{k-j}, \quad k=1, \dots, n; \quad d_0 = 1 \quad (5.197)$$

Relacije (5.196) i (5.197) praktično se primenjuju na sledeći način:

$$\begin{aligned} \tilde{y}(k) &= y(k) + Y^T(k) d^{i-1}; Y^T(k) = \{y(k-1) \dots y(k-1+n)\}; d^{i-1} = \{d_1^{i-1} \dots d_n^{i-1}\}^T \\ \tilde{u}(k) &= u(k) + U^T(k) d^{i-1}; Y^T(k) = \{u(k-1) \dots u(k-1+n)\} \\ \tilde{v}(k) &= v(k) + E^T(k) d^{i-1}; E^T(k) = \{v(k-1) \dots v(k-1+n)\} \end{aligned} \quad (5.198)$$

Težinska ili matrica pojačanja algoritma data je izrazom

$$\Gamma(i) = \Gamma(i-1) - \Gamma(i-1) Z_f(i) Z_f^T(i) \Gamma(i-1) \left[\alpha(i)^{-1} + Z_f^T(i) \Gamma(i-1) Z_f(i) \right]^{-1} \quad (5.199)$$

gde je $\alpha(i)$ parametar koji utiče na konvergenijska svojstva algoritma. U trećem poglavlju je pokazano da ovaj parametar omogućava i da se uspostavi vrlo važna veza između nelinearne transformacije reziduala $\psi(\cdot)$ i matrice pojačanja $\Gamma(\cdot)$, na bazi tehnike linearizacije, što rezultuje u sledeće mogućnosti za izbor ovog parametra: $\alpha(i) = \alpha = E\{\psi'(\xi(i))\}$ ili $\alpha(i) = \psi'(v(i))$ ili $\alpha(i) = \alpha = 1$. Primetimo da poslednji izbor rezultuje u modifikovanu verziju klasičnog algoritma generalisanih najmanjih kvadrata, koja se od osnovnog algoritma razlikuje samo po uvođenju nelinearne transformacije reziduala u jednačinu za estimaciju parametara (5.192), uz dodatno filtriranje podataka, dok je matrica pojačanja data sa Rikatiјevom jednačinom (5.199) za $\alpha(i) = \alpha = 1$ i ostaje neizmenjena.

Analiza konvergenije predloženog algoritma (5.192)-(5.199) može se, kao i do sada, zasnivati na teoriji martingala. Međutim, kod ovakve klase algoritama pogodnije je da se primeni pristup zasnovan na teoriji klasičnih diferencijalnih jednačina. U osnovi ovakvog prilaza nalazi se ideja da se rekurzivnoj stohastičkoj proceduri za estimaciju parametara pridruži odgovarajuća obična diferencijalna jednačina, pa da se dalje za analizu stabilnosti njenog rešenja koristi teorija stabilnosti Ljapunova. Otuda se ovakav pristup označava u literaturi skraćeno sa ODE (engleski Ordinary Differential Equation Approach) [15,63,93]. Postupci za pridruživanje obične diferencijalne jednačine odgovarajućim tipovima rekurzivnih stohastičkih estimacionih procedura dobro su razradjeni u literaturi [45,63]. Da bi se ovi rezultati

primenili na razmatrani algoritam, neophodno je da se jednokoračni prediktor izlaza sistema i njegov gradijent prikažu u pogodnoj formi u prostoru stanja. Da bismo to postigli, takodje je potrebno da se osnovni algoritam (5.192)-(5.199) prikaže u nešto drugačijoj formi, kao što je prikazano u nastavku. Naime, koristeći lemu o inverziji matrica, izraz (5.199) se zapisati u obliku

$$\Gamma^{-1}(i) = \Gamma^{-1}(i-1) + \alpha(i) Z_f(i) Z_f^T(i) \quad (5.200)$$

Uvodeći dalje smenu

$$\Gamma^{-1}(i) = \bar{R}(i) \quad (5.201)$$

izraz (5.200) postaje

$$\bar{R}(i) = \bar{R}(i-1) + \alpha(i) Z_f(i) Z_f^T(i) \quad (5.202)$$

odnosno, usvajajući novu smenu

$$R(k) = \frac{1}{k} \bar{R}(k) \quad (5.203)$$

relacija (5.202) dobija oblik

$$iR(i) = (i-1)R(i-1) + \alpha(i) Z_f(i) Z_f^T(i) \quad (5.204)$$

odakle se dobija

$$R(i) = \frac{i-1}{i} R(i-1) + \frac{1}{i} \alpha(i) Z_f(i) Z_f^T(i) \quad (5.205)$$

ili u alternativnom obliku

$$R(i) = R(i-1) + \frac{1}{i} [\alpha(i) Z_f(i) Z_f^T(i) - R(i-1)] \quad (5.206)$$

Na taj način, procena nepoznatog vektora parametara u ARMAX modelu sistema (5.188)-(5.191) svodi se na

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \frac{1}{i} R^{-1}(i) Z_f(i) \psi(v(i)) \quad (5.207)$$

gde je matrica pojačanja definisana relacijom 5.206), vektor filtriranih podataka izrazima (5.198), a rezidual merenja relacijom (5.193). U trećem poglavlju je pokazano da je predikcija izlaza ARMAX modela definisana izrazom

$$\hat{y}(i/\theta) = \frac{B(q^{-1})}{C(q^{-1})} u(i) + \left[1 - \frac{A(q^{-1})}{C(q^{-1})} \right] y(i) = Z^T(i, \theta) \theta \quad (5.208)$$

gde su $d = n + m + l$ dimenzioni vektori parametara i pseudo-linearne regresije dati sledećim izrazima

$$\begin{aligned} \theta^T &= \{a_{l,L}, a_n, b_{l,L}, b_m, c_{l,L}, c_l\} \\ Z^T(i, \theta) &= \{-y(i-1), L, -y(i-n), u(i-1), L, u(i-m), v(i-1, \theta), v(i-l, \theta)\} \end{aligned} \quad (5.209)$$

Analiza rekurzivnih stohastičkih algoritma parametarske identifikacije: konvergencija
i konzistencija procena parametara i simulacija

pri čemu je greška predikcije

$$v(i, \theta) = y(i) - \hat{y}(i/\theta) = y(i) - Z^T(i, \theta)\theta \quad (5.210)$$

Predikcija izlaza (5.208) može se alternativno dobiti i filtriranjem ulazno-izlaznih mernih podataka kroz konačno dimenzioni linearni filter čiji je vektor stanja

$$\varphi^T(i, \theta) = Z^T(i, \theta) = \{-y(i-1), L, -y(i-n), u(i-1), L, u(i-m), v(i-1, \theta), v(i-l, \theta)\} \quad (5.211)$$

pobudni signal je dvodimenzioni vektor kolona koji sadrži tekuća merenja izlaza i ulaza $z^T(i) = \{y(i)u(i)\}$, dok predikcija (5.208) predstavlja izlazni signal. Na taj način se dobija za jadačinu stanja

$$\varphi(i+1, \theta) = F(\theta)\varphi(i, \theta) + E(\theta)z(i) \quad (5.212)$$

gde su odgovarajuće matrice funkcije nepoznatog vektora parametara θ dimenzije $d=n+m+l$ date izrazima

$$F(\theta) = \begin{bmatrix} 00 & \dots & 00 & 00 & \dots & 00 & 00 & \dots & 00 \\ 10 & \dots & 00 & 00 & \dots & 00 & 00 & \dots & 00 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 00 & \dots & 10 & 00 & \dots & 00 & 00 & \dots & 00 \\ & & & & & & & & \\ 00 & \dots & 00 & 00 & \dots & 00 & 00 & \dots & 00 \\ 00 & \dots & 00 & 10 & \dots & 00 & 00 & \dots & 00 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 00 & \dots & 00 & 00 & \dots & 10 & 00 & \dots & 00 \\ & & & & & & & & \\ -a_1 \dots & -a_n & -b_1 & \dots & -b_m & -c_1 & \dots & -c_l \\ 00 & \dots & 00 & 00 & \dots & 00 & 10 & \dots & 00 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 00 & \dots & 00 & 00 & \dots & 00 & 00 & \dots & 00 \end{bmatrix}; E(\theta) = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ & & & & & & & \\ 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ & & & & & & & \\ 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \quad (5.213)$$

dok je jednačina izlaza data relacijom (5.208), odnosno

$$\hat{y}(i/\theta) = H(\theta)\varphi(i, \theta); H(\theta) = \theta^T \quad (5.214)$$

Sledeći metodologiju izloženu u literaturi [63], diferencirajmo izraze (5.212) i (5.214) po vektoru parametara θ . Pošto je $F(\theta)$ matrica a θ je vektor, to će izvod $F(\theta)$ po θ biti veličina sa tri indeksa, odnosno tenzor [2,5]. Ovako kompleksna notacija može se izbeći uvođenjem $d \times d$ matrice

$$M(\theta, \varphi, z) = \frac{\partial}{\partial \theta} [F(\theta)\varphi + E(\theta)z] \quad (5.215)$$

čija i-ta kolona ima d vrsta i data je izrazom

$$\left[\frac{d}{d\theta_i} F(\theta) \right] \varphi + \left[\frac{d}{d\theta_i} E(\theta) \right] z; i = 1, \dots, d; d = n + m + l \quad (5.216)$$

kao i $1 \times d$ vector vrstu

$$D(\theta, \varphi) = \frac{\partial}{\partial \theta} [H(\theta) \varphi] \quad (5.217)$$

čija i-ta kolona data je relacijom

$$\left[\frac{d}{d\theta_i} H(\theta) \right] \varphi; i = 1, \dots, d; d = n + m + l \quad (5.218)$$

Uz navedeno, uvedimo dodatno i $d \times d$ matricu

$$\zeta(i, \theta) = \frac{d}{d\theta} \varphi(i, \theta) = \{ \varphi^{(1)}(i, \theta) \cdots \varphi^{(d)}(i, \theta) \}; d = n + m + l \quad (5.219)$$

čija je j-ta kolona sa d vrsta i ima oblik

$$\varphi^{(j)}(i, \theta) = \frac{d\varphi(i, \theta)}{d\theta_j}; j = 1, \dots, d = n + m + l \quad (5.220)$$

Na taj način, izvod izraza (5.212) po θ može se pisati u obliku

$$\zeta(i+1, \theta) = F(\theta) \zeta(i, \theta) + M(\theta, \varphi(i, \theta), z(i)) \quad (5.221)$$

Takodje, uvodeći oznaku za d dimenzioni vektor kolonu

$$\eta(i, \theta) = \frac{d}{d\theta} \hat{y}(i/\theta) \quad (5.222)$$

jednačina (5.214) može se napisati u sledećem obliku, nakon diferenciranja

$$\eta^T(i, \theta) = H(\theta) \zeta(i, \theta) + D(\theta, \varphi(i, \theta)) \quad (5.223)$$

gde je vector $H(\cdot)$ definisan relacijom (5.214) a matrice $\zeta(\cdot)$ i $D(\cdot)$ izrazima (5.219, 5.220) i (5.217, 5.218), respektivno. Uvodeći prošireni vector stanja $q(\cdot)$ i prošireni vector izlaza $\gamma(\cdot)$

$$q(i, \theta) = \begin{bmatrix} \varphi(i, \theta)_{d \times 1} \\ \text{col } \zeta(i, \theta)_{d^2 \times 1} \end{bmatrix}_{d(d+1) \times 1}; \gamma(i, \theta) = \begin{bmatrix} \hat{y}(i/\theta)_{1 \times 1} \\ \eta(i, \theta)_{d \times 1} \end{bmatrix}_{(d+1) \times 1}; \text{col } \zeta(i, \theta) = \begin{Bmatrix} \varphi^{(1)}(i, \theta)_{d \times 1} \\ \vdots \\ \varphi^{(d)}(i, \theta)_{d \times 1} \end{Bmatrix}_{d^2 \times 1} \quad (5.224)$$

gde $\text{col } A$ označava vektor kolonu koji je formiran od matrice A tako što su njene kolone, počevši od prve, poredjane jedna za drugom, relacije (5.212), (5.221), (5.214) i (5.223) mogu se objediniti u okviru sledeće dve vektorske jednačine, koje definišu prošireni model u prostoru stanja

$$\begin{aligned} q(i+1, \theta) &= A(\theta)q(i, \theta) + B(\theta)z(i) \\ \gamma(i, \theta) &= C(\theta)q(i, \theta) \end{aligned} \quad (5.225)$$

gde su $d(d+1) \times d(d+1)$ matrica A i $d(d+1) \times 2$ matrica B date izrazima

$$A(\theta) = \begin{bmatrix} F(\theta) & 0_{d \times d} & \cdots & 0_{d \times d} \\ G_1(\theta) & F(\theta) & \cdots & 0_{d \times d} \\ G_2(\theta) & 0_{d \times d} & \cdots & 0_{d \times d} \\ \vdots & \vdots & \ddots & \vdots \\ G_d(\theta) & 0_{d \times d} & \vdots & F(\theta) \end{bmatrix}_{d(d+1) \times d(d+1)} \quad ; \quad B(\theta) = \begin{bmatrix} E(\theta) \\ 0_{d \times 2} \\ \vdots \\ 0_{d \times 2} \end{bmatrix}_{d(d+1) \times 2} \quad (5.226)$$

pri čemu su matrice F i E definisane izrazima (5.213), dok $0_{m \times n}$ označava nula matricu dimenzije $m \times n$. Takodje, na osnovu (5.213) i (5.216) dobija se

$$\frac{\partial}{\partial \theta_j} E(\theta) = 0; \quad j = 1, \dots, d \quad (5.227)$$

pa iz relacija (5.215) i (5.216) sledi

$$M(\varphi, z, \theta) = \frac{\partial}{\partial \theta} F(\theta) \varphi = \begin{bmatrix} 0_{n \times d} \\ 0_{m \times d} \\ -\varphi^T_{1 \times d} \\ 0_{(l-1)d} \end{bmatrix} = [M_1 \cdots M_d]_{d \times d} \quad (5.228)$$

$$M_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \varphi_i = \begin{bmatrix} 00 \cdots 0 & 00 \cdots 0 \\ 00 \cdots 0 & 00 \cdots 0 \\ \vdots & \vdots \\ 00 \cdots 0 & 00 \cdots 0 \\ 00 \cdots 0 & -10 \cdots 0 \\ 00 \cdots 0 & 00 \cdots 0 \end{bmatrix} \varphi = G_i(\theta) \varphi; \quad i = 1, \dots, d = n + m + l \quad (5.229)$$

gde je $\varphi_i(\cdot)$ i -ta komponenta regresionog vektora u (5.212). Sa druge strane, na osnovu izraza (5.221) sledi

$$\text{col} \zeta(i+1, \theta) = F(\theta) \text{col} \zeta(i, \theta) + \text{col} M = F(\theta) \text{col} \zeta(i, \theta) + [M_1 \cdots M_i \cdots M_d] \quad (5.230)$$

gde su kolone matrice M definisane izrazom (5.229). Uzimajući u obzir (5.212), (5.219), (5.224) i (5.230) dobija se direktno izraz (5.226) za matrice A i B u proširenom modelu u prostoru stanja (5.225).

Na sličan način, na osnovu (5.214) i (5.218) dobija se

$$\frac{\partial}{\partial \theta_i} H(\theta) = \frac{\partial}{\partial \theta_i} \theta^T = \{0 \cdots 10 \cdots 0\}_{1 \times d} \quad (5.231)$$

pa iz (5.217) sledi

$$D(\varphi, \theta) = \varphi^T \quad (5.232)$$

Na taj način, imajući u vidu (5.219) i (5.232), relacija (5.223) može se zapisati u obliku

$$\eta = H(\theta) \text{col} \zeta + \varphi \quad (5.233)$$

pa na osnovu (2.208), (5.211), (5.222), (5.224) i (5.233) matrica C u jednačini izlaza u proširenom modelu u prostoru stanja (5.225) određena je kao

$$C(\theta) = \begin{bmatrix} H(\theta) & 0_{1 \times d} & 0_{1 \times d} & \cdots & 0_{1 \times d} \\ & H(\theta) & 0_{1 \times d} & \cdots & 0_{1 \times d} \\ I_{d \times d} & 0_{1 \times d} & H(\theta) & \cdots & 0_{1 \times d} \\ & \vdots & \vdots & \ddots & \vdots \\ & 0_{1 \times d} & 0_{1 \times d} & \cdots & H(\theta) \end{bmatrix} \quad (5.234)$$

Primetimo da je matrica u (5.226) trougaona i da sadrži matricu F u svakom od $(d+1)$ -og dijagonalnog blok elementa. Ovakva struktura matrice A ima za posledicu da su sopstvene vrednosti matrice A identične sa sopstvenim vrednostima matrice F u (5.213), samo će stepen višestrukosti sopstvenih vrednosti matrice A biti veći u poredjenju sa stepenom višestrukosti odgovarajućih sopstvenih vrednosti matrice F . Analizom matrice F pokazuje se da ona poseduje nultu sopstvenu vrednost višestrukosti $m+n$, dok preostalih l sopstvenih vrednosti predstavljaju nule polinoma $C(\cdot)$. Na taj nači, skup

$$D_s = \{ \theta : \text{sopstvene vrednosti } F(\theta) \text{ leže unutar jediničnog kruga } |z| < 1 \} \quad (5.235)$$

identičan je skupu

$$D_s = \{ \theta : \text{koreni jednačine } C(z^{-1}) = 0 \text{ leže unutar jediničnog kruga } |z| < 1 \} \quad (5.236)$$

Drugim rečima, stabilnost proširenog modela u prostoru stanja (5.224)-(5.226), (5.234) uslovljena je stabilnošću linearnog prediktora (5.211)-(5.213), a ovaj sistem će biti stabilan pod uslovom da vektor parametara pripada skupu (5.235), odnosno (5.236). Primetimo, takodje, da je određivanje predikcije izlaza i njenog gradijenta na osnovu jednačine izlaza u proširenom modelu u prostoru stanja (5.225) ne efikasno u praktičnom smislu, s obzirom da su matrice koje definišu ovaj model retke, odnosno sadrže veliki broj elemenata čija je vrednost jednaka nuli. S toga se ovaj model koristi za teoretske analize, ali ne i za praktičnu implementaciju parametarskog etimatora. Označimo, dalje, sa $M(\theta)$ skup razmatranih modela, koji se dobija variranjem vrednosti vektora parametara θ u ARMAX reprezentaciji (5.183)-(5.187) u okviru nekog podskupa $D_\mu \in R^d$, gde R^d označava d dimenzioni Euklidov prostor. Pošto je generisanje predikcije stabilno jedino ako $\theta \in D_s$, to je prirodno da se pretpostavi da je $D_\mu \subset D_s$. Medjutim, osnovnu poteškoću u primeni prediktora (5.225) u okviru rekurzivne procedure za estimaciju vektora parametara θ predstavlja činjenica da algoritam (5.225) u opštem slučaju nije rekurzivan. Ovakav zaključak

zasniva se na činjenici da izračunavanje vektora stanja $q(i, \hat{\theta}(i-1))$ na osnovu (5.225), gde je $\hat{\theta}(i-1)$ procena vektora parametara θ u $(i-1)$ -vom diskretnom trenutku vremena, ne zahteva samo tekuća merenja ulaza i izlaza sistema $z(i)$, već zahteva pamćenje i prethodnog skupa mernih podataka $\{z(k), k \leq i\}$, sve dok matrica $A(\hat{\theta}(i-1))$ ne postane nulpotentna, odnosno dok ne bude ispunjeno $A^n(\hat{\theta}(i-1)) = 0$ za neko n . Ovaj problem može se prevazići na sledeći način. Naime, rešenje vektorske diferencne jednačine stanja (5.225) dato je sa [16,24,48]

$$q(i, \hat{\theta}(i-1)) = A^i(\hat{\theta}(i-1))q(0) + \sum_{k=0}^{i-1} A^{i-k-1}(\hat{\theta}(i-1))B(\hat{\theta}(i-1))z(k) \quad (5.237)$$

Ako je dalje zadovoljeno da $\hat{\theta}(i-1) \in D_s$, tada će $A^k(\hat{\theta}(i-1)) \rightarrow 0$ sa porastom indeksa k , te će dominantni članovi u sumi (5.237) imati vrednost indeksa $k = i-1, \dots, i-K$. Sa druge strane, ako se za rekurzivnu estimaciju parametara u usvojenom skupu modela primeni algoritam (5.206)-(5.210), skalarni factor $1/i$ u algoritmu za estimaciju parametara (5.207) biće mali broj za dovoljno veliku vrednost indeksa vremena i , pa će u navedenom slučaju biti vrlo male razlike između procena vektora parametara $\hat{\theta}(i-1), \dots, \hat{\theta}(i-K)$. Ove dve činjenice opravdavaju uvođenje sledeće aproksimacije izraza (5.237)

$$q(i, \hat{\theta}(i-1)) \approx q(i) = \sum_{k=0}^{i-1} \left[\prod_{s=k+1}^{i-1} A(\hat{\theta}(s)) \right] B(\hat{\theta}(k))z(k) \quad (5.238)$$

na osnovu koje se dobija pogodna rekurzivna aproksimacija izraza (5.225) u obliku relacije

$$q(i+1) = A(\hat{\theta}(i))q(i) + B(\hat{\theta}(i))z(i) \quad (5.239)$$

Izvedeni rezultat omogućava da se modifikuje i jednačina izlaza u proširenom modelu u prostoru stanja (5.225) na sledeći način

$$\gamma(i, \hat{\theta}(i-1)) = \begin{bmatrix} \hat{y}(i/\hat{\theta}(i-1)) \\ \eta(i/\hat{\theta}(i-1)) \end{bmatrix} \approx \begin{bmatrix} \hat{y}(i) \\ \eta(i) \end{bmatrix} = C(\hat{\theta}(i-1))q(i) \quad (5.240)$$

U izvedenoj rekurzivnoj aproksimaciji (5.239), (5.240) linearnog prediktora u prostoru stanja (5.225) matrice $A(\cdot)$ i $B(\cdot)$ definisane su relacijama (5.226), dok je matrica $C(\cdot)$ definisana izrazom (5.234). Da bismo linearni predictor (5.239), (5.240) primenili u okviru rekurzivne stohastičke procedure za estimaciju vektora parametara (5.194), (5.196), (5.206)-(5.211) proširimo vector stanja još i sa tekućim merenjima ulaza i izlaza $z(i)$, čime se dobija modifikovani linearni prediktor u obliku jednačine stanja

$$\begin{aligned}
 x(i+1) &= \begin{bmatrix} \varphi(i+1)_{d \times 1} \\ \text{col} \zeta(i+1)_{d^2 \times 1} \\ z(i+1)_{2 \times 1} \end{bmatrix}_{(d^2+d+2) \times 1} = \begin{bmatrix} q(i+1)_{d(d+1)} \\ z(i+1)_{2 \times 1} \end{bmatrix}_{d(d+1)+2} \\
 &= \begin{bmatrix} A(\hat{\theta}(i))_{d(d+1) \times d(d+1)} & B(\hat{\theta}(i))_{d(d+1) \times 2} \\ 0_{2 \times d(d+1)} & 0_{2 \times 2} \end{bmatrix} x(i) + \begin{bmatrix} 0_{d(d+1) \times 2} \\ I_{2 \times 2} \end{bmatrix} z(i) \\
 &= A^*(\hat{\theta}(i))x(i) + B^*(\hat{\theta}(i))z(i)
 \end{aligned} \tag{5.241}$$

gde su matrice $A(\cdot)$ i $B(\cdot)$ definisane relacijama (5.226) a novi vektor stanja dat je sa

$$x(i) = \begin{bmatrix} q(i)_{d(d+1) \times 1} \\ z(i)_{2 \times 1} \end{bmatrix}_{(d+d^2+2) \times 1} = \begin{bmatrix} \varphi(i)_{d \times 1} \\ \text{col} \zeta(i)_{d^2 \times 1} \\ z(i)_{2 \times 1} \end{bmatrix}_{(d+d^2+2) \times 1} = \begin{bmatrix} Z(i)_{d \times 1} \\ \text{col} \frac{dZ(i)}{d\theta}_{d^2 \times 1} \\ z(i)_{2 \times 1} \end{bmatrix}_{(d+d^2+2) \times 1} ; z(i) = \begin{bmatrix} y(i) \\ u(i) \end{bmatrix}_{2 \times 1} \tag{5.242}$$

dok je jednačina novog izlaza, na osnovu (5.196) i (5.208)-(5.211)

$$\kappa(i) = \begin{bmatrix} v(i)_{1 \times 1} \\ Z_f(i)_{d \times 1} \end{bmatrix} = \begin{bmatrix} 0_{1 \times d} & 0_{1 \times d^2} & \left\{ \frac{A^{i-1}(q^{-1})}{B^{i-1}(q^{-1})} - \frac{B^{i-1}(q^{-1})}{C^{i-1}(q^{-1})} \right\}_{1 \times 2} \\ \frac{1}{C^{i-1}(q^{-1})} I_{d \times d} & 0_{d \times d^2} & 0_{d \times 2} \end{bmatrix} x(i)_{(d+d^2+2) \times 1} \tag{5.243}$$

U poslednjoj relaciji gornji indeks $(i-1)$ označava da su $A(\cdot)$, $B(\cdot)$ i $C(\cdot)$ polinomi definisani relacijom (5.185), samo su nepoznati koeficijenti zamenjeni njihovim procenama $\hat{\theta}(i-1)$ u $(i-1)$ -vom koraku. Kompletan algoritam za estimaciju parametara ARMAX modela (5.185) definisan je relacijama (5.206), (5.207), dok se vektor regresije $Z(\cdot)$, greška predikcije $v(\cdot)$ i filtrirani vektor regresije $Z_f(\cdot)$ izračunavaju na osnovu jednačina linearnog prediktora (5.241)-(5.243). Međutim, kao što je već istaknuto, neophodno je da se obezbedi da procena vektora parametara $\hat{\theta}(i)$ ostane u okviru skupa (5.235). Ovo se može postići uvođenjem projekcionog mehanizma u algoritam (5.207)

$$\hat{\theta}(i) = \left[\hat{\theta}(i-1) + \frac{1}{i} R^{-1}(i) Z_f(i) \psi(v(i)) \right]_{D_\mu \subset D_s} \tag{5.244}$$

gde je

$$[\hat{\theta}]_{D_\mu} = \begin{cases} \hat{\theta} & \hat{\theta} \in D_\mu \\ \text{neko } \hat{\theta} \in D_\mu & \hat{\theta} \notin D_\mu \end{cases} \tag{5.245}$$

Kao što je pokazano uslov (5.245) svodi se na uslov da koreni jednačine $C^i(z^{-1}) = 0$ zadovoljavaju pretpostavku (5.236), a za proveru ovog zahteva može se koristiti neki od testova stabilnosti iz teorije linearnih diskretnih sistema, na primer Džurijev test [16]. Problem kontinuiranog praćenja stabilnosti inverznog filtra $D^i(z^{-1}) = 1/C^i(z^{-1})$, gde je z kompleksni broj, može se prevazići usvajanjem $C^{i-1}(q^{-1}) = 1$ u (5.196) i (5.243), kada se posmatrani algoritam svodi na rekurzivnu stohastičku proceduru bez filtriranja podataka (5.188)-(5.191), čija je konvergencija analizirana prethodno primenom teorije martingala. Tada je pokazano da ovaj algoritam konvergira pod pretpostavkom $H6$, koja se za slučaj linearne $\psi(\cdot)$ funkcije svodi na uslov striktno pozitivne realnosti polinoma $C(\cdot)$ u ARMAX modelu (5.185), što je dobro poznat rezultat u literature za pseudo-linearne regresione metode [28,63]. Takodje, razmatrani algoritam sa filtracijom podataka predstavlja nelinearnu ekstenziju u literature predloženog rekurzivnog algoritma maksimalne verodostojnosti tipa 2 [36,82].

Sledeći metodologiju ODE pristupa [63], sada se može formulisati sledeća teorema kojom se rekurzivnoj stohastičkoj estimacionoj proceduri pridružuje odgovarajući sistem običnih diferencijalnih jednačina (ODE sistem).

Teorema 5.6: Posmatrajmo model sistema (5.185)-(5.187) koji je poznat sa tačnošću do nepoznatog vektora parametara θ i rekurzivni algoritam (5.192)-(5.199) za estimaciju nepoznatog vektora parametara, koji se može napisati i u alternativnom obliku (5.196), (5.197), (5.206)-(5.211), pri čemu se vektor regresije $Z(\cdot)$, greška predikcije $v(\cdot)$ i filtrirani vektor regresije $Z_f(\cdot)$ mogu izračunati i na osnovu jednačina linearnog prediktora (5.241)-(5.243), odnosno

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \frac{1}{i} R^{-1}(i) Z_f(i) \psi(v(i)) \quad (5.246)$$

$$R(i) = R(i-1) + \frac{1}{i} [\alpha(i) Z_f(i) Z_f^T(i) - R(i-1)] \quad (5.247)$$

$$\begin{aligned} x(i+1) &= \begin{bmatrix} A(\hat{\theta}(i))_{d(d+1) \times d(d+1)} & B(\hat{\theta}(i))_{d(d+1) \times 2} \\ 0_{2 \times d(d+1)} & 0_{2 \times 2} \end{bmatrix} x(i) + \begin{bmatrix} 0_{d(d+1) \times 2} \\ I_{2 \times 2} \end{bmatrix} z(i) \\ &= A^*(\hat{\theta}(i)) x(i) + B^*(\hat{\theta}(i)) z(i) \end{aligned} \quad (5.248)$$

gde su matrice $A(\cdot)$ i $B(\cdot)$ definisane relacijama (5.226) a vektor stanja i vektor izlaza dati su sa

$$x(i) = \begin{bmatrix} Z(i)_{d \times 1} \\ \text{col} \frac{dZ(i)}{d\theta}_{d^2 \times 1} \\ z(i)_{2 \times 1} \end{bmatrix}_{(d+d^2+2) \times 1}; Z(i) = \begin{bmatrix} -y(i-1) \\ \vdots \\ -y(i-n) \\ u(i-1) \\ \vdots \\ u(i-m) \\ v(i-1) \\ \vdots \\ v(i-l) \end{bmatrix}; \text{col} \frac{dZ(i)}{d\theta} = \begin{bmatrix} \frac{dZ(i)}{d\theta_1}_{d \times 1} \\ \vdots \\ \frac{dZ(i)}{d\theta_d}_{d \times 1} \end{bmatrix}; z(i) = \begin{bmatrix} y(i) \\ u(i) \end{bmatrix} \quad (5.249)$$

$$\begin{bmatrix} v(i)_{1 \times 1} \\ Z_f(i)_{d \times 1} \end{bmatrix} = \begin{bmatrix} 0_{1 \times d} & 0_{1 \times d^2} & \left\{ \frac{A^{i-1}(q^{-1})}{B^{i-1}(q^{-1})} \quad -\frac{B^{i-1}(q^{-1})}{C^{i-1}(q^{-1})} \right\}_{1 \times 2} \\ \frac{1}{C^{i-1}(q^{-1})} I_{d \times d} & 0_{d \times d^2} & 0_{d \times 2} \end{bmatrix} x(i) \quad (5.250)$$

$$= \begin{bmatrix} C_1^*(\hat{\theta}(i-1)) \\ C_2^*(\hat{\theta}(i-1)) \end{bmatrix} x(i) = C^*(\hat{\theta}(i-1)) x(i)$$

U poslednjoj relaciji gornji indeks $(i-1)$ označava da su $A(\cdot)$, $B(\cdot)$ i $C(\cdot)$ polinomi definisani relacijom (5.185), samo su nepoznati koeficijenti zamenjeni njihovim procenama $\hat{\theta}(i-1)$ u $(i-1)$ -vom koraku. Pretpostavimo, takodje, da su ispunjene sledeće pretpostavke

P1: Skup razmatranih modela $M(\theta) = \{\theta : \theta \in D_\mu \in R^d\}$ opisanih sa (5.185)-(5.187) je takav da $D_\mu \subset D_s$, gde je skup D_s definisan sa (5.235) i (5.236), čime je obezbedjeno da se sve sopstvene vrednosti matrica $A^*(\theta)$ u (5.248) za svako $\theta \in D_\mu \subset D_s$ nalaze unutar jediničnog kruga $|z| < 1$ sa centrom u koordinatnom početku z -kompleksne ravni.

P2: Generisanje ulazno-izlaznih signala sistema $z(i)$ je takvo da poseduje asimptotski stacionarnu srednju vrednost, tako da za svako $\theta \in D_\mu \subset D_s$ postoje granične vrednosti

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N Z_f(i) v(i) = f(\theta)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N [\alpha(i) Z_f(i) Z_f^T(i) - R(i-1)] = F(R, \theta)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N [1 + \|z(i)\|]^3 < \infty$$

P3: Matrica pojačanja algoritma $R(i)$ u (5.206) je simetrična i za svaki indeks diskretnog vremena i zadovoljava uslov $R(i) > \delta I$, gde je I jedinična matrica odgovarajućih dimenzija, a $\delta > 0$ je neka pozitivna konstanta.

P4: Generisanje ulazno-izlaznih podataka je eksponencijalno stabilno, u smislu da su potonji podaci $z(k)$, $k < i$ skoro sigurno nezavisni od tekućih podataka $z(i)$, za k dovoljno manje od i .

P5: Kriterijumska funkcija $\rho(\cdot)$ u indeksu performance $J_i(\hat{\theta}) = i^{-1} \sum_{k=1}^i \rho(v(k, \hat{\theta}))$, čiji se minimum određuje rekurzivno primenom predloženog algoritma, je dva puta diferencijabilna, odnosno postoje izvodi $\psi(\cdot) = \rho'(\cdot)$ i $\psi'(\cdot) = \rho''(\cdot)$.

P6: Izvodi kriterijumske funkcije u P5 su glatke funkcije u smislu da zadovoljavaju uslove

$$\|\psi(v(i, \theta))\| \leq C(1 + \|v(\cdot)\|^2); \left\| \frac{\partial}{\partial \theta} \psi(v(i, \theta)) \right\| \leq C(1 + \|v(\cdot)\|); \left\| \frac{\partial}{\partial v} \psi(v(i, \theta)) \right\| \leq C(1 + \|v(\cdot)\|)$$

Tada se rekurzivnoj stohastičkoj estimacionoj proceduri (5.246)-(5.250) može pridružiti sledeći sistem običnih diferencijalnih jednačina (skraćeno ODE)

$$\frac{d\theta(\tau)}{d\tau} = R^{-1} f(\theta); \frac{dR(\tau)}{d\tau} = F(R, \theta) \quad (5.251)$$

gde su funkcije $f(\cdot)$ i $F(\cdot)$ definisane u hipotezi P2.

Dokaz: Predloženi algoritam predstavlja specijalan slučaj opšteg rekurzivnog algoritma tipa stohastičkog gradijenta

$$\begin{aligned} x^*(i) &= x^*(i-1) + \mu(i) Q(i, x^*(i-1), \varphi^*(i)) \\ \varphi^*(i) &= A(x^*(i-1)) \varphi^*(i-1) + B(x^*(i-1)) e(i) \end{aligned} \quad (5.252)$$

za koji je pretpostavljeno da ispunjava sledeće zahteve [63]:

C1: Funkcija $Q(i, x^*, \varphi^*)$ je neprekidna, u Lipšicovom smislu, po x^* i φ^* u bilo kojoj okolini $(\bar{x}, \bar{\varphi})$, gde je $\bar{x} \in D_R$, pri čemu je D_R podskup prostora parametara x^* a $\bar{\varphi}$ je proizvoljno, odnosno zadovoljava Lipšicov uslov

$$\|Q(i, x_1, \varphi_1) - Q(i, x_2, \varphi_2)\| \leq K(\bar{x}, \bar{\varphi}, \sigma, \nu) [\|x_1 - x_2\| + \|\varphi_1 - \varphi_2\|]$$

za $\|x_1 - \bar{x}\| \leq \sigma$ i $\|\varphi_1 - \bar{\varphi}\| \leq \nu$, gde je $\sigma = \sigma(\bar{x}) > 0$ i $\nu = \nu(\bar{\varphi}) > 0$.

C2: Matrične funkcije $A(x^*)$ i $B(x^*)$ su neprekidne, u Lipšicovom smislu, za $x^* \in D_R$.

C3: Diskretan niz $\{e(i)\}$ je takav da je ispunjeno

$$\lim_{i \rightarrow \infty} \sum_{k=1}^i \beta(i, k) Q(k, \bar{x}, \varphi^*(k, \bar{x})) = f(\bar{x}), \forall \bar{x} \in D_R$$

gde je

$$\varphi^*(i, \bar{x}) = A(\bar{x})\varphi^*(i-1, \bar{x}) + B(\bar{x})e(i); \varphi^*(0, \bar{x})$$

dok su težinski koeficijenti $\beta(i, k)$ povezani sa težinskim koeficijentima $\mu(i)$ na sledeći način

$$\beta(k, i) = \mu(i) \prod_{j=i+1}^k [1 - \mu(j)]$$

C4: Za svako $x^* \in D_R$ postoji neko $C(\bar{x}, \lambda, c)$ takvo da je

$$\lim_{i \rightarrow \infty} \sum_{k=1}^i \beta(i, k) [1 + \nu(k, \lambda, c)] K(\bar{x}, \varphi^*(k, \bar{x}), \sigma(\bar{x}), \nu(k, \lambda, c)) =$$

$$C(\bar{x}, \lambda, c); \nu(i, \lambda, c) = c \sum_{k=1}^i \lambda^{i-k} |e(k)|$$

gde je $K(\cdot)$ Lipšicova konstanta iz hipoteze C1, a λ je maksimalna sopstvena vrednost matrice matrice $A(\bar{x})$.

C5: Niz težinskih koeficijenata ispunjava uslove $\sum_{i=1}^{\infty} \mu(i) = \infty$; $\lim_{i \rightarrow \infty} \mu(i) = 0$.

Primetimo da hipoteze C1 i C2 obezbeđuju da su odgovarajuće funkcije od interesa dovoljno glatke, tako da male promene u x i φ neće prouzrokovati velike promene u procenama. Hipoteza C3 daje asimptotski usrednjeni pravac pretraživanja algoritma u x prostoru, dok je hipoteza C4 matematički iskaz zahteva da se neće pojaviti velika odstupanja od usrednjenog ponašanja. Prvi uslov u hipotezi C5 predstavlja matematičku formulaciju zahteva da procene mogu prelaziti proizvoljne distance, dok druga ralacija u C5 obezbeđuje konvergenciju procena bez obzira na slučajne fluktuacije u $Q(\cdot)$. Označimo, dalje, sa $\bar{D}$ zatvoreni podskup skupa D_R i pretpostavimo

da postoje konačna konstanta $C < \infty$ i niz diskretnih vremenskih trenutaka $\{i_k\}$, koji u opštem slučaju može zavisiti od pojedinačne sekvence $\{e(\cdot)\}$, takvi da $x^*(i_k) \in \bar{D}$, $|\phi^*(i_k)| < C$. Ako, uz navedeno, postoji i neka dva puta diferencijabilna funkcija $V(x^*)$ definisana na D_R , takva da je

$$V'(x^*) f(x^*) \leq 0, x^* \in D_R$$

u literature je dokazano da niz $\{x^*(i)\}$ generisan relacijom (5.252) konvergira na sledeći način [63]

$$x^*(i) \xrightarrow{i \rightarrow \infty} \left\{ \begin{array}{l} D_c = \{x^* : x^* \in D_R \wedge V'(x^*) f(x^*) = 0\} \\ \text{ili} \\ \text{klaster tački na granici skupa } D_R \end{array} \right\} \quad (5.254)$$

gde je $f(\cdot)$ srednji pravac pretraživanja algoritma (5.252) u C3. Postojanje klaster tačke na granici skupa D_R podrazumeva da postoji vremenski podniz $\{i_j\}$ takav da procena $x^*(i_j)$ u (5.252) teži ka granici skupa D_R kada indeks diskretnog vremena $i \rightarrow \infty$. U literature nije definisana posebna obična diferencijalna jednačina, ali su kasnije autori, imajući u vidu da je srednji pravac pretraživanja algoritma (5.252) u C3 dat sa $f(\cdot)$, rekurzivnoj stohastičkoj estimacionoj proceduri (5.252) pridružili običnu diferencijalnu jednačinu (ODE)

$$\frac{dx^*(\tau)}{d\tau} = f(x^*(\tau)) \quad (5.255)$$

koja ima ista asimptotska konvergencijska svojstva kao i rekurzivni stohastički algoritam (5.252) []. Uvedimo, dalje, opšti rekurzivni stohastički parametarski estimator, sa generalnim pravcem pretraživanja $H(\cdot)$ u prostoru parametara θ

$$R(i) = R(i-1) + \gamma(i) H(i, R(i-1), v(i), \eta(i)) \quad (5.256)$$

$$\theta(i) = \theta(i-1) + \gamma(i) R^{-1}(i) h(i, \theta(i-1), v(i), \eta(i)) \quad (5.257)$$

Specijalni slučaj ovog algoritma predstavlja Njutn-Rafsonov algoritam

$$\theta(i) = \theta(i-1) - \gamma(i) R^{-1}(i) \frac{d}{d\theta} \rho(i, \theta(i-1), v(i, \theta(i-1))) \quad (5.258)$$

$$\frac{d}{d\theta} \rho(i, \theta, v(i, \theta)) = \rho_\theta(i, \theta, v(i, \theta)) - \rho_v(i, \theta, v(i, \theta)) \eta(i, \theta) \quad (5.259)$$

$$v(i, \theta) = y(i) - y(i/\theta); \eta(i, \theta) = \frac{d}{d\theta} v(i, \theta) = -\frac{d}{d\theta} y(i/\theta) \quad (5.260)$$

$$R(i) = R(i-1) + \gamma(i) \left[\rho_{\theta\theta}(i, \theta(i-1), v(i, \theta(i-1))) + \eta(i, \theta(i-1)) \rho_{vv}(i, \theta(i-1), v(i, \theta(i-1))) \eta^T(i, \theta(i-1)) - R(i-1) \right] \quad (5.261)$$

koji rekurzivno minimizira generalni kriterijum

$$J(\theta) = \bar{E} \{ \rho(i, \theta, v(i, \theta)) \} = \lim_{i \rightarrow \infty} \frac{1}{i} \sum_{k=1}^i E \{ \rho(k, \theta, v(k, \theta)) \} \quad (5.262)$$

gde je matematičko očekivanje $E\{\cdot\}$ u odnosu na slučajni proces $\{z(i) = \{y(i), u(i)\}\}$ a $\bar{E}\{\cdot\}$ u (5.262) je generalisano matematičko očekivanje. U ovakvom algoritmu matrica pojačanja (5.261) predstavlja pogodnu aproksimaciju Hesijana kriterijuma (5.262), kojom se modifikuje gradijentni pravac pretraživanja (5.259), kako bi se dobio povoljniji pravac pretraživanja. Takodje, u izrazima (5.259) i (5.261) donji indeksi označavaju parcijalne izvode prvog i drugog reda u odnosu na odgovarajuće varijable. Predloženi algoritam (5.246), (5.247) predstavlja specijalni slučaj algoritma (5.258)-(5.262), kod koga kriterijumska funkcija $\rho(\cdot)$ u (5.262) zavisi samo od $v(\cdot)$ a ne i od θ , odnosno $\rho(i, \theta, v(i, \theta)) = \rho(i, v(i, \theta))$ tako da je $\rho_{\theta}(\cdot) = 0$ i $\rho_v(\cdot) = \rho'(\cdot) = \psi(\cdot)$, pri čemu je takodje uvedena aproksimacija $\eta(i, \theta(i-1)) \approx \eta(i) = Z_f(i)$. Na taj način, poredjenjem (5.246) i (5.247) sa (5.256) i (5.257), zaključuje se da je

$$h(i, \theta(i-1), v(i), \eta(i)) = h(i, v(i), \eta(i)) = \eta(i) \psi(v(i)) = Z_f(i) \psi(v(i)) \quad (5.263)$$

$$H(i, R(i-1), \theta(i-1), v(i), \eta(i)) = H(i, R(i-1), v(i), \eta(i)) = \alpha(i) \eta(i) \eta^T(i) - R(i-1) \quad (5.264)$$

Sledeći metodologiju predloženu u literature [63], uvedimo u generalni algoritam (5.256), (5.257) linearni prediktor u prostoru stanja za generisanje predikcije izlaza i njenog gradijenta

$$\zeta(i+1) = A(\theta(i-1)) \zeta(i) + B(\theta(i-1)) z(i) \quad (5.265)$$

$$\begin{bmatrix} \hat{y}(i+1) \\ \text{col} \eta(i+1) \end{bmatrix} = C(\theta(i)) \zeta(i+1) \quad (5.266)$$

$$v(i) = y(i) - \hat{y}(i); \eta(i) = -\frac{d}{d\theta} \hat{y}(i) \quad (5.267)$$

uz sledeće pretpostavke:

M1: Model prediktora (5.265)-(5.267) je stabilan za $\theta \in D_\mu$, gde je D_μ kompaktan podskup u okviru d-dimenzionog Euklidovog prostora R^d , odnosno za $\theta \in D_\mu$ sve sopstvene vrednosti matrice $A(\theta)$ nalaze se unutar jediničnog kruga $|z| < 1$ sa centrom u koordinatnom početku z-kompleksne ravni.

M2: Model prediktora je diferencijabilan po θ , za svako $\theta \in D_\mu$, odnosno matrice $A(\theta)$, $B(\theta)$ i $C(\theta)$ su neprekidno diferencijabilne po θ za svako $\theta \in D_\mu$.

C_h1 : Funkcije $h(\cdot)$ i $H(\cdot)$ u parametarskom estimatoru (5.256), (5.257) su dovoljno glatke, u smislu da je vektor $h(i, \theta, v, \eta)$ diferencijabilan po svim argumentima θ, v, η , tako da za neko $C < \infty$ i $\theta \in D_\mu$ važi

$$\|h(i, \theta, v, \eta)\| + \|h_\theta(i, \theta, v, \eta)\| \leq C[1 + |v|^2 + \|\eta\|^2]$$

$$\|h_v(i, \theta, v, \eta)\| + \|h_\eta(i, \theta, v, \eta)\| \leq C[1 + |v| + \|\eta\|]$$

dok je matrica $H(i, R, \theta, v, \eta)$ diferencijabilna po svim argumentima R, θ, v, η , tako da je

$$\|H(i, R, \theta, v, \eta)\| \leq C[1 + |v|^2 + \|\eta\|^2 + \|R\|]$$

$$\|H_R(i, R, \theta, v, \eta)\| + \|H_\theta(i, R, \theta, v, \eta)\| \leq C[1 + |v|^2 + \|\eta\|^2]$$

$$\|H_v(i, R, \theta, v, \eta)\| + \|H_\eta(i, R, \theta, v, \eta)\| \leq C[1 + |v| + \|\eta\|]$$

$C_\rho1$: Kriterijumska funkcija $\rho(i, \theta, v)$ je dovoljno glatka, u smislu da je tri puta neprekidno diferencijabilna po argumentima v, θ tako da je za neko $C < \infty$ i $\theta \in D_\mu$

$$|\rho_\theta(i, \theta, v)| + |\rho_{\theta\theta}(i, \theta, v)| \leq C[1 + |v|^2]$$

$$|\rho_v(i, \theta, v)| + |\rho_{v\theta}(i, \theta, v)| \leq C[1 + |v|]$$

$$|\rho_{vv}(i, \theta, v)| \leq C$$

$$|\rho_{vv}(i, \theta, v)| + |\rho_{vvv}(i, \theta, v)| + |\rho_{vv\theta}(i, \theta, v)| \leq C$$

$$|\rho_{\theta\theta}(i, \theta, v)| + |\rho_{\theta\theta\theta}(i, \theta, v)| \leq C[1 + |v|^2]$$

$$|\rho_{\theta\theta v}(i, \theta, v)| \leq C$$

D1: Generisanje ulazno-izlaznih podataka $z(i)$ je takvo da postoje asimptotski stacionarne srednje ili očekivane vrednosti

$$\bar{E}\{h(i, \theta, v(i, \theta), \eta(i, \theta))\} = f(\theta)$$

$$\bar{E}\{H(i, R, \theta, v(i, \theta), \eta(i, \theta))\} = F(R, \theta)$$

gde je generalisano matematičko očekivanje dato sa

$$\bar{E}\{f(i)\} = \lim_{i \rightarrow \infty} \frac{1}{i} \sum_{k=1}^i E\{f(i)\}$$

gde je $E\{\cdot\}$ klasično matematičko očekivanje u odnosu na dvodimenzionu slučajnu varijablu $z^T(i) = \{y(i), u(i)\}$. Uz navedeno, da bi obezbedili da su h_t i h_s nezavisni za $|t-s|$ dovoljno veliko, pretpostavimo da je generisanje podataka i eksponencijalno stabilno, u smislu da za svako t i s , $t \geq s$, postoji slučajni vektor $z_s^0(t)$ koji predstavlja aproksimaciju vektora $z(t)$, izračunatu na osnovu nove informacije u vremenskom intervalu (s, t) , a koji pripada σ -algebri generisanoj sa $z^t = \{z(i), i \leq t\}$ i nezavisan je od z^s (za $s=t$ usvaja se $z_s^0(t) = 0$), takav da je

$$E\{|z(t) - z_s^0(t)|^4\} < C\lambda^{t-s}, C < \infty, \lambda < 1.$$

R1: Generisanje matrice pojačanja algoritma $R(i)$ u (5.256) je takvo da je ova matrica simetrična i $R(i) \geq \delta I$ za neko $\delta > 0$ ($A \geq B$ ima smisao da je $A-B$ pozitivno semidefinitna matrica), čime je zagarantovano postojanje odgovarajuće inverzne matrice.

G1: Skalarni factor pojačanja u (5.256), (5.257) ispunjava uslov $\lim_{i \rightarrow \infty} i\gamma(i) = \mu > 0$.

Algoritam (5.256), (5.257), (5.265)-(5.267) može se svesti na oblik (5.252) uvodjenjem proširenog vektora stanja

$$\zeta^*(i) = \begin{bmatrix} \zeta(i) \\ z(i) \end{bmatrix}; A^*(\theta) = \begin{bmatrix} A(\theta) & B(\theta) \\ 0 & 0 \end{bmatrix}; C^*(\theta) = \begin{bmatrix} C_1^*(\theta) \\ C_2^*(\theta) \end{bmatrix} = \begin{bmatrix} -C_1(\theta) & I & 0 \\ C_2(\theta) & 0 & 0 \end{bmatrix}; B^*(\theta) = \begin{bmatrix} 0 \\ I \end{bmatrix} \quad (5.268)$$

nakon čega iz (5.265)-(5.267) sledi

$$\zeta^*(i) = A^*(\theta(i-1))\zeta^*(i-1) + B^*(\theta(i-1))z(i) \quad (5.269)$$

$$\begin{bmatrix} v(i) \\ col\eta(i) \end{bmatrix} = C^*(\theta(i-1))\zeta^*(i) \quad (5.270)$$

nakon uvođenja odgovarajućih smena promenljivih

$$x^* = \begin{bmatrix} \theta \\ colR \end{bmatrix}; \varphi^* = \zeta^*; Q(i, x^*, \varphi^*) = \begin{bmatrix} \mu(i)R^{-1}h(i, \theta, C_1^*(\theta)\zeta^*, C_1^*(\theta)\zeta^*) \\ \mu(i)colH(i, R, \theta, C_1^*(\theta)\zeta^*, C_1^*(\theta)\zeta^*) \end{bmatrix} \quad (5.271)$$

$$\gamma(i) = \frac{1}{i}; A(x^*) = A^*(\theta); B(x^*) = B^*(\theta); e(i) = z(i) \quad (5.272)$$

U literaturi je pokazano da je algoritam (5.256), (5.257), (5.265)-(5.267) pod uslovom da zadovoljava hipoteze M1, C_{h1} , $C_{\rho1}$, D1, R1, G1 identičan algoritmu (5.252), (5.268)-(5.272), kao i da poslednji algoritam zadovoljava uslove C1-C6. Zbog kompletnosti izvođenja, skiciraćemo ovaj dokaz u najkraćim crtama (za detaljno izvođenje videti str. 174-176 u [63]). Naime, uslov C1 sledi na osnovu direktnog izračunavanja, uz korišćenje hipoteza C_{h1} i M1, kao i činjenice da se na osnovu hipoteze R1 može pisati $\|R^{-1}\| < \delta^{-1}$, dok na osnovu hipoteze G1 sledi $\mu(i) < \bar{\mu}$, čime se konačno dobija

$$K(\bar{x}, \bar{\varphi}, \sigma, \nu) = K(\theta, R, \zeta, \sigma, \nu) = C[1 + \nu + \|\zeta\|]^2; \theta \in D_\mu$$

gde je C konstanta dobijena na osnovu $\delta, \bar{\mu}$ i konstanti u okviru C_{h1} .

Uslov C2 je direktna posledica hipoteze M1.

Dokaz uslova C3 zasniva se na hipotezi D1. Naime, pokazuje se da na osnovu D1 sledi [63]

$$\lim_{i \rightarrow \infty} \frac{1}{i} \sum_{k=1}^i h(k, \theta, \nu(k, \theta), \eta(k, \theta)) = f(\theta)$$

$$\lim_{i \rightarrow \infty} \frac{1}{i} \sum_{k=1}^i H(k, R, \theta, \nu(k, \theta), \eta(k, \theta)) = F(R, \theta)$$

$$\limsup_{i \rightarrow \infty} \frac{1}{i} \sum_{k=1}^i [1 + \|z(k)\|]^3 < \infty$$

tako da se prve dve relacije razlikuju od C3 samo po faktoru $\mu(i)$ i R^{-1} . Pokazuje se da C3 sledi na osnovu gornjih izraza, uz korišćenje Holderove nejednakosti, kao i hipoteza R1, G1 i C_{h1} .

Uslov C4 sledi na osnovu definicionog izraza u C4, ograničenosti podataka prema poslednjoj relaciji, definicionog izraza za vektor stanja $\zeta(\cdot)$ i hipoteze M1.

Uslovi C5 i C6 su automatski ispunjeni za izabrani faktor pojačanja $\gamma(i) = 1/i$.

Predloženi algoritam u okviru Teoreme 5.6 definisan je relacijama (5.246)-(250) i predstavlja specijalan slučaj generalnog algoritma (5.256), (5.257), (5.265)-(5.267), za specifičan izbor funkcija $h(\cdot)$ i $H(\cdot)$ u (5.263) i (5.264), respektivno. S toga, uvedene pretpostavke P1-P6 predstavljaju tehnički iskaz hipoteza M1, $C_h 1, C_\rho 1$, D1, R1 i G1 za navedeni konkretni slučaj, tako da su uslovi C1-C6 zadovoljeni i za pretpostavke P1-P6, tako da na osnovu definicije vektora stanja u (5.271) i pridružene diferencijalne jednačine u (5.255) direktno sledi izraz (5.251), čime je dokaz teoreme završen.

Pretpostavke M1, M2, $C_h 1, C_\rho 1$, R1 i G1, odnosno odgovarajuće hipoteze u okviru teoreme 5.6 odnose se na funkcije i varijable koje bira sam projektant, tako da se one uvek mogu potpuno verifikovati. Međutim, uslovi D1, odnosno ekvivalentne pretpostavke P2 i P4 u Teoremi 5.6, odnose se na podatke, te na njih treba gledati kao na hipoteze, pošto nije moguće da se izvrši njihova formalna verifikacija u praksi. Jedino je moguće da se uvedu neke pretpostavke o nepoznatom mehanizmu kojim se generišu ulazni i izlazni signali sistema koji se identifikuje, pa da se naknadno proveriti da li uvedene pretpostavke dovode do uslova koji su nam potrebni. U tom smislu su i uvedene odgovarajuće prirodne pretpostavke o načinu generisanja ulazno-izlaznih podataka, čime su obezbeđene asimptotski stacionarne potrebne očekivane vrednosti, kao i i eksponencijalna stabilnost.

Primetimo da uslov (5.253) ukazuje da je uvedena funkcija $V(\cdot)$ u suštini funkcija Ljapunova, koja se koristi za analizu stabilnosti rešenja običnih diferencijalnih jednačina [16,93]. Prema teoriji Ljapunova, ako je $V(x)$ pozitivna skalarna funkcija vektorskog argumenta, tj.

$$V(x) \geq 0 \quad , za \forall x \quad (5.273)$$

i to takva da kada se izračuna duž trajektorije $x(\tau)$ diferencijalne jednačine

$$x' = \frac{dx(\tau)}{d\tau} = f(x) \quad (5.274)$$

ova funkcija opada, odnosno zadovoljava uslove

$$\frac{d}{d\tau} V(x(\tau)) = \frac{d}{dx} V(x(\tau)) \frac{dx(\tau)}{d\tau} = V'(x(\tau)) f(x(\tau)) \leq 0 \quad , za \forall x(\tau) \quad (5.275)$$

$$\frac{d}{d\tau} V(x(\tau)) = 0 \Rightarrow x(\tau) \in D_c = \{x : x(0) \in D_c \Rightarrow x(\tau) \in D_c za \forall \tau\} \quad (5.276)$$

gde je D_c invarijantni skup date diferencijalne jednačine, tada se takva funkcija $V(x)$ naziva funkcijom Ljapunova. Kao što je matematički iskazano u (5.276), invarijantni skup zadovoljava uslov da bilo koja trajektorija koja započne unutar ovog skupa i ostaje u njemu. Van ovog skupa funkcija $V\{x(\tau)\}$ je striktno opadajuća kao funkcija vremenskog argumenta τ . Međutim, pošto je funkcija Ljapunova ograničena sa donje strane, to ona ne može da opada neograničeno, već će trajektorija, odnosno rešenje diferencijalne jednačine, $x(\tau)$ konvergirati ka invarijantnom skupu. Na taj način, uslovi (5.273)-(5.276) garantuju da će invarijantni skup D_c biti asimptotski globalno stabilan. Takodje, stacionarna tačka x^* diferencijalne jednačine (5.274) zadovoljava uslov

$$f(x^*) = 0 \quad (5.277)$$

i predstavlja jedan invarijantan skup, pošto je u važnosti

$$x(0) = x^* \Rightarrow x' = f(x^*) = 0 \Rightarrow x(\tau) = x^*, \text{ za } \forall \tau \quad (5.278)$$

Uz navedeno, bilo koji invarijantan skup D_c ima domen atrakcije D_A , koji je takav da bilo koja trajektorija koja započne u njemu konvergiraće ka njegovom invarijantnom skupu, t.j.

$$x(0) \in D_A \Rightarrow x(\tau) \xrightarrow{\tau \rightarrow \infty} D_c, D_c \subset D_A \quad (5.279)$$

Ako je D_A jednak celom skupu na kome je definisana diferencijalna jednačina, tada je skup D_c globalno asimptotski stabilan, dok ako je D_A striktno veći od D_c radi se o stabilnom invarijantnom skupu. Na taj način, sve trajektorije (rešenja) diferencijalne jednačine koje započnu u Euklidovom prostoru D_R ili će ostati u njemu ili će konvergirati ka invarijantnom skupu D_c kada vreme neograničeno raste. Dakle, uslov stabilnosti trajektorije (rešenja) posmatrane diferencijalne jednačine je uslov Ljapunova (5.275), odnosno

$$V'(x) f(x) \leq 0, x \in D_R \quad (5.280)$$

Ukratko izložena teorija Ljapunova može se direktno primeniti na običnu diferencijalnu jednačinu (5.251) koja je pridružena rekurzivnoj stohastičkoj estimacionoj proceduri (5.246)-(5.250), za analizu stabilnosti odgovarajuće trajektorije (rešenja) diferencijalne jednačine, s obzirom da su asimptotska svojstva predloženog rekurzivnog stohastičkog algoritma ekvivalentna asimptotskim svojstvima analizirane trajektorije. Rezultati su sumirani u okviru sledeće teoreme.

Teorema 5.7: Posmatrajmo algoritam (5.264)-(5.250), uz pretpostavku da zadovoljava hipoteze P1-P6 Teoreme 5.6, kao i da su ispunjeni dodatni uslovi da nelinearna transformacija rezidala merenja $\psi(\cdot)$ ima ograničen izvod, a da realni deo polinoma

$C(q^{-1})$ definisanog sa (5.196) i (5.197) pozitivan, t.j. $\operatorname{Re}\{C(q^{-1})\} > 0$. Tada procena vektora parametara (5.264) ARMAX modela (5.185)-(5.187) konvergira sa verovatnoćom 1 (s.v.1) ili skoro sigurno (s.s) ka stacionarnoj tački funkcije Ljapunova

$$V(\theta, R) = \frac{1}{2} E \{ \psi^2(v(i)) \} \quad (5.281)$$

gde je $v(\cdot)$ residual merenja u (5.250), a matematičko olekivanje je definisano u odnosu na dvodimenzioni vektor izlazno-ulaznih podataka $z^T(i) = \{y(i), u(i)\}$.

Dokaz: Dokaz se svodi na dokazivanje činjenice da je funkcija (5.281) zaista funkcija Ljapunova, koja zadovoljava uslove (5.273) i (5.275). Pošto je na osnovu definicije (5.281) ova funkcija pozitivna, to je uslov (5.273) ispunjen, te ostaje da se pokaže, u skladu sa zahtevom (5.275) i da je njen izvod nepozitivan. Diferenciranjem izraza (5.281) po vremenskoj varijabli τ dobija se

$$\frac{d}{d\theta} V(\theta, R) = E \{ \psi(v(i)) \psi'(v(i)) \nabla_{\theta}^T v(i) \} \frac{d\theta(\tau)}{d\tau} \quad (5.282)$$

gde je na osnovu (5.186), (5.187), (5.190) i (5.193)

$$v(i) = -Z^T(i)(\theta - \theta^*) + \xi(i) \quad (5.283)$$

U izrazu (5.283) simbol $*$ označava tačnu vrednost vektora parametara, što podrazumeva da i realan sistem pripada usvojenoj klasi modela i da je razlika između sistema i modela samo u vrednosti vektora parametara. Ukoliko to nije slučaj, vektor parametara sa gornjim indeksom $*$ označava samo minimum kriterijuma u P5. Imajući u vidu definiciju filtriranog regresionog vektora u (5.196), t.j.

$$Z_f(i) = \frac{1}{C(q^{-1})} Z(i) = D(q^{-1}) Z(i) \quad (5.284)$$

na osnovu (5.283) i (5.284) sledi

$$\nabla_{\theta} v(i) = -Z(i) = -C(q^{-1}) Z_f(i) \quad (5.285)$$

Zamenom izraza za gradijent reziduala ili greške predikcije (5.285) i izraza (5.251) za diferencijalnu jednačinu pridruženu rekurzivnoj stohastičkoj estimacionoj proceduri u relaciju (5.283), uz korišćenje pretpostavke da je $|\psi'(\cdot)| \leq k$ gde je k konačna, pozitivna konstanta, dobija se

$$\frac{d}{d\theta} V(\theta, R) = -k E \{ C(q^{-1}) \psi(v(i)) Z_f^T(i) \} R^{-1}(\tau) f(\tau) \quad (5.286)$$

gde je na osnovu pretpostavke P2 u teoremi 5.6

$$f(\theta) = E\{Z_f(i)\psi(v(i))\} \quad (5.287)$$

Na osnovu P2 takodje sledi da je funkcija $F(\theta, R)$ u (5.251) data sa

$$F(\theta, R) = E\{\alpha(i)Z_f(i)Z_f^T(i)\} - R, \quad \alpha(i) = \psi'(v(i)) \quad (5.288)$$

Uvrštavanjem izraza (5.287) u (5.286) može se pisati

$$\frac{d}{d\theta} V(\theta, R) = -kC(q^{-1})f^T(\tau)R^{-1}(\tau)f(\tau) \leq 0 \quad (5.289)$$

Relacija (5.289) je realna, te izvedena nejednakost sledi na osnovu usvojene pretpostavke da polinom $C(q^{-1})$, gde q sada ima značenje kompleksne varijable, zadovoljava uslov pozitivnosti realnog dela $\text{Re}\{C(q^{-1})\} > 0$ i hipoteze P3 u Teoremi 5.6, na osnovu koje je $R(i) > 0$, tako da je kvadratna forma u (5.289) nenegativna, t.j. $f^T R^{-1} f \geq 0$. U izrazu (5.289) znak jednakosti važi samo za neko $\theta = \theta^*$ za koje je $f(\theta^*) = 0$, odnosno koje predstavlja stacionarnu tačku funkcije Ljapunova $V(\cdot)$. Ukoliko je rešenje $\theta = \theta^*$ posmatrane nelinearne algebarske jednačine jedinstveno, procene vektora parametara (5.264) konvergiraju ka njemu skoro sigurno, u smislu da je

$$P\left\{\lim_{i \rightarrow \infty} \theta(i) = \theta^*\right\} = 1 \quad (5.290)$$

Ovim je dokaz teoreme završen.

Primetimo, na kraju, da se po analogiji sa Teoremama 5.6 i 5.7, koje se odnose na ARMAX reprezentaciju modela sistema, mogu formulisati i ekvivalentne teoreme za slučaj ARX i FIR modela sistema, koji su analizirani u prethodnom poglavlju primenom teorije konvergencije martingala. Jedina razlika u takvoj analizi u odnosu rezultate prikazane u ovom poglavlju predstavlja oblik matrica $F(\cdot)$ i $E(\cdot)$ u izrazu (5.213), koje u slučaju ARX modela postaju

$$F(\theta) = \begin{bmatrix} 00 & \dots & 00 & 00 & \dots & 00 \\ 10 & \dots & 00 & 00 & \dots & 00 \\ \vdots & & \vdots & \vdots & & \vdots \\ 00 & \dots & 10 & 00 & \dots & 00 \\ 00 & \dots & 00 & 00 & \dots & 00 \\ 00 & \dots & 00 & 10 & \dots & 00 \\ \vdots & & \vdots & \vdots & & \vdots \\ 00 & \dots & 00 & 00 & \dots & 10 \end{bmatrix}; E(\theta) = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \quad (5.291)$$

dok se za slučaj FIR modela svode na oblik

$$F(\theta) = \begin{bmatrix} 00 & \dots & 00 \\ 10 & \dots & 00 \\ \vdots & & \vdots \\ 00 & \dots & 10 \end{bmatrix}; E(\theta) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \quad (5.292)$$

U oba razmatrana slučaja matrica $M(\cdot)$ u (5.228) jednaka je nuli, te matrica $A(\cdot)$ u (5.226) postaje blok dijagonalna. Takođe, regresioni vektor u slučaju ARX modela ima oblik

$$\varphi^T(i) = Z^T(i) = \{-y(i-1), L, -y(i-n), u(i-1), L, u(i-m)\} \quad (5.293)$$

dok je isti vektor kod FIR modela dat sa

$$\varphi^T(i) = Z^T(i) = \{u(i-1), L, u(i-m)\} \quad (5.294)$$

Dalja analiza u potpunosti sledi korake izložene u ovom poglavlju.

Rezultati analize izvršene u ovom i prethodnom poglavlju imaju asimptotski karakter, odnosno u važnosti su na uzorku velike dužine, koja je u teorijskom smislu neograničena. Nažalost, u teoriji ne postoji kriterijum na osnovu koga se može odrediti neophodna veličina mernog uzorka koja će garantovati izvedene asimptotske osobine algoritama parametarske estimacije. Odgovor na postavljeno pitanje, kao i na mnogobrojna druga pitanja koja se odnose na osobine stohastičkih procedura parametarske estimacije na mernim uzorcima male dužine, kao i u različitim uslovima rada bliskih realnosti, mogu se dobiti samo putem eksperimentalne analize na bazi računarske simulacije. Rezultati ovakvih istraživanja prikazani su u narednom poglavlju.

5.4 Eksperimentalna Analiza na osnovu računarske simulacije

U ovom poglavlju prikazani su praktični rezultati koji su dobijeni na osnovu implementacije predloženih algoritama parametarske estimacije na računaru opšte namene, korišćenjem MatLab programskog paketa i metodologije Monte Karlo simulacije. Analizirane su dve ulazno-izlazne reprezentacije modela sistema, od kojih prva predstavlja specijalni oblik opšte forme ARX model (5.80)-(5.84), četvrtog reda kod koga su odgovarajući polinomi

$$A(q^{-1}) = 1 - q^{-1} + 0.18q^{-2} - 0.78q^{-3} + 0.656q^{-4}, \quad B(q^{-1}) = 2q^{-1} \quad (5.295)$$

tako da vektor linearnne regresije (5.293) dobija oblik

$$Z^T(i) = [-y(i-1) - y(i-2) - y(i-3) - y(i-4) \ u(i-1)] \quad (5.296)$$

dok je vektor pravih parametara

$$\Theta^{*T} = [1 \ -0.18 \ 0.78 \ -0.656 \ 2] \quad (5.297)$$

Pobudni signal $\{u(i)\}$ generisan je kao diskretan beli gausovski stohastički proces, nulte srednje vrednosti i jedinične varijanse, dok je stohastički poremećaj, t.j. aditivni šum merenja, $\{\xi(i)\}$ generisan kao realizacija ε -kontaminirane gausovske funkcije gustine verovatnoće (f.g.v)

$$p(\xi(i)) = p_\varepsilon(\xi(i)) = (1-\varepsilon)N(\xi(i)/0, \sigma^2) + \varepsilon h(\xi(i)), \quad 0 \leq \varepsilon \leq 1 \quad (5.298)$$

gde je $N(. / a, b)$ normalna f.g.v srednje vrednosti a i varijanse b , dok je $h(.)$ proizvoljna f.g.v. Uobičajeno je da se usvoji i da je f.g.v $h(.)$ takodje normalna f.g.v, nulte srednje vrednosti i varijanse σ_0^2 znatno veće od nominalne varijanse σ^2 . Uzorak ovakve f.g.v može se praktično realizovati generisanjem realizacije uniformno raspodeljene slučajne varijable r na intervalu $(0,1)$, pozivanjem odgovarajuće bibliotečke funkcije MatLab, te ukoliko je $r < 1 - \varepsilon$ treba pozvati odgovarajući generator pseudo-slučajnih brojeva u cilju generisanja uzorka sa nominalne normalne f.g.v $N(. / 0, \sigma^2)$, dok u protivnom treba pozvati generator pseudo-slučajnih brojeva da bi se generisao uzorak sa usvojenom f.g.v $h(.)$. Primetimo, takodje, da MatLab poseduje biblioteku funkciju kojom se generišu realizacije standardne normalne f.g.v $N(. / 0, 1)$, te ukoliko sa x označimo takvu realizaciju, tada je realizacija y slučajne varijable $N(. / 0, \sigma^2)$ data sa $y = \sigma x$. Za proizvoljnu f.g.v $h(.)$ potrebno je da se napiše MatLab programski kod za željeni generator pseudo-slučajnih brojeva, koristeći odgovarajući algoritam[49].

Cilj eksperimenta je bio da se uporede osobine predloženog rekurzivnog nelinearnog stohastičkog algoritma (5.85)-(5.89) za estimaciju parametara u usvojenom modelu sistema (5.295)-(5.297) sa standardnim rekurzivnim algoritmom najmanjih kvadrata, koji sledi iz prethodnog algoritma kada se usvoji linearna transformacija reziduala merenja, t.j. $\psi(v) = v$, a vrednost parametra α odredi na osnovu poznate vrednosti

nominalne varijanse šuma u modelu (5.298), odnosno $\alpha = \sigma^{-2}$. Prvi algoritam označen je skraćeno sa RRA, a drugi sa RLS. U prvom algoritmu parametar α predstavlja koeficijent statističke linearizacije i izračunava se kao matematičko očekivanje izvoda nelinearne transformacije reziduala (5.88) u odnosu na usvojenu f.g.v aditivnog šuma merenja (5.298), t.j. $\alpha = E\{\psi'(v(i))\}$. U opštem slučaju ovo očekivanje se ne može odrediti egzaktno, s obzirom da vrednosti stepena kontaminacije ε , kao i kontaminirajući član $h(\cdot)$ i njegova varijanse σ_0^2 u modelu šuma (5.298) nisu nikad tačno poznati u praksi. Kao što je diskutovano u četvrtom poglavlju, jedna mogućnost je da se ovo matematičko očekivanje izračuna na osnovu najnepovoljnije f.g.v p_0 u okviru usvojene klase (5.298), koja nosi minimum informacije o estimiranim parametrima, u smislu Fišerove količine informacije $I(p)$. U šestom poglavlju je pokazano da je ovakva f.g.v normalna u središnjem delu i eksponencijalna u repovima, na osnovu čega se dobija

$$p_0 = \arg \min_{p \in p_\varepsilon} I(p), I(p) = E\left\{\left(\frac{p'}{p}\right)^2\right\}, \alpha = I(p_0) = [2(1-\varepsilon) \operatorname{erf}(k) \sigma^2]^{-1} \quad (5.299)$$

gde je $\operatorname{erf}(\cdot)$ funkcija greške, dok k označava podešljivi parametar u izrazu za $\psi(\cdot)$ funkciju (5.88), a σ^2 predstavlja varijansu normalnog člana u familiji f.v.g (5.298). Alternativna mogućnost je da se uvede klizeći prozor oko tekućeg diskretnog trenutka vremena i odgovarajuće dužine L i matematičko očekivanje aproksimira odgovarajućom aritmetičkom sredinom

$$\alpha = \alpha(i) = \frac{1}{L} \sum_{k=i-L+1}^i \psi'(v(k)) \quad (5.300)$$

gde je $v(\cdot)$ rezidual merenja u (5.86).

Efikasna robusnost algoritama (odgovarajući engleski termin je Efficiency Robustness), u smislu njihove neosetljivosti na prisustvo realizacija impulsnog šuma, ili outlier-a, koje su generisane kontaminirajućom f.g.v $h(\cdot)$ u (5.298), ilustrovana je rezultatima prikazanim u okviru Tabela 5.1, 5.2 i 5.3. Kao kriterijum za poredjenje algoritama korišćena je ukupna srednje-kvadratna greška izračunata usrednjavanjem rezultata 100 Monte Karlo realizacija. Dobijeni rezultati pokazuju da je RLS algoritam nešto efikasniji od RRA algoritma u slučaju čisto normalnog šuma merenja (Tabela 5.1; $\varepsilon = 0$ u (5.298)), što je u skladu sa dobro poznatom činjenicom da je RLS metod u navedenom slučaju optimalan, u smislu minimalne totalne varijanse greške estimacije [28,49,63]. Oba algoritma zadržavaju visoku efikasnost i u uslovima prisustva outlier-a male varijanse σ_0^2 (Tabela 5.2), dok je RRA algoritam znatno efikasniji od RLS algoritma u slučaju većih vrednosti varijanse outlier-a σ_0^2 , čak i za relativno male vrednosti stepena kontaminacije ε u (5.298) (Tabele 5.1 i 5.2). Sa praktičnog stanovišta je posebno važno da se istakne da je efikasnost RRA algoritma zadovoljavajuće visoka za proizvoljan tip kontaminirajuće f.g.v u (5.298) (videti Tabelu 5.3).

Sa aspekta implementacije RRA algoritma je, takodje, važno da se obezbedi njegova relativna neosetljivost na izbor vrednosti slobodnih parametara u algoritmu, kao što su podešljivi parametar k u izrazu za $\psi(\cdot)$ funkciju (5.88), varijansa σ^2 nominalnog normalno raspodeljenog člana u familiji f.v.g (5.298) i početni uslovi $\theta(0)$ i $\Gamma(0)$ u (5.85) i (5.87), respektivno. U praksi nepoznata varijansa σ^2 nominalnog normalnog poremećaja procenjena je na osnovu robusnog medijana estimatora (5.89), te je ostalo da se pogodnim izborom vrednosti parametra k u (5.88) obezbedi robusna efikasnost, u smislu da RRA algoritam poseduje približno isti kvalitet procena kao RLS algoritam u uslovima čistog normalno raspodeljenog poremećaja, koju zadržava i u situacijama kada je normalno raspodeljeni merni uzorak kontaminiran outlierima, odnosno realizacijama impulsnog šuma, koji su generisani nepoznatom f.g.v $h(\cdot)$ u (5.298). Uticaj parametra k u izrazu za $\psi(\cdot)$ funkciju (5.88) na kvalitet procena parametara (5.85) ilustrovan je rezultatima Monte Karlo simulacije koji su prikazani u Tabeli 5. Na osnovu dobijenih rezultata može se zaključiti da se izborom vrednosti ovog parametra u intervalu (1.4,1.5) postiže zadovoljavajuća efikasnost RRA algoritma, za umerene vrednosti stepena kontaminacije $\varepsilon \leq 0.2$, bez obzira na vrednost varijanse σ_0^2 kontaminirajućeg člana u (5.298). Ovaj rezultat je zadovoljavajući u mnogim primenama, s obzirom da realni merni podaci različite fizičke prirode po pravilu sadrže do 10% outliera [4]. Primetimo da je varijansa šuma generisanog sa (5.298) data sa $(1-\varepsilon)\sigma^2 + \varepsilon \text{var}\{h(\cdot)\}$, te da je $\text{var}\{L(0,1)\} = 2$, $\text{var}\{C(0,1)\} = \infty$, $\text{var}\{N(0,10)\} = 10$, $\text{var}\{R(0,10)\} = 10^2/12$. Sa druge strane, izbor početnih pogađanja $\theta(0)$ i $\Gamma(0)$ može uticati na konvergenijska svojstva RRA algoritma, s obzirom na njegovu nelinearnu prirodu. Takodje, osetljivost RRA algoritma na početne uslove je nešto veća ukoliko se koristi vremenski promenljivi faktor konvergencije (5.300) umesto fiksnog faktora konvergencije (5.299), za umerene vrednosti stepena kontaminacije $\varepsilon \leq 0.2$. Ovaj problem se može prevazići izborom dobrih početnih procena parametara, koje su generisane nekim nerekurzivnim algoritmom, kao što su nerekurzivni metod najmanjih kvadrata ili nerekurzivni metod ponderisanih najmanjih kvadrata, koji su opisani u šestom poglavlju. Inače problem adaptivne estimacije, sa jednovremenom procenom nepoznatih statističkih parametara šuma i parametara modela sistema, biće razmatran u sedmom poglavlju.

Algoritam	ε			
	0	0.02	0.05	0.3
RLS	0.012 2	0.037 6	0.071 2	0.181 1
RRA	0.013 9	0.013 5	0.015 6	0.213

Tabela 5.1: Srednje kvadratna greška estimacije za različite vrednosti statističkog parametra šuma ε ($h = N(0,10)$, $\sigma^2 = 1$, $k = 1.5$, $\hat{\Theta}(0) = 0$, $\Gamma(0) = 0.1I$)

Algoritam	σ_0^2			
	3	6	10	100
RLS	0.017 2	0.050 6	0.123 6	0.8113
RRA	0.015 3	0.015 8	0.016 0	0.0224

Tabela 5.2: Srednje kvadratna greška estimacije za različite vrednosti statističkog parametra šuma σ_0^2 ($h = N(0, \sigma_0^2)$, $\sigma^2 = 1$, $\varepsilon = 0.1$, $k = 1.5$, $\hat{\Theta}(0) = 0$, $\Gamma(0) = 0.1I$)

Algor.	$h(\cdot)$			
	$L(0,1)$	$C(0,1)$	$R(0,10)$	$N(0,10)$
RLS	0.1249	0.9721	0.0317	0.1235
RRA	0.0167	0.0194	0.0131	0.0160

Tabela 5.3: Srednje kvadratna greška estimacije za različite statistike outliera h (L – Laplasova, C – Košijeva, R – uniformna f.g.v; $\sigma^2 = 1$, $\varepsilon = 0.1$, $k = 1.5$, $\hat{\Theta}(0) = 0$, $\Gamma(0) = 0.1I$)

Algoritam	k			
	1	2	3	4
RRA	0.0273	0.0164	0.0189	0.0337

Tabela 5.4: Srednje kvadratna greška estimacije za različite vrednosti nivoa zasićenja k nelinearnosti $\psi(\cdot)$ ($h = N(0,10)$, $\sigma^2 = 1$, $\varepsilon = 0.1$, $\hat{\Theta}(0) = 0$, $\Gamma(0) = 0.1I$)

Da bi se ispitao uticaj korelisanog šuma na performanse RRA algoritma simuliran je ARMAX model sistema (5.185), kod koga su $A(\cdot)$ i $B(\cdot)$ polinomi definisani izrazima (5.295), dok je obojeni šum generisan filtriranjem normalnog belog šuma, nulte srednje vrednosti i jedinične varijanse, kroz polinom (filar)

$$C(q^{-1}) = 1 + 0.2q^{-1} \quad (5.301)$$

Na taj način, regresioni vektor (5.190) dobija oblik

$$Z^T(i) = [-y(i-1) - y(i-2) - y(i-3) - y(i-4) \ u(i-1) \ v(i-1)] \quad (5.302)$$

dok je vektor tačnih vrednosti parametara

$$\Theta^{*T} = [1 \ -0.18 \ 0.78 \ -0.656 \ 2 \ 0.2] \quad (5.303)$$

RRA algoritam sada je definisan izrazima (5.302), (5.303), (5.188)-(5.191), (5.88) i (5.89). Svi ostali uslovi ostali su neizmenjeni u odnosu na prethodni eksperiment. Performanse ovog algoritma poredjene su sa RLS algoritmom, koji koristi linearnu transformaciju reziduala umesto (5.88) i faktor konvergencije $\alpha = d^{-1}$, gde je varijansa nominalnog normalnog šuma u modelu (5.298) izračunata na osnovu (5.89). U RRA algoritmu korišćena je relacija (5.299) ili (5.300) za izračunavanje faktora konvergencije. Rezultati Monte Karlo simulacije prikazani su u Tabeli 5.5. Takođe, implementirana je i modifikacija RRA algoritma sa filtracijom podataka (5.192)-(5.199) sa nelinearnom transformacijom reziduala (5.88), estimacijom nominalne varijanse šuma (5.99) i faktorom konvergencije (5.299) ili (5.300). Ovaj algoritam označen je sa RRAF.

Algor.	$h(\cdot)$			
	$L(0,1)$	$C(0,1)$	$N(0,1)$	$N(0,10)$
RLS	0.1544	1.9151	0.0220	0.1635
RRA	0.0360	0.0484	0.0309	0.0278
RRAF	0.0329	0.0415	0.0312	0.0233

Tabela 5.5: Srednje kvadratna greška estimacije za različite statistike outliera h (L – Laplasova, C – Košijeva, R – uniformna f.g.v; $\sigma^2 = 1$, $\varepsilon = 0.1$, $k = 1.5$, $\hat{\theta}(0) = 0$, $\Gamma(0) = 0.1I$)

Slično kao u prvom eksperimentu, RLS algoritam je dao najbolje rezultate u slučaju čisto gausovskog poremećaja (Tabela 5.5; $h(\cdot) = N(0,1)$ in (5.298)). U slučaju kontaminiranog šuma merenja (Tabela 5.5, $h(\cdot) = L(0,1), C(0,1), N(0,10)$) RLS algoritam generiše pomerene procene parametara zbog korelisanog šuma, dok RRA algoritam zadržava visoku efikasnost, kao i RRAF algoritam. Ovo se može objasniti delovanjem nelinearne transformacije reziduala $\psi(\cdot)$ u relaciji za estimaciju parametara, ali i adekvatnim mehanizmom za generisanje matrice pojačanja algoritma $\Gamma(i)$, gde faktor konvergencije α deluje tako da sopstvene vrednosti ove matrice ne opadnu suviše, čime se obezbeđuje adekvatan pravac pretraživanja u prostoru parametara i relativna neosetljivost na uticaj šuma merenja. Filtracijom mernih podataka poboljša je kvalitet procena parametara, ali je takav algoritam složeniji za implementaciju.

6. OPTIMALNA PARAMETARSKA IDENTIFIKACIJA

U petom poglavlju razmatrana je klasa rekurzivnih stohastičkih procedura Newton-Raphsonovog tipa za parametarsku estimaciju linearnih, vremenski-invarijantnih, diskretnih, dinamičkih stohastičkih sistema. Kao što je pokazano, ovi algoritmi rekurzivno minimiziraju opšti nelinearni empirijski indeks performanse, a analiza konvergencije algoritama data je u petom poglavlju. Sam izbor nelinearne kriterijumske funkcije u indeksu performanse predstavlja slobodan korak i različiti izbori nelinearne transformacije greške predikcije ili reziduala merenja rezultovaće u algoritme parametarske estimacije različitih svojstava. U ovom poglavlju je pokazano da se ova nelinearna funkcija ipak može odrediti preciznije zavisno od raspoložive apriorne informacije o statističkim karakteristikama aditivnog šuma merenja. Kao što je poznato iz literature, ukoliko postoji asimptotski efikasan estimator, u smislu minimalno moguće totalne asimptotske varijanse greške estimacije, odnosno maksimalno moguća asimptotska brzina konvergencije, onda je to estimator maksimalne verodostojnosti, kod koga je nelinearna transformacija greške predikcije (reziduala merenja) definisana funkcijom verodostojnosti, koja predstavlja negativan prirodni logaritam funkcije gustine verovatnoće šuma merenja [49,60,63,89]. Ovakav rezultat sledi direktno na osnovu Cramer-Raove nejednakosti, koja definiše minimalno moguću vrednost totalne asimptotske varijanse greške estimacije za zadate stohastičke karakteristike relevantnih slučajnih procesa [49,60,63,89]. Drugim rečima, estimator maksimalne verodostojnosti dostiže Cramer-Raovu donju među, koja je definisana Fisherovom informacionom matricom. Međutim, u mnogim primenama realna funkcija raspodele verovatnoće šuma nije poznata, te se optimalna kriterijumska funkcija verodostojnosti ne može odrediti. Sledeći metodologiju robusne estimacije [4,33,39,40,72,75,76,78,89], i u navedenim slučajevima može se definisati optimalna kriterijumska funkcija, zavisno od raspoložive apriorne statističke informacije o statističkim karakteristikama šuma merenja. Rešenje ovako postavljenog zadatka sastoji se u sledećem. Na osnovu raspoložive apriorne informacije o statističkim karakteristikama šuma definiše se izvesna klasa raspodele P , kojoj pripada i realna funkcija raspodele verovatnoće šuma. Zatim se određuje najnepovoljnija funkcija gustine verovatnoće u okviru pretpostavljene klase, koja nosi minimum informacije o estimiranim parametrima, odnosno koja maksimizira Cramer-Raovu donju među. Optimalna kriterijumska funkcija tada predstavlja funkciju verodostojnosti, određenu prema najnepovoljnijoj funkciji gustine verovatnoće u okviru date klase. Analizom ukupne asimptotske varijanse greške estimacije može se pokazati da je ovako izabrana kriterijumska funkcija optimalna u smislu mini-maksa ukupne asimptotske varijanse greške estimacije. Drugim rečima, predložena procedura za izbor optimalne kriterijumske funkcije u okviru unapred specificirane klase raspodele poremećaja zasniva se na primeni Cramer-Raove teoreme na izvedeni izraz za asimptotku matricu kovarijanse greške estimacije i maksimizacije Cramer-Raove dalje međe za najnepovoljniju raspodelu u okviru date klase. Tada će ukupna asimptotska varijansa greške estimacije takvog estimatora biti statistički ograničena i neće prelaziti garantovanu granicu robusnosti, koja je minimalno moguća u okviru

pretpostavljene klase raspodele šuma i klase algoritama koji generišu nepomerene procene parametara u usvojenoj reprezentaciji sistema. Ovakav estimator je u suštini estimator maksimalne verodostojnosti za najnepovoljniju funkciju raspodele verovatnoće šuma u okviru specificirane klase, a za svaku drugu raspodelu verovatnoće unutar usvojene klase asimptotska varijansa greške estimacije neće biti veća nego za slučaj najnepovoljnije funkcije raspodele verovatnoće, koja je određena na osnovu minimizacije Fisherove informacione matrice, tj. Cramer-Raove dalje međe.

6.1 IZBOR OPTIMALNOG INDEKSA PERFORMANSE NA BAZI RASPOLOŽIVE APRIORNE INFORMACIJE O STATISTIČKIM KARAKTERISTIKAMA ŠUMA

Kao što je pokazano u četvrtom poglavlju, procena nepoznatog vektora parametara u usvojenoj reprezentaciji modela sistema koja odgovara FIR, ARX ili ARMAX strukturi dobija se kao rešenje odgovarajućeg nelinearnog sistema jednačina, koji se generalno ne može rešiti analitički. U praksi je jedino moguće da se na osnovu konačnog skupa merenja ulaza i izlaza sistema $\{u(j), y(j), 1 \leq j \leq i\}$ generiše, na neki način, procena $\hat{\theta}(i)$ koja bi kad i neograničeno raste konvergirala ka optimalnom rešenju, koje minimizira usvojeni indeks performanse, odnosno ka pravoj vrednosti vektora parametara u razmatranom modelu. Kao kriterijum za ocenu kvaliteta takve procene može se usvojiti asimptotska matrica kovarijanse greške

$$V(F, p) = \lim_{i \rightarrow \infty} i E \left\{ \left(\hat{\theta}(i) - \theta^* \right) \left(\hat{\theta}(i) - \theta^* \right)^T \right\} \quad (6.1)$$

koja karakteriše asimptotsku brzinu konvergencije procene $\hat{\theta}(i)$ ka tačnoj vrednosti parametara θ^* , odnosno minimumu usvojenog kriterijuma. Navedena matrica, kao što je u izrazu (6.1) i naznačeno, generalno zavisi od funkcije gubitaka $F(\cdot)$ u indeksu performanse (4.5) i realne gustine verovatnoće poremećaja $p(\cdot)$ u modelu (3.). Najjednostavniji način generisanja takve procene sastoji se u aproksimaciji indeksa performanse (4.5) aritmetičkom sredinom, odnosno

$$J_i(\theta) = \frac{1}{i} \sum_{j=1}^i F[v(j, \theta)] \quad (6.2)$$

gde je v rezidual ili greška predikcije. U opštem slučaju $v(i, \theta)$ je dato jednačinom linearne regresije za FIR i ARX strukturu modela, $v(i, \theta) = y(i) - \varphi^T(i)\theta$, odnosno jednačinom pseudolinearne regresije $v(i, \theta) = y(i) - \varphi^T(i, \theta)\theta$ za ARMAX modele. Na taj način, procena $\hat{\theta}(i)$ dobija se direktno iz uslova minimuma empirijskog kriterijuma (6.2), tj.

$$\hat{\theta}(i) = \arg \min_{\theta} J_i(\theta) \quad (6.3)$$

Navedena aproksimacija (6.2) omogućava da se približno izračuna asimptotska matrica kovarijanse greške optimalne procene (6.3) i da se funkcija gubitaka $F(\cdot)$ izabere na optimalan način u smislu postizanja maksimalno moguće asimptotske brzine konvergencije. Osnovni rezultat je prikazan sledećom teoremom.

Teorema 6.1. Neka su zadovoljene sledeće pretpostavke:

P1: Funkcija rizika $F(\cdot)$ u indeksu performanse (6.2) je parna, dva puta diferencijalna i konveksna.

P2: Funkcija raspodele poremećaja $\xi(i)$ u modelu zadovoljava pretpostavke $E\{\xi(i)\} = 0$, $E\{\xi(i)\xi(j)\} = \sigma_\xi^2 \delta_{ij}$, $P(-z) + P(z) = 1$, gde je δ_{ij} Kronekerov delta simbol, pri čemu postoji apsolutno neprekidna funkcija gustine verovatnoće $p(z) = dP(z)/dz$.

P3: Matrica

$$B(\theta^*, \sigma_\xi^2) = E\{\nabla_\theta v(i, \theta^*) \nabla_\theta^T v(i, \theta^*)\} = B_1(\theta^*) + \sigma_\xi^2 B_2(\theta^*) \quad (6.4)$$

je pozitivno-definitna, gde je σ_ξ^2 disperzija poremećaja, a $v(i, \theta)$ je greška predikcije ili rezidual merenja.

Tada je asimptotska matrica kovarijanse greške procene (6.3) definisana izrazom

$$V(F, p) = \frac{\alpha}{\beta^2} B^{-1}(\theta^*, \sigma_\xi^2) \quad (6.5)$$

gde je

$$\alpha = \alpha(\psi, p) = E_\xi\{\psi^2(z)\} = \int_{-\infty}^{\infty} \psi^2(z) p(z) dz; \quad \psi(z) = F'(z) \quad (6.6)$$

$$\beta = \beta(\psi, p) = E_\xi\{\psi'(z)\} = \int_{-\infty}^{\infty} \psi'(z) p(z) dz. \quad (6.7)$$

Dokaz: Slično kao u poglavlju 4, razvojem funkcionala (6.2) u Tajlorov red oko tačke $\theta = \theta^*$, zadržavajući se samo na kvadratnom članu, dobija se

$$J_i(\theta) \approx J_i(\theta^*) + \nabla_\theta J_i(\theta^*) (\theta - \theta^*) + \frac{1}{2} (\theta - \theta^*)^T \nabla_\theta^2 J_i(\theta^*) (\theta - \theta^*) \quad (6.8)$$

Na osnovu relacije (6.2) sledi

$$J_i(\theta^*) = \frac{1}{i} \sum_{j=1}^i F(\xi(j)) \quad (6.9)$$

$$\nabla_{\theta} J_i(\theta^*) = \frac{1}{i} \sum_{j=1}^i \psi(\xi(j)) \nabla_{\theta} v(j, \theta^*) \quad (6.10)$$

$$\begin{aligned} \nabla_{\theta}^2 J_i(\theta^*) &= \frac{1}{i} \sum_{j=1}^i \psi'(\xi(j)) \nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*) + \frac{1}{i} \sum_{j=1}^i \psi(\xi(j)) \nabla_{\theta}^2 v(j, \theta^*) \\ &\approx \frac{1}{i} \sum_{j=1}^i \psi'(\xi(j)) \nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*) \end{aligned} \quad (6.11)$$

Prilikom izvođenja relacije (6.11) zanemarena je matrica drugih izvoda $\nabla^2 v$, što je uobičajen postupak uprošćavanja u parametarskoj identifikaciji [28,63,72]. Minimum kriterijuma (6.2) dostiže se pri uslovu $\nabla_{\theta} J_i(\hat{\theta}(i)) = 0$, koji se s obzirom na (6.8) može napisati u obliku

$$\nabla_{\theta} J_i(\theta^*) + \nabla_{\theta}^2 J_i(\theta^*) (\hat{\theta}(i) - \theta^*) = 0 \quad (6.12)$$

odakle se za grešku estimacije dobija

$$\hat{\theta}(i) - \theta^* = -[\nabla^2 J_i(\theta^*)]^{-1} \nabla_{\theta} J_i(\theta^*). \quad (6.13)$$

Pošto na osnovu zakona velikih brojeva [69,71]

$$\nabla_{\theta}^2 J_i(\theta^*) \xrightarrow{i \rightarrow \infty} \nabla_{\theta}^2 J(\theta^*) = E\{\nabla_{\theta}^2 F(v(j, \theta^*))\} \quad (6.14)$$

i uzimajući u obzir da je prema pretpostavci P2 ξ beli šum koji je nekorelisan sa gradijentom reziduala merenja $\nabla_{\theta} v$, na osnovu toga sledi

$$\begin{aligned} \nabla_{\theta}^2 E\{F(v(j, \theta^*))\} &= E\{\nabla_{\theta} [\nabla_{\theta} F(v(j, \theta^*))]\} \\ &= E\{\nabla_{\theta} [\psi(v(j, \theta^*)) \nabla_{\theta} v(j, \theta^*)]\} \\ &= E\{\psi'(v(j, \theta^*)) \nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*)\} + E\{\psi(v(j, \theta^*)) \nabla_{\theta}^2 v(j, \theta^*)\} \end{aligned}$$

ili pošto je na osnovu P1 i P2 $E\{\psi(\xi(j))\} = 0$ dalje se može pisati

$$E\{\psi(v(j, \theta^*)) \nabla_{\theta}^2 v(j, \theta^*)\} = E\{\psi(\xi(j)) \nabla_{\theta}^2 v(j, \theta^*)\} = E\{\psi(\xi(j))\} E\{\nabla_{\theta}^2 v(j, \theta^*)\} = 0$$

odakle se konačno dobija

$$E\{\nabla_{\theta}^2 F(v(j, \theta^*))\} = E\{\psi'(\xi(j))\} E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*)\} \quad (6.15)$$

Na taj način, iz (6.14) i (6.15) se dobija

$$\nabla_{\theta}^2 J_i(\theta^*) \xrightarrow{i \rightarrow \infty} E\{\psi'(\xi(j))\} E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*)\} \quad (6.16)$$

Na osnovu (6.10), (6.13) i (6.16) dalje sledi

$$\hat{\theta}(i) - \theta^* = -\frac{1}{E_{\xi}\{\psi'(z)\}} \left[E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*)\} \right]^{-1} \frac{1}{i} \sum_{j=1}^i \psi(\xi(j)) \nabla_{\theta} v(j, \theta^*) \quad (6.17)$$

Uzimajući u obzir (6.17) i definicioni izraz (6.1), može se takođe pisati

$$\begin{aligned} V(F, p) &= \frac{i}{E_{\xi}^2\{\psi'(z)\}} \left[E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*)\} \right]^{-2} \\ &\cdot E \left\{ \frac{1}{i^2} \sum_{j=1}^i \sum_{k=1}^i \psi(\xi(j)) \psi(\xi(k)) \nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(k, \theta^*) \right\} \end{aligned} \quad (6.18)$$

Takođe, pošto je na osnovu P2 slučajni proces ξ beli šum, to su slučajne veličine $\xi(j)$ i $\xi(k)$ međusobno nekorelisane, a takođe su nekorelisane i sa gradijentom greške predikcije (reziduala merenja), te je

$$\begin{aligned} &E \left\{ \frac{1}{i^2} \sum_{j=1}^i \sum_{k=1}^i \psi(\xi(j)) \psi(\xi(k)) \nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(k, \theta^*) \right\} \\ &= \frac{1}{i^2} \sum_{j=1}^i E\{\psi^2(\xi(j))\} E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*)\} \\ &= \frac{1}{i} E\{\psi^2(\xi(j))\} E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*)\} \end{aligned} \quad (6.19)$$

Prilikom izvođenja (6.19) iskorišćena je činjenica da je:

$$\begin{aligned} &E\{\psi(\xi(j)) \psi(\xi(k)) \nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(k, \theta^*)\} = \\ &E\{\psi(\xi(j)) \psi(\xi(k))\} E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(k, \theta^*)\} = \\ &E\{\psi(\xi(j))\} E\{\psi(\xi(k))\} E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(k, \theta^*)\} = 0 \end{aligned}$$

pošto je po pretpostavci P1 funkcija ψ neparna, a na osnovu P2 funkcija gustine verovatnoće poremećaja ξ je parna funkcija, te je $E\{\psi(\cdot)\} = 0$.

Zamenom (6.19) u (6.18) konačno se dobija

$$V(F, p) = \frac{E_{\xi}\{\psi^2(z)\}}{E_{\xi}^2\{\psi'(z)\}} \left[E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*)\} \right]^{-1} \quad (6.20)$$

Uvodeći oznaku $B(\theta^*, \sigma_\xi^2) = E\{\nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*)\}$ i pišući matricu B u razvijenom obliku, slično kao kod analize ocene najmanjih kvadrata u četvrtom poglavlju, dobija se da je $B = B_1(\theta^*) + \sigma_\xi^2 B_2(\theta^*)$, gde su $B_{1,2}$ pozitivno-definitne simetrične matrice, čime je dokaz teoreme završen. Ukoliko je realna raspodela poremećaja tačno poznata funkcija $F(\cdot)$ u kriterijumu (6.2) može se izabrati na optimalan način u smislu maksimalne brzine konvergencije procene (6.3). Njen optimalan izbor praktično se zasniva na primeni Kramer-Raove teoreme na dobijeni izraz (6.20) za asimptotsku matricu kovarijanse greške estimacije. Naime, ukoliko je procena (6.3) nepomerena, a gustina raspodele poremećaja zadovoljava pretpostavke teoreme 6.1, asimptotska matrica kovarijanse greške (6.20) zadovoljava Kramer-Raovu nejednakost [63,84,89]

$$V(F, p) \geq [I(p)B(\theta^*, \sigma_\xi^2)]^{-1} \quad (6.21)$$

gde je matrica B definisana izrazom (6.4), a $I(p)$ je Fišerova količina informacije.

$$I(p) = E_\xi \left\{ \left(\frac{p'}{p} \right)^2 \right\} = \int_{-\infty}^{\infty} \frac{p'^2(z)}{p(z)} dz \quad (6.22)$$

Ako se za kriterijumsku funkciju usvoji

$$F(z) = F_0(z) = k_1 \log p(z) + k_2 \quad (6.23)$$

skalarni faktori α i β definisani izrazima (6.6) i (6.7), respektivno, postaju $\alpha = \beta = I(p)$ a u relaciji (6.21) važi znak jednakosti, odnosno asimptotska matrica kovarijanse greške procene (6.3) dostiže minimalno moguću vrednost (Kramer-Raovu donju među). Zaista, primenom parcijalne integracije i nejednakosti Koši-Švarc-Bunjakovskog lako se pokazuje da je [8,56,58]

$$\begin{aligned} \beta^2 &= \left[\int_{-\infty}^{\infty} \psi'(z) p_\xi(z) dz \right]^2 = \left[- \int_{-\infty}^{\infty} \psi(z) p_\xi'(z) dz \right]^2 \leq \\ &\leq \left[\int_{-\infty}^{\infty} \psi^2(z) p_\xi(z) dz \right] \cdot \\ &\cdot \left[\int_{-\infty}^{\infty} \frac{p_\xi'^2(z)}{p_\xi(z)} dz \right] = \alpha \int_{-\infty}^{\infty} \frac{p_\xi'^2(z)}{p_\xi(z)} dz \end{aligned}$$

odakle na osnovu (6.22) sledi

$$\frac{\alpha}{\beta^2} = \frac{\int_{-\infty}^{\infty} \psi^2(z) p_\xi(z) dz}{\left[\int_{-\infty}^{\infty} \psi'(z) p_\xi(z) dz \right]^2} \geq \left[\int_{-\infty}^{\infty} \frac{p_\xi'^2(z)}{p_\xi(z)} dz \right]^{-1} = I^{-1}(p_\xi). \quad (6.24)$$

U (6.24) znak jednakosti važi za

$$\psi(z) = F'(z) = -[\log p_\xi(z)], \quad (6.25)$$

odakle direktno sledi (6.23). Prilikom izvođenja izraza (6.23) implicitno je pretpostavljeno da je u Kramer-Raovoj donjoj međi na desnoj strani relacije (6.21) matrica B približno konstantna, tako da ova međa zavisi od funkcije gustine verovatnoće poremećaja p poglavito preko Fisherove količine informacije $I(p)$.

Izraz (6.23) pokazuje da optimalna kriterijumska funkcija nije jedinstvena, već zavisi od parametara $k_{1,2}$. Dobijeni rezultat je posledica činjenice da asimptotska matrica kovarijanse greške (6.5) zavisi i od izvoda kriterijumske funkcije $\psi_0 = F'_0$ i ψ'_0 . Na osnovu jednostavnih geometrijskih razmatranja sleduje da parametar k_1 mora biti negativan. Naime, parametar k_2 određuje minimalnu vrednost funkcije rizika za nultu vrednost argumenta $F_0(0)$. Ako je $k_2 > k_1 \log p(z)$ i $k_1 < 0$ biće $F_0(z) > 0$ za svako z . Analitički je najjednostavnije da se usvoje vrednosti $k_1 = -1$ i $k_2 = 0$. Na taj način optimalna kriterijumska funkcija je funkcija verodostojnosti

$$F_0(z) = -\log p(z), \quad (6.26)$$

odnosno optimalni estimator minimalne totalne varijanse greške estimacije algoritam maksimalne verodostojnosti. Posebno, ukoliko su poremećaji raspodeljeni po Gausu funkcija (6.26) je kvadratna i algoritam maksimalne verodostojnosti svodi se na metod najmanjih kvadrata, koji je analiziran u četvrtom poglavlju.

Međutim, realna gustina raspodele poremećaja je retko potpuno poznata u praksi, ali se često na osnovu raspoložive apriorne informacije o statističkim karakteristikama poremećaja može odrediti izvesna klasa raspodele kojoj i realna raspodela poremećaja pripada. Sledeći metodologiju robusne estimacije, može se definisati optimalna kriterijumska funkcija u okviru date klase, u smislu mini-maksa asimptotske matrice kovarijanse greške (6.5). Pokazuje se da je procena (6.3) u tom slučaju robusna u smislu statističke ograničenosti kovarijacione matrice greške estimacije, koja ne prelazi određenu minimalno moguću granicu robusnosti, s obzirom da se oslanja na Kramer-Raovu donju među. Ovaj rezultat je preciznije formulisan u okviru sledeće teoreme.

Teorema 6.2. Neka su ispunjeni sledeći uslovi

P1: Realna raspodela poremećaja $\xi(i)$ u reprezentaciji modela sistema pripada klasi P konveksnih i simetričnih funkcija raspodele koje poseduje apsolutno neprekidne funkcije gustine verovatnoće $p(z) = dP(z)/dz$.

P2: Postoji najnepovoljnija funkcija gustine verovatnoće u okviru date klase P , koja se dobija kao rešenje minimizacionog problema

$$p_0(z) = \arg \min_{p \in P} I(p) B(\theta^*, \sigma_\xi^2(p)), \quad (6.27)$$

gde je pozitivna definitna matrica B definisano izrazom (6.4), a Fisherova količina informacije $I(p)$ data je sa (6.22), pretpostavljajući da ovaj definicioni integral postoji.

P3: Funkcija gubitaka (cene ili rizika)

$$F_0(z) = -\log p_0(z) \quad (6.28)$$

zadovoljava pretpostavke teoreme 6.1, odnosno predstavlja parnu, konveksnu i dva puta diferencijalnu funkciju.

Tada par (p_0, F_0) zadovoljava uslov sedlaste tačke asimptotske matrice kovarijanse greške (6.5) u okviru usvojene klase raspodele poremećaja P i klase parnih, konveksnih i dva puta diferencijabilnih kriterijumskih funkcija $F(\cdot)$, tj.

$$V(F, p_0) \geq V(F_0, p_0) \geq V(F_0, p). \quad (6.29)$$

Dokaz: Asimptotska matrica kovarijanse greške procene (6.3) definisana je, na osnovu teoreme 6.1, izrazima (6.4) – (6.7). Leva strana dvostruke nejednakosti (6.29) sledi direktno na osnovu Kramer-Raove nejednakosti (6.21), s obzirom da je

$$V(F_0, p_0) = [I(p_0)B(\theta^*, \sigma_\xi^2(p_0))]^{-1} \quad (6.30)$$

odnosno da $V(F, p_0)$ za $F(\cdot) = F_0(\cdot)$ dostiže Kramer-Raovu donju među.

Desna nejednakost u (6.29) može se dokazati koristeći metodologiju izloženu u dokazu teoreme Hubera, gde je razmatran problem estimacije parametara lokacije odnosno ocena skalarnog parametra θ u modelu $y(i) = \theta + \xi(i)$ na osnovu merne sekvence y , gde je ξ beli šum merenja [40]. Naime, pošto je na osnovu pretpostavki teoreme 6.1 matrica B u izrazu (6.4) pozitivno-definitna, umesto desne nejednakosti (6.29) može se posmatrati ekvivalentna nejednakost

$$V^{-1}(F_0, p) \geq V^{-1}(F_0, p_0) \quad (6.31)$$

pri čemu, na osnovu (6.28) $V(F_0, p_0)$ zavisi samo od najnepovoljnije funkcije gustine verovatnoće $p_0(\cdot)$ a $V(F_0, p)$ je funkcija kako najnepovoljnije funkcije gustine verovatnoće $p_0(\cdot)$, tako i realne funkcije gustine verovatnoće poremećaja $p(\cdot)$. Na taj način, imajući u vidu (6.28), izraz (6.31) može se napisati u obliku

$$V^{-1}(p_0, p) \geq V^{-1}(p_0) \quad (6.32)$$

gde je na osnovu (6.5)

$$V(F_0, p) = V(p_0, p) = \frac{\alpha(p_0, p)}{\beta^2(p_0, p)} B^{-1}(\theta^*, \sigma^2(p)) \quad (6.33)$$

gde je

$$\alpha(p_0, p) = E_p \{ \psi_0^2(z) \} = \int_{-\infty}^{\infty} \frac{p_0^2(z)}{p_0^2(z)} p(z) dz; \quad \psi_0(z) = \frac{dF_0(z)}{dz} \quad (6.34)$$

i

$$\beta(p_0, p) = E_p \{ \psi_0'(z) \} = \int_{-\infty}^{\infty} \left[\frac{p_0'(z)}{p_0(z)} \right] p(z) dz = - \int_{-\infty}^{\infty} \frac{p_0'(z)}{p_0(z)} p'(z) dz \quad (6.35)$$

a $V(F_0, p_0) = V(p_0)$ je definisana izrazom (6.30). Poslednji izraz u relaciji (6.35) dobijen je primenom parcijalne integracije.

Posmatrajmo, dalje, funkcionalne $V^{-1}(p_0)$ i $V^{-1}(p_0, p)$ pri varijaciji oko funkcije $p_0(\cdot)$. U tu svrhu zamenimo funkciju $p_0(\cdot)$ funkcijom $p_\varepsilon(\cdot) = (1-\varepsilon)p_0(\cdot) + \varepsilon p(\cdot)$, gde je $0 \leq \varepsilon \leq 1$, a $p(\cdot)$ proizvoljna funkcija gustina verovatnoće u okviru date klase. Na osnovu pretpostavke $p_\varepsilon(\cdot)$ je konveksna i simetrična funkcija, te i ona pripada klasi P . Označimo dalje sa $Q(\varepsilon)$ funkcional $V^{-1}(p_\varepsilon, p_0)$. Tada je, s obzirom na relacije (6.33) – (6.35)

$$Q(\varepsilon) = V^{-1}(p_\varepsilon, p_0) = \frac{\beta^2(p_\varepsilon, p_0)}{\alpha(p_\varepsilon, p_0)} B(\theta^*, \sigma^2(p_0)) \quad (6.36)$$

gde je

$$\alpha(p_\varepsilon, p_0) = \int_{-\infty}^{\infty} \left[\frac{p_0'(z) + \varepsilon(p'(z) - p_0'(z))}{p_0(z) + \varepsilon(p(z) - p_0(z))} \right]^2 p_0(z) dz \quad (6.37)$$

i

$$\beta(p_\varepsilon, p_0) = - \int_{-\infty}^{\infty} \frac{p_0'(z) + \varepsilon[p'(z) - p_0'(z)]}{p_0(z) + \varepsilon[p(z) - p_0(z)]} p'(z) dz \quad (6.38)$$

odnosno

$$\left. \frac{dQ(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = Q'(0) = I^{-1}(p_0) \left[2 \left. \frac{d\beta(p_\varepsilon, p_0)}{d\varepsilon} \right|_{\varepsilon=0} - \left. \frac{d\alpha(p_\varepsilon, p_0)}{d\varepsilon} \right|_{\varepsilon=0} \right] B \quad (6.39)$$

Prilikom izvođenja izraza (6.39) uzeto je u obzir da je na osnovu definicionih obrazaca (6.34) i (6.35) $\alpha(p_0, p_0) = \beta(p_0, p_0) = I(p_0)$. Pošto je, dalje, na osnovu relacija (6.37) i (6.38)

$$2 \left. \frac{d\beta(p_\varepsilon, p_0)}{d\varepsilon} \right|_{\varepsilon=0} = \left. \frac{d\alpha(p_\varepsilon, p_0)}{d\varepsilon} \right|_{\varepsilon=0} \quad (6.40)$$

sledi $Q'(0)=0$. Posmatrajmo, dalje, funkcional $\rho = \beta^2 / \alpha$ i pokažimo da je navedeni funkcional konveksna funkcija argumenta $p(\cdot)$. To će dalje implicirati da je i $Q(\varepsilon)$ konveksna funkcija. Dobijeni rezultat $Q'(0)=0$ tada označava da je $\varepsilon = 0$ stacionarna tačka funkcije $Q(\varepsilon)$, odnosno da je $Q(0) \leq Q(\varepsilon)$. Imajući u vidu da je $Q(0) \leq V^{-1}(p_0, p_0)$ i $Q(\varepsilon) \leq V^{-1}(p_\varepsilon, p_0)$, dalje, sledi $V^{-1}(p_0, p_0) \leq V^{-1}(p, p_0)$ za proizvoljnu funkciju gustine verovatnoće $p(\cdot)$ u okviru date klase P , čime je dokaz teoreme završen. Ostaje još da pokažemo da je pomenuti funkcional $\rho = \beta^2 / \alpha$ konveksna funkcija argumenta $p(\cdot)$. U tu svrhu, uvedimo oznake

$$I(\varepsilon) = I(p_\varepsilon) = \int_{-\infty}^{\infty} \frac{[(1-\varepsilon)p'_0 + \varepsilon p']^2}{(1-\varepsilon)p_0 + \varepsilon p} dz \quad (6.41)$$

gde je $I(\cdot)$ Fisherova količina informacije u (6.22), a posmatrani funkcional definisan je izrazom

$$\rho(\varepsilon) = \rho(p_\varepsilon) = \frac{\beta^2(\psi_0, p_\varepsilon)}{\alpha(\psi_0, p_\varepsilon)} = \frac{\left[(1-\varepsilon) \int_{-\infty}^{\infty} \psi'_0(z) p_0(z) dz + \varepsilon \int_{-\infty}^{\infty} \psi'_0(z) p(z) dz \right]}{(1-\varepsilon) \int_{-\infty}^{\infty} \psi_0^2(z) p_0 dz + \varepsilon \int_{-\infty}^{\infty} \psi_0^2(z) p(z) dz} \quad (6.42)$$

Ako se u (6.42) uvedu sledeće oznake

$$u_1 = \int \psi'_0 p dz, \quad u_2 = \int \psi'_0 p_0 dz, \quad v_1 = \int \psi_0^2 p_1 dz, \quad v_2 = \int \psi_0^2 p_0 dz, \quad (6.43)$$

tada je

$$\begin{aligned} \rho(\varepsilon) &= \frac{[\varepsilon u_1 + (1-\varepsilon)u_2]^2}{\varepsilon v_1 + (1-\varepsilon)v_2} = \\ &= [\varepsilon v_1 + (1-\varepsilon)v_2] \left[\frac{\varepsilon v_1}{\varepsilon v_1 + (1-\varepsilon)v_2} \frac{u_1}{v_1} + \left(1 - \frac{\varepsilon v_1}{\varepsilon v_1 + (1-\varepsilon)v_2} \right) \frac{u_2}{v_2} \right]^2 \leq \\ &\leq [\varepsilon v_1 + (1-\varepsilon)v_2] \left[\frac{\varepsilon v_1}{\varepsilon v_1 + (1-\varepsilon)v_2} \frac{u_1^2}{v_1^2} + \left(1 - \frac{\varepsilon v_1}{\varepsilon v_1 + (1-\varepsilon)v_2} \right) \frac{u_2^2}{v_2^2} \right] = \\ &= \varepsilon \frac{u_1^2}{v_1} + (1-\varepsilon) \frac{u_2^2}{v_2} \end{aligned} \quad (6.44)$$

Na taj način je pokazano, imajući u vidu izraze (6.42) – (6.44), da je

$$\rho(p_\varepsilon) = \rho((1-\varepsilon)p_0 + \varepsilon p) \leq (1-\varepsilon)\rho(p_0) + \varepsilon\rho(p). \quad (6.45)$$

Relacija (6.45) je Jensenova nejednakost i predstavlja potreban i dovoljan uslov konveksnosti funkcije $\rho(\cdot)$ [1,74]. Primetimo da se na isti način može pokazati da je

$$\frac{p_\varepsilon'^2}{p_\varepsilon} = \frac{[(1-\varepsilon)p_0'' + \varepsilon p']^2}{(1-\varepsilon)p_0 + \varepsilon p} \leq (1-\varepsilon)\frac{p_0'^2}{p_0} + \varepsilon\frac{p'^2}{p}, \quad (6.46)$$

odnosno i funkcija $p_\varepsilon'^2 / p_\varepsilon$ je konveksna, pa je i $I(\varepsilon)$ u (6.41), kao integral konveksne funkcije, takođe konveksna funkcija, tj.

$$I(p_\varepsilon) = I((1-\varepsilon)p_0 + \varepsilon p) \leq (1-\varepsilon)I(p_0) + \varepsilon I(p). \quad (6.47)$$

Ovim je dokaz teoreme konačno završen.

Teoreme 6.1 i 6.2 dokazane su pri pretpostavci da je funkcija gubitaka $F(\cdot)$ apsolutno neprekidna. Rezultati navedenih teorema važe i ukoliko je $F(\cdot)$ neprekidna u svim sem konačnom skupu tačaka, a funkcija gustine verovatnoće poremećaja $p(\cdot)$ je apsolutno neprekidna, pri čemu Fisherova količina informacija $I(p)$ u (6.22) postoji i konačna je. Primetimo, takođe, da ukoliko je $F(\cdot)$ konveksna funkcija, ona je i unimodalna, a rešenje θ je jedinstveno. Takođe, diferencijabilna funkcija je konveksna ako je Hessian matrica pozitivno-definitna [1,8,56,58,79].

Najnepovoljnija gustina verovatnoće (6.27) u okviru date klase nosi minimum informacije o estimiranom vektoru parametara, s obzirom da se dobija iz uslova maksimuma desne strane Kramer-Raove nejednakosti (6.21), odnosno maksimalne Kramer-Raove donje međe. Optimalan izbor funkcije rizika (6.28) u okviru pretpostavljene klase garantuje da za proizvoljnu gustinu verovatnoće poremećaja asimptotska kovarijansa greške estimatora na levoj strani (6.21) neće preći granicu robusnosti $V(F_0, p_0)$ u (6.30), koja je minimalna u okviru svih asimptotskih kovarijansi $V(F, p_0)$, gde je $F(\cdot)$ proizvoljna parna funkcija gubitaka.

Određivanje najnepovoljnije gustine verovatnoće u okviru date klase zahteva rešavanje neklasičnog varijacionog problema (6.27). Njegovo rešenje može se naći analitički samo u slučaju FIR modela sistema ($A(q^{-1})=1$, $C(q^{-1})=1$ u ARMAX strukturi). Za slučaj identifikacije parametara modela autoregresije ($B(q^{-1})=0$) pokazano je da rešenje postoji i u okviru klase kontaminiranih normalnih raspodela u specifičnom slučaju skalarnog parametra i belog šuma merenja, tj. $C(z^{-1})=1$ []. U opštem slučaju rešenje postoji samo za klasu raspodela poremećaja sa ograničenom disperzijom, ali njegovo nalaženje zahteva primenu metoda numeričke analize. Kao što je već istaknuto, ako se pretpostavi da je matrica B u (6.27) približno konstantna, tada se razmatrani problem svodi na minimizaciju Fisherove količine informacija $I(p)$ u (6.22), o čemu će više biti reči a u nastavku. Uz navedeno, ukoliko rešenje $p_0(\cdot)$ zadatka minimizacije matričnog kriterijuma (6.27) postoji, ono se poklapa sa rešenjem koje minimizira odgovarajući skalarni kriterijum

$$p_0(z) = \arg \min_{p \in P} T_r \{ I(p) B(\theta^*, \sigma_\varepsilon^2) \} \quad (6.48)$$

gde $T_r \{ \cdot \}$ označava trag matrice. Ako rešenje matričnog problema optimizacije (6.27) ne postoji matrični princip mini-maksne optimalnosti (6.29) u okviru pretpostavljene klase više ne važi. Međutim, i u poslednjem slučaju ostaje u važnosti skalarni princip optimalnosti

$$T_r \{V^{-1}(F, p_0)\} \leq T_r \{V^{-1}(F_0, p_0)\} \leq T_r \{V^{-1}(F_0, p)\} \quad (6.49)$$

gde je optimalna robusna kriterijumska funkcija i dalje definisana relacijom (6.28), ali se sada najnepovoljnija gustina verovatnoće određuje iz uslova (6.48).

Za generisanje robusne procene (6.3) u praksi, koja minimizira optimalni indeks performanse (6.2), (6.28), može se koristiti bilo koja od rekurzivnih stohastičkih procedura predloženih u četvrtom poglavlju. Kao što je pokazano, greške predikcije ili rezidual $v(\cdot)$ je linearna funkcija estimiranog vektora parametara θ samo u slučaju FIR i ARX strukture modela, što rezultuje u relacije $\nabla v(i, \theta) = -\varphi(i)$ i $\nabla^2 v(i, \theta) = 0$, gde je $\varphi(\cdot)$ regresioni vektor. U opštem slučaju korelisanih poremećaja u ARMAX modelu važi (3.), odnosno

$$C(q^{-1})v(i, \theta) = A(q^{-1})y(i) - B(q^{-1})u(i) \Rightarrow v(i, \theta) = y(i) - \varphi^T(i, \theta)\theta \quad (6.50)$$

gde je

$$\begin{aligned} \varphi^T(i, \theta) &= \{-y(i-1) \dots -y(i-n)u(i-1) \dots u(i-n)v(i-1, \theta) \dots v(i-n, \theta)\}, \\ \theta^T &= \{a_1 \dots a_n b_1 \dots b_n, c_1 \dots c_n\} \end{aligned}$$

odakle je očigledno da je rezidual v nelinearna funkcija vektora parametara θ . Ukoliko se, međutim, pretpostavi da su procene bliske pravoj vrednosti nepoznatog vektora parametara θ^* , može se usvojiti aproksimacija $\nabla v(i, \theta) \approx -\varphi(i, \theta)$ i $\nabla^2 v(i, \theta) \approx 0$, pri čemu sada regresioni vektor $\varphi(i, \theta)$ pored pomeraja ulaza i izlaza zavisi od vektora parametara koji se procenjuje. Do istog rezultata može se doći i direktnom linearizacijom izraza (6.50), kao što će biti pokazano u nastavku teksta.

6.2. ASIMPTOTSKA ANALIZA BRZINE KONVERGENCIJE I ODREĐIVANJE OPTIMALNE TEŽINSKE MATRICE KOD REKURZIVNIH STOHAISTIČKIH ALGORITAMA

Kao što je pokazano u četvrtom poglavlju, sintezi robusnih algoritama identifikacije u realnom vremenu (rekurzivnih robusnih stohastičkih procedura) može se prići na bazi opšte formulacije rekurzivnih algoritama tipa stohastičke aproksimacije ili Newton Raphsonove stohastike procedure. Za iterativno pronalaženje korena nelinearnog sistema jednačina u realnom vremenu, koji se dobija iz potrebnog uslova minimuma $\text{grad} J_i(\theta) = 0$ kriterijuma (6.2), može se primeniti

rekurzivni algoritam tipa stohastičke aproksimacije koji je razmatran u četvrtom poglavlju

$$\hat{\theta}(i) = \hat{\theta}(i-1) - \Gamma(i) \nabla_{\theta} v(i, \hat{\theta}(i-1)) \mu_0(v(i, \hat{\theta}(i-1))); \quad \psi_0(z) = F_0'(z) \quad (6.51)$$

gde je rezidual $v(\cdot)$ definisan relacijom (6.50), $\psi_0(\cdot) = F_0'(\cdot)$ relacijom (6.28), a $\Gamma(\cdot)$ je matrica pojačanja, koja utiče na osobinu konvergencije procena. Matrica $\Gamma(\cdot)$ predstavlja slobodan parametar algoritma i njenim pogodnim izborom može se obezbediti ista brzina konvergencije kao i u slučaju nerekurzivnog robusnog estimatora (6.3). Drugim rečima, može se pokazati da je i asimptotska kovarijansa greške procene (6.51) definisana izrazom (6.20), te su oba estimatora asimptotski ekvivalentna i poseduju maksimalno moguću asimptotsku brzinu konvergencije u uslovima nepotpune apriorne informacije o statističkim karakteristikama poremećaja, kada je poznata samo klasa raspodele kojoj pripada stvarna funkcija raspodele verovatnoće poremećaja. Rezultati su sistematizovani u obliku sledeće teoreme. Dobijeni rezultat biće iskorišćen da se izvede izraz za optimalnu matricu pojačanja $\Gamma(\cdot)$ u (6.51), u smislu maksimalno moguće asimptotske brzine konvergencije za zadate uslove stohastičkog okruženja.

Teorema 6.3. Neka su ispunjeni sledeći uslovi:

P1: Realna funkcija raspodele verovatnoće $P(\cdot)$ poremećaja $\xi(i)$ pripada klasi funkcija raspodela koje zadovoljavaju pretpostavke teoreme 6.1, odnosno

$$E\{\xi(i)\} = 0, \quad E\{\xi(i)\xi(j)\} = \sigma_{\xi}^2 \delta_{ij}, \quad P(-z) + P(z) = 1.$$

P2: Kriterijumska funkcija $F(\cdot)$ u indeksu performanse (6.2) data je izrazom (6.28), pri čemu najnepovoljnija funkcija gustine verovatnoće $p_0(\cdot)$ u okviru date klase, u smislu minimuma kriterijuma (6.27), postoji i zadovoljava pretpostavke P1, dok kriterijumska funkcija $F(\cdot)$ zadovoljava pretpostavke teoreme 6.1, odnosno ova funkcija je parna, dva puta diferencijabilna i konveksna.

P3: Parametri rekurzivnog algoritma stohastičke aproksimacije (6.51), koji iterativno u realnom vremenu minimizira kriterijum (6.2), su: nelinearna transformacija reziduala

$$F(z) = F_0(z) = -[\log p_0(z)]; \quad \psi_0(z) = F_0'(z) \quad (6.52)$$

i matrica pojačanja

$$\Gamma(i) = i^{-1} \Gamma_0 \quad (6.53)$$

pri čemu je konstanta, simetrična i pozitivno-definitna matrica Γ_0 data relacijom

$$\Gamma_0^{-1} = \beta(p_0, p) B(\theta^*, \sigma_{\xi}^2) \quad (6.54)$$

gde je $\beta(\cdot)$ definisano izrazom (6.35), a matrica B relacijom (6.4).

Tada asimptotska matrica kovarijanse greške algoritma stohastičke aproksimacije (6.51) zadovoljava uslov sedlaste tačke

$$V(\theta_s, p_0) \geq V(\theta_0, p_0) \geq V(\theta_0, p) \quad \text{za} \quad \forall p \in P \quad (6.55)$$

gde je sa $V(\theta_s, p)$ označena asimptotska matrica kovarijanse greške proizvoljnog estimatora θ_s koji generiše nepomerene procene vektora parametara θ , a θ_0 je procena vektora parametara generisana algoritmom stohastičke aproksimacije (6.51).

Dokaz: Ako se od obe strane relacije (6.51) oduzme prava vrednost vektora parametara θ^* i za grešku estimacije uvede oznaka $\tilde{\theta}(i) = \hat{\theta}(i) - \theta^*$, (6.51) se može napisati u obliku

$$\tilde{\theta}(i) = \tilde{\theta}(i-1) - i^{-1} \Gamma_0 \nabla_{\theta} F_0 \left[v(i, \theta^* + \tilde{\theta}(i-1)) \right]. \quad (6.56)$$

Ukoliko se, dalje, gradijent kriterijumske funkcije $F_0(\cdot)$ razvije u red oko tačke $\theta = \theta^*$, zadržavajući se samo na prvom članu razvoja, dobija se

$$\nabla_{\theta} F_0 \left[v(i, \hat{\theta}(i-1)) \right] \approx \nabla_{\theta} F_0 \left[v(i, \theta^*) \right] + \nabla_{\theta}^2 F_0 \left[v(i, \theta^*) \right] \tilde{\theta}(i-1). \quad (6.57)$$

Zamenom (6.57) u (6.56), posle elementarnog pregrupisanja članova i množenjem obe strane relacije (6.56) sa $\sqrt{i}$, dobija se

$$\sqrt{i} \tilde{\theta}(i) = \left[I - i^{-1} \Gamma_0 \nabla_{\theta}^2 F_0 \left(v(i, \theta^*) \right) \right] \sqrt{\frac{i}{i-1}} \sqrt{i-1} \tilde{\theta}(i-1) - i^{-1} \sqrt{i} \Gamma_0 \nabla_{\theta} F_0 \left(v(i, \theta^*) \right) \quad (6.58)$$

gde je I jedinična matrica odgovarajućih dimenzija. Ako se, dalje, na osnovu razvoja u Maklerenov red usvoji aproksimacija $\sqrt{i/i-1} \approx 1 + 1/(2i)$, uvede oznaka

$$A(i, \theta^*) = \frac{I}{2} - \Gamma_0 \nabla_{\theta}^2 F_0 \left(v(i, \theta^*) \right) \quad (6.59)$$

i zanemare članovi uz $1/i^2$, pretpostavljajući da je i dovoljno veliko, relacija (6.58) dobija oblik

$$\sqrt{i} \tilde{\theta}(i) \approx \sqrt{i-1} \tilde{\theta}(i-1) + \frac{1}{i} A(i, \theta^*) \sqrt{i-1} \tilde{\theta}(i-1) - \frac{\sqrt{i}}{i} \Gamma_0 \nabla_{\theta} F_0 \left(v(i, \theta^*) \right). \quad (6.60)$$

Nalaženjem transponovane vrednosti $\sqrt{i} \tilde{\theta}^T(i)$ i množenjem relacije (6.60) sa dobijenim izrazom za $\sqrt{i} \tilde{\theta}^T(i)$, zanemarujući ponovo članove uz $1/i^2$ za dovoljno veliko i , sledi

$$i \tilde{\theta}(i) \tilde{\theta}^T(i) \approx (i-1) \tilde{\theta}(i-1) \tilde{\theta}^T(i-1) + \frac{1}{i} A(i, \theta^*) (i-1) \tilde{\theta}(i-1) \tilde{\theta}^T(i-1) -$$

$$\begin{aligned}
 & -\frac{1}{i}\sqrt{i(i-1)}\Gamma_0\nabla_{\theta}F_0(v(i,\theta^*))\tilde{\theta}^T(i-1)+\frac{1}{i}(i-1)\tilde{\theta}(i-1)\tilde{\theta}^T(i-1)A^T(i,\tilde{\theta}^*)- \\
 & -\frac{1}{i}\sqrt{i(i-1)}\tilde{\theta}(i-1)\nabla_{\theta}^TF_0(v(i,\theta^*))\Gamma_0+\frac{1}{i}\Gamma_0\nabla_{\theta}F_0(v(i,\theta^*))\nabla_{\theta}^TF_0(v(i,\theta^*))\Gamma_0 \quad (6.61)
 \end{aligned}$$

Važno je uočiti da drugi sabirak na desnoj strani izraza (6.61) predstavlja transponovanu vrednost četvrtog sabirka, a treći petog, te je dovoljno odrediti samo matematičko očekivanje drugog i trećeg člana. Pošto je na osnovu usvojene ARMAX reprezentacije modela sistema

$$\nabla_{\theta}F_0(v(i,\theta^*))=-\psi_0(\xi(i))\varphi(i,\theta^*) \quad (6.62)$$

to je

$$E\{\nabla_{\theta}F_0(v(i,\theta^*))\tilde{\theta}^T(i-1)\}=E_{\xi}\{\psi_0(\xi(i))\}E\{\varphi(i,\theta^*)\tilde{\theta}^T(i-1)\}=0 \quad (6.63)$$

Prilikom izvođenja (6.63) uzeto je u obzir da je slučajna veličina $\xi(i)$ nezavisna od slučajnih vektora $\varphi(i,\theta^*)$ i $\tilde{\theta}(i-1)$, pri čemu je $\psi_0(\cdot)$ neparna funkcija, dok je funkcija gustine verovatnoće $p(\cdot)$ šuma $\xi(\cdot)$ parna funkcija. Na taj način, matematička očekivanja trećeg i petog sabirka u (6.61) je nula.

Imajući u vidu da je

$$\nabla_{\theta}^2F_0(v(i,\theta^*))\approx-\psi_0'(\xi(i))\varphi(i,\theta^*)\varphi^T(i,\theta^*) \quad (6.64)$$

i pošto je na osnovu (6.51)

$$\tilde{\theta}(i)=\tilde{\theta}(0)+\sum_{j=1}^i\Gamma(j)\varphi(j,\theta^*)\psi_0(v(j,\tilde{\theta}(j-1))) \quad (6.65)$$

zaključuje se da su $\nabla_{\theta}^2F_0(v(i,\theta^*))$ i $\tilde{\theta}(i-1)$ statistički nezavisne slučajne veličine (zbog opadajuće prirode matrice $\Gamma(i)$ može se zanemariti uticaj φ na $\tilde{\theta}$ za dovoljno veliko i), što dalje implicira i statističku nezavisnost slučajnih veličina $A(i,\theta^*)$ i $\tilde{\theta}(i-1)$, pa je

$$E\{A(i,\theta^*)\tilde{\theta}(i-1)\tilde{\theta}^T(i-1)\}=AE\{\tilde{\theta}(i-1)\tilde{\theta}^T(i-1)\} \quad (6.67)$$

gde je prema (6.59)

$$A=E\{A(i,\theta^*)\}=\frac{I}{2}-\Gamma_0E\{\nabla_{\theta}^2F_0(v(i,\theta^*))\}. \quad (6.68)$$

S obzirom na (6.50), (6.62) i statističku nezavisnost slučajnih veličina $\xi(i)$ i $\varphi(i,\theta^*)$, takođe sledi

$$E\{\nabla_{\theta} F_0(v(i, \theta^*)) \nabla_{\theta}^T F_0(v(i, \theta^*))\} = \alpha(p_0, p) B(\theta^*, \sigma^2(p)) \quad (6.69)$$

gde je $\alpha(p_0, p)$ dato izrazom (6.34).

Na osnovu relacija (6.61), (6.63), (6.67), (6.68) i (6.69), dalje, sledi

$$V(i) = V(i-1) + \frac{1}{i} [AV(i-1) + V(i-1)A^T + \Gamma_0 B(\theta^*, \sigma^2(p)) \Gamma_0 \alpha(p_0, p)] \quad (6.70)$$

gde je sa $V(i)$ označena kovarijaciona matrica slučajnog vektora $\sqrt{i} \tilde{\theta}^T(i)$ u (6.1), tj.

$$V(i) = E\{\tilde{\theta}(i) \tilde{\theta}^T(i)\} \quad (6.71)$$

Izraz (6.70) iterativno definiše matricu kovarijanse greške estimacije nepomerene procene (6.51). U stacionarnom stanju kad i neograničeno raste $V(i-1) = V(i) = V$, gde je sa V označena asimptotska matrica kovarijanse greške procene parametara (6.51). S obzirom na (6.70), matrica V mora u stacionarnom stanju zadovoljavati matričnu jednačinu Ljapunova [49,60,63,89]

$$AV + VA^T = -\alpha(p_0, p) \Gamma_0 B(\theta^*, \sigma^2(p)) \Gamma_0 \quad (6.72)$$

Imajući u vidu (6.64) i (6.52) – (6.54), takođe se dobija

$$E\{\nabla^2 F_0(v(i, \theta^*))\} = \beta(p_0, p) B(\theta^*, \sigma^2(p)) = \Gamma_0^{-1}. \quad (6.73)$$

Pošto je prema (6.73) i (6.68) $A = -I/2$, iz (6.72) sledi da je asimptotska matrica kovarijanse greške $V = V(\theta_0, p)$, $\theta_0 = \lim_{i \rightarrow \infty} \hat{\theta}(i)$ estimatora stohastičke aproksimacije (6.51) data izrazom (6.5).

Primetimo da leva strana dvostruke nejednakosti (6.55) sledi direktno na osnovu Kramer-Raove nejednakosti (6.21), a desna na osnovu izvedenog izraza za asimptotsku matricu kovarijanse greške (6.5) i rezultata teoreme 6.1, čime je dokaz teoreme završen.

U širem kontekstu relacija (6.72) definiše približno asimptotsku matricu kovarijanse greške algoritma stohastičke aproksimacije (6.51) za proizvoljan izbor nelinearne transformacije reziduala $\psi_0(\cdot)$, pod uslovom da je neparna funkcija, i matričnog koeficijenta pojačanja u obliku $\Gamma(i) = i\Gamma_0$, gde je Γ_0 proizvoljna, ograničena, simetrična i pozitivno-definitna matrica. Posebno, ukoliko se pretpostavljena klasa raspodela sastoji samo od jednog elementa [21,36,63,82,89] tada je $p_0 = p$ i robusni algoritam (6.51) je rekurzivan algoritam maksimalne verodostojnosti. S obzirom da je u navedenom slučaju $\alpha(p_0, p) = \beta(p_0, p) = I(p_0)$ asimptotska matrica kovarijanse greške algoritma maksimalne verodostojnosti definisana je izrazom (6.30). Rezultati teoreme 6.3 takođe pokazuju da robusni algoritam (6.51), (6.52) – (6.54) poseduje minimalno moguću asimptotsku

kovarijacionu matricu greške, odnosno maksimalno moguću asimptotsku brzinu konvergencije, ukoliko je realna raspodela poremećaja najnepovoljnija raspodela u datoj klasi. Za svaku drugu raspodelu u okviru pretpostavljene klase asimptotska matrica kovarijanse greške biće manja nego u slučaju najnepovoljnije raspodele i neće preći granicu robusnosti (6.30), koja je minimalno moguća, s obzirom da dostiže Kramer-Raovu donju među.

Realizacija rekurzivnog robusnog algoritma identifikacije (6.51) – (6.54), optimalnog u smislu mini-maksa asimptotske matrice kovarijanse greške (6.55), zahteva da se vektor $\nabla_v(i, \theta^*) = -\varphi(i, \theta^*)$ aproksimira izrazom $-\varphi(i, \hat{\theta})$, gde je $\hat{\theta}$ procena nepoznatog vektora parametara θ^* . Međutim, i u tom slučaju algoritam se ne može realizovati u praksi, s obzirom da matični koeficijent pojačanja (6.53) pretpostavlja poznavanje realne raspodele poremećaja i prave vrednosti nepoznatog vektora parametara θ^* koji se identifikuje. Jedna od mogućnosti da se prevaziđu navedene teškoće sastoji se u izračunavanju skalarnog faktora (6.35) prema najnepovoljnijoj gustini verovatnoće poremećaja u okviru pretpostavljene klase, dobijenoj rešavanjem varijacionog zadatka (6.27), i aproksimaciji matrice (6.4) aritmetičkom sredinom, tj.

$$\beta(p_0, p) \approx \beta(p_0) = I(p_0) = \int_{-\infty}^{\infty} \frac{p_0'(z)}{p_0(z)} dz \quad (6.74)$$

$$B(\theta^*, \sigma^2(p)) \approx \frac{1}{i} \sum_{j=1}^i \varphi(j, \hat{\theta}(j-1)) \varphi^T(j, \hat{\theta}(j-1)) \quad (6.75)$$

gde je vektor pseudolinearne regresije $\varphi(\cdot)$, pretpostavljajući da su procene parametara $\hat{\theta}$ bliske njihovoj tačnoj vrednosti θ^* , definisan izrazom (6.50). Zamenom (6.74), (6.75) u (6.53), za suboptimalnu matricu pojačanja algoritma dobija se izraz

$$\Gamma_0(i) = (I(p_0)) \sum_{j=1}^i \varphi(j, \hat{\theta}(j-1)) \varphi^T(j, \hat{\theta}(j-1))^{-1} \quad (6.76)$$

ili, nakon primene leme o inverziji matrica, dobija se rekurzivna varijanta izraza (6.76)

$$\Gamma_0(i) = \Gamma_0(i-1) - \frac{\Gamma_0(i-1) \varphi(i, \hat{\theta}(i-1)) \varphi^T(i, \hat{\theta}(i-1)) \Gamma_0(i-1) I(p_0)}{1 + \varphi^T(i, \hat{\theta}(i-1)) \Gamma_0(i-1) \varphi(i, \hat{\theta}(i-1)) I(p_0)} \quad (6.77)$$

Relacije (6.50), (6.51), (6.52) i (6.77) definišu suboptimalni rekurzivni robusni algoritam identifikacije u realnom vremenu konstantnog vektora parametara u ARMAX reprezentaciji modela sistema, koji je već razmatran u četvrtom poglavlju.

Matrična relacija (6.72) omogućava da se približno proceni asimptotska brzina konvergencije suboptimalnog algoritma. Naime, pošto na osnovu zakona velikih brojeva, ukoliko $\hat{\theta}(i)$ konvergira ka θ^* kod i neograničeno raste, matrični koeficijent

pojačanja (6.76) suoptimalnog algoritma konvergira ka $\Gamma_0 = [I(p_0)B(\theta^*, \sigma^2(p))]^{-1}$, to je na osnovu (6.68) i (6.73) matrica A u izrazu (6.72)

$$A = \frac{I(p_0) - 2\beta(p_0, p)}{2I(p_0)} I \quad (6.78)$$

gde je I jedinična matrica odgovarajućih dimenzija. Rešavajući matricnu jednačinu Ljapunova (6.72) za asimptotsku matricu kovarijanse greške razmatrane suboptimalne procene dobija se izraz

$$V_s(\psi_0, p) = \frac{\alpha(p_0, p)}{2\beta(p_0, p)I(p_0) - I^2(p_0)} B^{-1}(\theta^*, \sigma^2(p)). \quad (6.79)$$

S obzirom da je $[\beta(p_0, p) - I(p_0)]^2 \geq 0$, dalje sledi

$$\frac{1}{2\beta(p_0, p)I(p_0) - I^2(p_0)} \leq \frac{1}{\beta^2(p_0, p)} \quad (6.80)$$

odakle se dobija

$$V_s(\psi_0, p) \leq \frac{\alpha(p_0, p)}{\beta^2(p_0, p)} B^{-1}(\theta^*, \sigma^2(p)) = V(\psi_0, p) \quad (6.81)$$

gde je sa $V(\psi_0, p)$ označena asimptotska matrica kovarijanse greške optimalnog min-maksnog robusnog algoritma (6.51) – (6.54). Imajući u vidu da je $V_s(\psi_0, p_0) = V(\psi_0, p_0)$ i rezultate teoreme 6.3, sledi da i suboptimalni algoritam približno zadovoljava uslov sedlaste tačke (6.55). Za početnu matricu pojačanja algoritma $\Gamma(0)$, može se usvojiti $\Gamma(0) = a^2 I$, gde je a pogodno odabran broj, a I jedinična matrica odgovarajućih dimenzija.

Posebno, ukoliko se i skalarni faktor $I(p_0)$ u izrazu (6.77) aproksimira aritmetičkom sredinom

$$I(p_0) \approx \frac{1}{i} \sum_{j=1}^i \psi_0'(\nu(j, \theta(j-1))) \quad (6.82)$$

i pretpostavi da su procene parametara bliske njihovoj pravoj vrednosti θ^* , tada je $I(p_0) \approx \psi_0'(\nu(i, \theta(i-1)))$ i izraz (6.77) se svodi na oblik

$$\Gamma_0(i) = \Gamma_0(i-1) - \frac{\Gamma_0(i-1)\varphi(i, \hat{\theta}(i-1))\varphi^T(i, \hat{\theta}(i-1))\Gamma_0(i-1)\psi_0'(\nu(i, \hat{\theta}(i-1)))}{1 + \varphi^T(i, \hat{\theta}(i-1))\Gamma_0(i-1)\varphi(i, \hat{\theta}(i-1))\psi_0'(\nu(i, \theta(i-1)))}. \quad (6.83)$$

Suboptimalni robusni algoritam, definisan relacijama (6.50), (6.51), (6.52) i (6.83), takođe je već razmatran u četvrtom poglavlju. U slučaju da je osnovna reprezentacija modela sistema definisana ARX umesto ARMAX strukturom, teorema 6.3 može se

formulisati u strožijem obliku, odnosno može se pokazati da u navedenom slučaju greška estimacije rekurzivne stohastičke procedure (6.50)-(6.54) ima asimptotski normalnu raspodelu, nulte srednje vrednosti i matrice kovarijanse koja je definisana izrazom (6.20), koji se za najnepovoljniju funkciju gustine verovatnoće poremećaja, u smislu minimuma Fisherove informacione matrice (Cramer-Raove donje međe), svodi na izraz (6.30). Ovakav rezultat direktna je posledica sledeće dve leme, a formulisan je i dokazan u okviru teoreme 6.4.

Lema 6.1 [44]: Neka je $\{b_n\}$ niz skalarnih veličina takvih da je $b_n > 0$, $\lim_{n \rightarrow \infty} b_n = 0$ i $nb_n < \left(\sum_{k=1}^{n-1} b_k^2 + n^{\alpha_1} \right)$ za neko $c > 0$ i $\alpha_1 \in (0,1)$. Tada je

$$\sum_{k=1}^{n-1} b_k^2 < cn^\beta; \quad \beta = \begin{cases} \max & (0, 2\alpha_1 - 1) & \text{za } \alpha_1 \neq 1/2 \\ \varepsilon & (\varepsilon > 0) & \text{za } \alpha_1 = 1/2 \end{cases}$$

Neka je takođe $\{L(n), F_n\}$ L^2 ograničen martingal [1]. Tada teorema Doob-a (1953) garantuje postojanje veličine L takve da $L(n)$ konvergira ka L sa verovatnoćom 1 (s.v.1) i po kriterijumu L^2 [28]. Na taj način, može se pisati

$$L(n) = \sum_{i=1}^n X(i)$$

gde je

$$X(i) = L(i) - L(i-1), \quad L(\cdot) = 0$$

Drugim rečima, niz $\sum_i X(i)$ konvergiće sa verovatnoćom 1 (s.v.1) i po kriterijumu L^2 ka graničnoj vrednosti L , pri čemu je diferencija

$$L - L(n) = \sum_{i=n+1}^{\infty} X(i)$$

Sada se može formulisati sledeća lema:

Lema 6.2 [28]: Neka je $\{x(t)\}$ skalarni stohastički proces, pri čemu je $V_n^2 = \sum_{t=n+1}^{\infty} E\{x^2(t) | F_{t-1}\}; \quad S_n^2 = E\{V_n^2\} = E\{[L - L(n)]^2\}$ i zadovoljene su sledeće pretpostavke:

$$H1) \quad P\left\{\left|\frac{V_n^2}{S_n^2}\right| > \varepsilon\right\} < \delta \quad \text{za neko } \varepsilon > 0, \text{ tj., } \frac{V_n^2}{S_n^2} \xrightarrow[n \rightarrow \infty]{} 1 \quad \text{s.v.1.}$$

$$H2) \quad \lim_{n \rightarrow \infty} \frac{1}{S_n^2} \sum_{t=n+1}^{\infty} E\{x^2(t) I(|x(t)| \geq \varepsilon S_n)\} = 0$$

gde je I indikatorska funkcija. Tada je u važnosti:

$$\frac{V_n}{S_n} = \frac{L - L(n)}{S_n} \xrightarrow{D} N(0,1)$$

gde znak $\xrightarrow{D}$ označava konvergenciju po raspodeli, dok $N(0,1)$ predstavlja standardnu normalnu raspodelu, nulte srednje vrednosti i jedinične varijanse.

Sada se može formulisati sledeća teorema:

Teorema 6.4. Posmatrajmo ARX model sistema i algoritam (6.50), (6.51) i (6.77), za koji su zadovoljene sledeće pretpostavke:

- A1) Sve nule polinoma $A(q^{-1})$ nalaze se unutar jediničnog kruga.
- A2) $\{\xi(i)\}$ je niz nezavisnih i identički raspodeljenih slučajnih varijabli, koje imaju simetričnu funkciju raspodele i zadovoljavaju uslove $E\{\xi(i)\} = 0$, $E\{\xi^2(i)\} = \sigma_\xi^2 < \infty$.
- A3) $E\{Z(i)Z^T(i)\}$ je pozitivno definitna matrica, gde je $Z(i) = \varphi(i, \hat{\theta}(i-1))$.
- A4) Nelinearna funkcija $\psi(\cdot)$ u (6.51) je neparna i ograničena.
- A5) Prvi i drugi izvod funkcije ψ su ograničeni.
- A6) Procena $\hat{\theta}(i)$ u (6.51) je konzistentna, tj.

$$P\left\{\lim_{i \rightarrow \infty} \hat{\theta}(i) = \theta^*\right\} = 1.$$

Tada $\sqrt{i}[\hat{\theta}(i) - \theta^*] \xrightarrow{D} N(0, P)$, gde je $N(0, P)$ normalna raspodela, nulte srednje vrednosti i matrice kovarijanse

$$P = \frac{E\{\psi^2[e(i)]\}}{E^2\{\psi[e(i)]\}} [E\{Z(i)Z^T(i)\}]^{-1}. \quad (6.84)$$

Dokaz. Označimo sa $\hat{\theta}(i) = \hat{\theta}(i) - \theta^*$, i napišimo (6.51) i (6.57) u obliku

$$\hat{\theta}(i) = \hat{\theta}(i-1) + \bar{R}^{-1}(i)Z(i)\psi[v(i, \hat{\theta}(i-1))] \quad (6.85)$$

$$\bar{R}(i) = \bar{R}(i-1) + \alpha Z(i)Z^T(i) \quad (6.86)$$

gde je

$$\Gamma(i) = \bar{R}^{-1}(i) \quad (6.87)$$

Na osnovu (6.85) i (6.86) dalje sledi

$$\bar{R}(i)\tilde{\theta}(i) = \bar{R}(i-1)\tilde{\theta}(i-1) + \alpha Z(i)Z^T(i)\tilde{\theta}(i-1) + Z(i)\psi[v(i, \hat{\theta}(i-1))] \quad (6.88)$$

Sumirajući (6.88) od 0 do i , može se pisati

$$\bar{R}(i)\tilde{\theta}(i) = \bar{R}(0)\tilde{\theta}(0) + \sum_{k=1}^i Z(k)\psi[\xi(k)] + \sum_{k=1}^i Z(k)\{\alpha Z^T(k)\tilde{\theta}(k-1) + \psi[v(k, \hat{\theta}(k-1))] - \psi[\xi(k)]\} \quad (6.89)$$

Takođe, uvodeći smenu $\bar{R}(i) = iR(i)$, na osnovu (6.89) sledi

$$\begin{aligned} \sqrt{i}\hat{\theta}(i) &= \frac{1}{\sqrt{i}}R^{-1}(i)\bar{R}(0)\tilde{\theta}(0) + \frac{1}{\sqrt{i}}R^{-1}(i)\sum_{k=1}^i Z(k)\psi[\xi(k)] + \\ &+ \frac{1}{\sqrt{i}}R^{-1}(i)\sum_{k=1}^i Z(k)\{\alpha Z^T(k)\tilde{\theta}(k-1) + \psi[v(k, \hat{\theta}(k-1))] - \psi[\xi(k)]\} \end{aligned} \quad (6.90)$$

Pošto uticaj početnog pogađanja $\hat{\theta}(0)$ na procene vektora parametra θ^* iščezava sa vremenom dovoljno je da posmatramo poslednja dva člana na desnoj strani (6.90). Da bismo analizirali uticaj trećeg člana, posmatrajmo izraz

$$\begin{aligned} &\|Z(k)\{\alpha Z^T(k)\tilde{\theta}(k-1) + \psi[v(k, \hat{\theta}(k-1))] - \psi[\xi(k)]\}\| \leq \\ &\|Z(k)\|\|\alpha Z^T(k)\tilde{\theta}(k-1) + \psi[v(k, \hat{\theta}(k-1))] - \psi[\xi(k)]\| \end{aligned} \quad (6.91)$$

gde $\|\cdot\|$ označava Euklidovu normu. Linearizujući $\psi[\xi(k, \theta)]$ u okolini θ^* i uzimajući $\theta = \hat{\theta}(k-1)$, dobija se

$$\begin{aligned} &|\psi[v(k, \hat{\theta}(k-1))] - \psi[\xi(k)] + \psi'[\xi(k)]Z^T(k)\tilde{\theta}(k-1)| = \\ &= \frac{1}{2}|\psi''[\xi(k)]\{Z^T(k)\tilde{\theta}(k-1)\}^2|. \end{aligned} \quad (6.92)$$

Na osnovu pretpostavke A5, zaključuje se da je

$$|\psi[v(k, \hat{\theta}(k-1))] - \psi[\xi(k)] + \psi'[\xi(k)]Z^T(k)\tilde{\theta}(k-1)| \leq k_2[Z^T(k)\tilde{\theta}(k-1)]^2 \quad (6.93)$$

za neko $k_{1,2} \in (0, \infty)$. Na osnovu pretpostavki A1, A2, kao i relacija (6.91) – (6.93), dalje sledi

$$\|Z(k)\{\alpha Z^T(k)\tilde{\theta}(k-1) + \psi[v(k, \hat{\theta}(k-1))] - \psi[\xi(k)]\}\| \leq k_3\|\tilde{\theta}(k-1)\|^2 \quad (6.94)$$

gde je k_3 konačna i pozitivna konstanta.

Analizirajmo sada veličinu

$$S(i) = \sum_{k=1}^i k^{-\frac{1}{2}-\delta} Z(k) \psi[e(k)] \quad (6.95)$$

Pošto je na osnovu A2 i A4 $E\{\psi[e(i)]\} = 0$, zaključuje se da je

$$\begin{aligned} E\{S(i) | F_{i-1}\} &= E\left\{ \sum_{k=1}^{i-1} k^{-\frac{1}{2}-\delta} Z(k) \psi[\xi(k)] | F_{i-1} \right\} + \\ &+ i^{-\frac{1}{2}-\delta} Z(i) E\{\psi[\xi(i)]\} = S(i-1). \end{aligned} \quad (6.96)$$

Na osnovu (6.95) takođe sledi

$$\|S(i)\|^2 = \sum_{k=1}^i \left(k^{-\frac{1}{2}-\delta} \right)^2 \|Z(k)\|^2 \|\psi[\xi(k)]\|^2 \quad (6.97)$$

Primenjujući operator matematičkog očekivanja $E\{\cdot\}$ na (6.97) i uzimajući u obzir pretpostavke A1, A2 i A4, na osnovu kojih sledi da su slučajne veličine $Z(k)$ i $\psi[\xi(k)]$ nekorelisane i ograničene, može se pisati

$$E\{\|S(i)\|^2\} = \sum_{k=1}^i \left(k^{-\frac{1}{2}-\delta} \right)^2 E\{\|Z(k)\|^2\} E\{\|\psi[\xi(k)]\|^2\} \leq k_4 \sum_{k=1}^{\infty} k^{-1-2\delta} < \infty \quad (6.98)$$

za neko $k_4 \in (0, \infty)$. Zadnja nejednakost u (6.98) sledi na osnovu činjenice da $\sum_{k=1}^{\infty} k^{-1-2\delta}$ konvergira [44]. Na osnovu (6.98) i (6.96) zaključuje se da je slučajni proces $\{S(i)\}$ martingal, tako da se može primeniti teorema Neveu-a (videti poglavlje 4), na osnovu koje je [70]

$$\lim_{i \rightarrow \infty} S(i) = S_{\infty} < \infty \quad \text{s.v.1} \quad (6.99)$$

Na osnovu Kronecker-ove leme, koja je takođe izložena u četvrtom poglavlju, zaključuje se da je

$$\lim_{i \rightarrow \infty} \frac{1}{i^{\frac{1}{2}+\delta}} \sum_{k=1}^i Z(k) \psi[\xi(k)] = 0 \quad \text{s.v.1} \quad (6.100)$$

Primitimo, takođe, da se na osnovu (6.89) može pisati

$$i \tilde{\theta}(i) = R^{-1}(i) R(0) \tilde{\theta}(0) + i^{\frac{1}{2}+\delta} \frac{1}{i^{\frac{1}{2}+\delta}} R^{-1}(i) \sum_{k=1}^i Z(k) \psi[\xi(k)] +$$

$$+ R^{-1}(i) \sum_{k=1}^i Z(k) \left\{ \alpha Z^T(k, \tilde{\theta}(k-1)) + \psi \left[v(k, \hat{\theta}(k-1)) \right] - \psi[\xi(k)] \right\} \quad (6.101)$$

Na taj način, primenjujući pretpostavku A6 i relacije (6.94), (6.100) i (6.101), dobija se

$$\begin{aligned} i \|\tilde{\theta}(i)\| &\leq \|R^{-1}(i)\| \|R(0)\| \|\tilde{\theta}(0)\| + k_5 i^{\frac{1}{2}+\delta} \|R^{-1}(i)\| + \\ &k_3 \|R^{-1}(i)\| \sum_{k=1}^i \|\tilde{\theta}(k-1)\|^2 \leq k_6 + k_7 i^{\frac{1}{2}+\delta} + k_8 \sum_{k=1}^i \|\tilde{\theta}(k-1)\|^2 \leq \\ &k_9 \left(i^{\frac{1}{2}+\delta} + \sum_{k=1}^i \|\tilde{\theta}(k-1)\|^2 \right) \end{aligned} \quad (6.102)$$

za neke konstante $k_5 - k_9 \in (0, \infty)$.

Ako se uvede smena $b_i = \|\tilde{\theta}(i)\|$ i $\alpha_1 = 1/2 + \delta$, na izraz (6.102) može se primeniti lema 6.1, na osnovu čega sledi

$$\sum_{k=1}^i \|\tilde{\theta}(k-1)\|^2 < k_9 i^{2\delta} \quad (6.103)$$

Na taj način, treći član na desnoj strani izraza (6.90) zadovoljava uslov

$$\begin{aligned} &\left\| R^{-1}(i) \frac{1}{\sqrt{i}} \sum_{k=1}^i Z(k) \left\{ \alpha Z^T(k) \tilde{\theta}(k-1) + \psi \left[v(k, \hat{\theta}(k-1)) \right] - \psi[\xi(k)] \right\} \right\| \leq \\ &\leq \|R^{-1}(i)\| \frac{1}{\sqrt{i}} k_3 k_9 i^{2\delta} \leq k_{10} \frac{1}{i^{\frac{1}{2}-2\delta}} \xrightarrow{i \rightarrow \infty} 0 \end{aligned} \quad (6.104)$$

gde je k_{10} konačna i pozitivna konstanta. Dobijeni rezultat u (6.104) pokazuje da asimptotska raspodela greške estimacije zavisi samo od drugog člana na desnoj strani izraza (6.90). Da bismo izveli potreban izazr, primenimo lemu 6.2 na svaki element vektorske veličine

$$L(i) = \frac{1}{\sqrt{i}} \sum_{k=1}^i Z(k) \psi[\xi(k)] \quad (6.105)$$

kao i uvedimo oznake

$$V_n^2 = \sum_{k=1}^n E \left\{ [Z(k) \psi(\xi(k))]^2 | F_{k-1} \right\} = \sum_{k=1}^n Z(k) Z^T(k) E \left\{ \psi^2[\xi(k)] | F_{k-1} \right\} \quad (6.106)$$

$$S_n^2 = E\{V_n^2\} = \sum_{k=1}^n E\{Z(k)Z^T(k)\}E\{\psi^2[\xi(k)]\} \quad (6.107)$$

Analizirajući dalje izraze $n^{-1}V_n^2$ i $n^{-1}S_n^2$. Ako primenimo zakon velikih brojeva na (6.106) i (6.107), zaključujemo da je

$$\frac{V_n^2}{n} \xrightarrow{n \rightarrow \infty} E\{Z(k)Z^T(k)\}E\{\psi^2(\xi(k))\} \quad (6.108)$$

$$\frac{S_n^2}{n} \xrightarrow{n \rightarrow \infty} E\{Z(k)Z^T(k)\}E\{\psi^2(\xi(k))\} \quad (6.109)$$

Relacija (6.108) i (6.109) pokazuje da je prvi uslov leme 6.2 zadovoljen. Takođe, pošto je $\lim_{n \rightarrow \infty} S_n = \infty$ a veličina $Z(k)\psi(\xi(k))$ su na osnovu A1, A2 i A4, ograničene, zaključuje se da je

$$\frac{1}{S_n^2} \sum_{k=1}^n E\left\{ \left[Z(k)Z^T(k)\psi^2(\xi(k)) \right]^2 \right\} P\left\{ \|Z(k)\psi(\xi(k))\| > \varepsilon S_n \right\} \xrightarrow{n \rightarrow \infty} 0 \quad (6.110)$$

Relacija (6.110) pokazuje da je i drugi uslov leme 6.2 ispunjen, te se ova lema može primeniti na (6.105), na osnovu čega se dobija

$$L - L(i) = \frac{1}{\sqrt{i}} \sum_{k=1}^i Z(k)\psi[\xi(k)] \xrightarrow{D} N(0, Q) \quad (6.111)$$

$$Q = E\{Z(k)Z^T(k)\}E\{\psi^2[\xi(k)]\}. \quad (6.112)$$

Takođe, na osnovu relacije (6.86) i smene $\bar{R}(i) = iR(i)$, može se pisati

$$R(i) = \frac{1}{i} \bar{R}(0) + \frac{1}{i} \sum_{k=1}^i \alpha Z(k)Z^T(k) \quad (6.113)$$

Primenjujući zakon velikih brojeva na (6.113), zaključuje se da je

$$R(i) = E\{\alpha Z(i)Z^T(i)\} = E\{\psi'(\xi(i))\}E\{Z(i)Z^T(i)\} \quad (6.114)$$

gde je $\alpha = E\{\psi'(\cdot)\}$. Konačno, na osnovu (A3) i relacija (6.112) – (6.114) dobija se

$$\frac{1}{\sqrt{i}} R^{-1}(i) \sum_{k=1}^i Z(k)\psi[\xi(k)] \xrightarrow{D} N(0, P); \quad (6.115)$$

gde je $P = R^{-1}QR^{-1}$ definisano relacijom (6.84), čime je kompletiran dokaz.

6.3. STATISTIČKO MODELIRANJE POREMEĆAJA

Ključni korak u izloženom postupku sinteze optimalnih estimatora na zadatoj klasi raspodela, bilo nerekurzivnih ili rekurzivnih, predstavlja definisanje, na osnovu raspoložive apriorne informacije, klase raspodela P kojoj pripada i stvarna raspodela poremećaja i određivanje najnepovoljnije gustine raspodele u okviru pretpostavljene klase, iz uslova minimuma Fišerove informacione matrice. Na osnovu izvedenog analitičkog izraza za najnepovoljniju funkciju gustine verovatnoće (6.27) direktno sledi kriterijumska funkcija (6.52) i optimalna matrica pojačanja (6.53), (6.54) za rekurzivni robusni estimator stohastičke aproksimacije (6.51). Ova matrica se dalje aproksimira izrazom (6.77) ili (6.83), respektivno, čime se dolazi do praktično primenjivog izraza za suboptimalnu matricu pojačanja rekurzivne stohastičke procedure parametarske identifikacije. U ovom odeljku navedeni su neki primeri klasa raspodela, koje u praksi predstavljaju najrasprostranjenije tipove raspoložive apriorne informacije o statističkim karakteristikama poremećaja, i određena je aproksimativno najnepovoljnija funkcija gustine verovatnoće u okviru svake od klasa, na osnovu koje se dalje određuje optimalna kriterijumska funkcija.

Određivanje najnepovoljnije gustine raspodele zahteva rešavanje izuzetno kompleksnog problema optimizacije (6.27), koji se u opštem slučaju ne može rešiti u zatvorenoj formi. Ovaj problem se može rešiti u analitičkom obliku jedino ako se pretpostavi da Fisherova informaciona matrica sa desne strane izraza (6.27) zavisi od nepoznate funkcije gustine verovatnoće p jedino preko Fisherove količine informacija $I(p)$ u (6.22), dok je matrica B približno konstantna. Tada se razmatrani problem određivanja najnepovoljnije funkcije gustine verovatnoće svodi na

$$p_0 = \arg \min_{p \in P} I(p) \quad (6.116)$$

U stohastičkoj literaturi postoje različiti postupci za rešavanje postavljenog optimizacionog problema, od kojih se najčešće koriste postupci Hubera [40] i Cipkina i Poljaka [72,89]. U nastavku ćemo razmotriti pomenute postupke.

Polazeći od metodologije Hubera [40], označimo sa

$$p_\varepsilon = (1 - \varepsilon)p_0 + \varepsilon p; \quad \varepsilon \in (0,1) \quad (6.117)$$

proizvoljnu funkciju gustine verovatnoće u okviru pretpostavljene klase P , gde je p_0 najnepovoljnija funkcija gustine verovatnoće za koju je $I(p_0) < I(p_\varepsilon)$. Odgovarajuće funkcije raspodele verovatnoće P_0 i P su simetrične i konveksne funkcije, te je i raspodela P_ε simetrična i konveksna funkcija i pripada usvojenoj klasi P . Definišimo dalje $I(\varepsilon) = I(p_\varepsilon)$ i $\rho(\varepsilon) = \rho(p_\varepsilon)$ na osnovu izraza (6.41) i (6.42), respektivno. Na osnovu relacija (6.45) i (6.47) ove funkcije zadovoljavaju Jensenovu nejednakost, te su konveksne. Takođe, pošto je $\rho(\varepsilon)$ konveksna funkcija, ukoliko je $d\rho(\varepsilon)/d\varepsilon \geq 0$ u tački $\varepsilon = 0$ imaće globalni minimum. Polazeći od (6.42) može se pisati

$$\left. \frac{d\rho(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = \int_{-\infty}^{\infty} [-2\psi_0(z)g'(z) - \psi_0^2(z)g(z)]dz; \quad g = p - p_0. \quad (6.118)$$

gde je

$$\psi_0(z) = -[\log p_0(z)]' \quad (6.119)$$

Na isti način se pokazuje, diferenciranjem (6.41), da je

$$\left. \frac{dI(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = \int_{-\infty}^{\infty} [-2\psi_0(z)g'(z) - \psi_0^2(z)g(z)]dz; \quad g = p - p_0. \quad (6.120)$$

S obzirom na (6.46), dalje, sledi

$$\frac{1}{\varepsilon} \left[\frac{p_\varepsilon'^2}{p_\varepsilon} - \frac{p_0'^2}{p_0} \right] \leq \frac{p'^2}{p} - \frac{p_0'^2}{p_0}. \quad (6.121)$$

Leva strana u (6.121) je integrabilna po pretpostavci, pa je

$$\frac{1}{\varepsilon} \int_{-\infty}^{\infty} \left[\frac{p_\varepsilon'^2}{p_\varepsilon} - \frac{p_0'^2}{p_0} \right] dz \leq \int_{-\infty}^{\infty} \left[\frac{p'^2}{p} - \frac{p_0'^2}{p_0} \right] dz = I(p) - I(p_0) \geq 0. \quad (6.122)$$

Ovde je iskorišćena činjenica da je p_0 najnepovoljnija funkcija gustine verovatnoće u smislu minimuma Fisherove količine informacije, $I(p_0) \leq I(p)$, a i leva strana nejednakosti u (6.122) je na osnovu toga nenegativna, tj.

$$\frac{1}{\varepsilon} \int_{-\infty}^{\infty} \left[\frac{p_\varepsilon'^2}{p_\varepsilon} - \frac{p_0'^2}{p_0} \right] dz = \frac{1}{\varepsilon} [I(p_\varepsilon) - I(p_0)] \geq 0, \quad (6.123)$$

te su ispunjeni uslovi za prelaz limesa pod znak integrala, odnosno

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_{-\infty}^{\infty} \left[\frac{p_\varepsilon'^2}{p_\varepsilon} - \frac{p_0'^2}{p_0} \right] dz = \int_{-\infty}^{\infty} \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left[\frac{p_\varepsilon'^2}{p_\varepsilon} - \frac{p_0'^2}{p_0} \right] dz. \quad (6.124)$$

Ako se u (6.124) p_ε zameni definicionim izrazom (6.117) dobija se

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left[\frac{p_\varepsilon'^2}{p_\varepsilon} - \frac{p_0'^2}{p_0} \right] = 2 \frac{p_0'}{p_0} [p' - p_0'] - \frac{p_0'^2}{p_0^2} [p - p_0]. \quad (6.125)$$

S obzirom na (6.119), (6.123) i (6.125), dalje, sledi

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_{-\infty}^{\infty} \left[\frac{p_\varepsilon'^2}{p_\varepsilon} - \frac{p_0'^2}{p_0} \right] dz = \int_{-\infty}^{\infty} [-2\psi_0 g' - \psi_0^2 g] dz \geq 0. \quad (6.126)$$

gde je $g = p - p_0$. Pošto je leva strana nejednakosti (6.126) jednaka $\rho'(0)$ (relacija (6.118)) to je $\rho'(0) \geq 0$.

Imajući u vidu (6.120) važi i obrnut stav. Naime, ako je gustina $p_0(\cdot)$ takva da je

$$\int_{-\infty}^{\infty} [-2\psi_0(z)g'(z) - \psi_0^2(z)g(z)] dz \geq 0; \quad g(z) = p(z) - p_0(z) \quad (6.127)$$

za svako $p \in P$ tada je $I'(0) \geq 0$, i pošto je $I(p_\varepsilon)$ konveksna funkcija biće $I(0) = I(p_0) \geq I(p)$ za $\forall p \in P$, odnosno gustina $p_0(\cdot)$ minimiziraće Fišerovu količinu informacija $I(p)$. Drugim rečima, uslov (6.127) predstavlja potreban i dovoljan uslov da bi funkcija gustine raspodele $p_0(\cdot)$ minimizirala $I(p)$. Pošto je $I(p)$ konveksna funkcija i funkcija p_0 je konveksna, tj. $p_0 \in P$.

Ovde se ništa ne govori o egzistenciji rešenja $p_0(\cdot)$ i o tome šta se dešava ukoliko raspodele $P_\varepsilon \in P$ ne poseduju apsolutno neprekidne gustine $p_\varepsilon(\cdot)$. Može se pokazati da je $I(p_\varepsilon) < \infty$ ako i samo ako sve raspodele u okviru klase P poseduju apsolutno neprekidne gustine, za koje je $\int p_\varepsilon'^2 / p_\varepsilon dz < \infty$. Ako je, dalje skup P kompaktan i konveksan i $\inf_{p_\varepsilon \in P} I(p_\varepsilon) = a < \infty$ postojaće jedinstveno rešenje p_0 takvo da je $I(p_0) = a < \infty$ [40].

Prilaz Cipkina i Poljaka je znatno jednostavniji za primenu u praksi, a u njegovoj osnovi nalazi se nejednakost Koši-Švarc-Bunjakovskog [1,60,79,86]

$$\left[\int_{-\infty}^{\infty} \phi(z) G(z) p(z) dz \right]^2 \leq \int_{-\infty}^{\infty} G^2(z) p(z) dz \int_{-\infty}^{\infty} \phi^2(z) p(z) dz. \quad (6.128)$$

Ako se u (6.128) usvoji $\phi(z) = p'_z(z) / p_\xi(z)$ sledi

$$I(p_\xi) = \int_{-\infty}^{\infty} \frac{p_\xi'^2(z)}{p_\xi(z)} dz \geq \frac{\left[\int_{-\infty}^{\infty} G(z) p_\xi'(z) dz \right]^2}{\int_{-\infty}^{\infty} G^2(z) p_\xi(z) dz} = \frac{\left[\int_{-\infty}^{\infty} G'(z) p_\xi(z) dz \right]^2}{\int_{-\infty}^{\infty} G^2(z) p_\xi(z) dz} \quad (6.129)$$

gde je $p_\xi \in P$ i $G(\cdot)$ proizvoljna funkcija. Desna jednakost u (6.129) dobijena je primenom parcijalne integracije. Za najnepovoljniju gustinu $p_0 \in P$ u nejednakosti (6.129) treba da važi znak jednakosti. Sa druge strane, lako je pokazati da se znak jednakosti u relaciji (6.129) dobija za

$$\frac{p'_0(z)}{p_0(z)} = [\log p_0(z)] = kG(z) \quad (6.130)$$

Na taj način, najnepovoljnija gustina dobija se kao rešenje diferencijalne jednačine (6.130), gde funkcija $G(\cdot)$ zavisi od pretpostavljene klase P . Dobijeno rešenje je jedinstveno ukoliko su sve funkcije raspodele verovatnoće unutar klase P konveksne i simetrične funkcije za koje postoje apsolutno neprekidne funkcije gustine verovatnoće, a za koje Fišerova količina informacija $I(p)$ postoji i konačna je. Razmatramo dalje neke klase raspodela, koje predstavljaju uobičajenu apriornu informaciju o statističkim karakteristikama šuma.

a) P_1 - klasa funkcija gustine verovatnoće neprekidnih u koordinatnom početku

$$P_1 = \{p_\varepsilon(z): p(0) \geq 1/2a > 0\}. \quad (6.131)$$

Klasa P_1 definiše praktično minimalno mogući obim apriorne statističke informacije o poremećaju u modelu, odnosno uslov (6.131) praktično znači potpuno odsustvo apriorne informacije o statističkim karakteristikama poremećaja. Ako se u (6.130) usvoji $G(z) = \operatorname{sgn} z$ i reši diferencijalna jednačina dobija se $p_0(z) = C \exp\{-k|z|\}$.

Konstanta integracije C određuje se iz uslova normiranja $\int_{-\infty}^{\infty} p_0(z) dz = 1$, odakle sledi $C = k/2$. S obzirom da je i $p_0(z) \in P$, to je $p_0(0) = k/2 \geq 1/2a$, odnosno $k = 1/a$. Na taj način, najnepovoljnija gustina verovatnoće u klasi P_1 je Laplasova funkcija gustine verovatnoće

$$p_0(z) = \frac{1}{2a} \exp\left\{-\frac{1}{a}|z|\right\}. \quad (6.132)$$

Optimalna kriterijumska funkcija u algoritmu parametarske estimacije, na osnovu (6.26), je

$$F_0(z) = -\log p_0(z) = \log 2a + \frac{1}{a}|z|, \quad (6.133)$$

a njen prvi izvod, tzv. funkcija uticaja [33,40,76]

$$\psi_0(z) = -[\log p_0(z)] = \frac{1}{a} \operatorname{sgn} z \quad (6.134)$$

Fišerova količina informacija za najnepovoljniju funkciju gustine verovatnoće p_0 data je izrazom

$$I(p_\xi) \geq I(p_0) = \frac{\left[\int_{-\infty}^{\infty} \operatorname{sgn}(z) p_0'(z) dz \right]^2}{\int_{-\infty}^{\infty} p_0(z) dz} = 4p_0^2(0) = a^{-2}. \quad (6.135)$$

b) P_2 - klasa gustina raspodela verovatnoće sa ograničenom disperzijom

$$P_2 = \left\{ p_\xi(z) : \int_{-\infty}^{\infty} z^2 p_\xi(z) dz \leq \sigma^2 \right\}. \quad (6.136)$$

Ako se u (6.130) usvoji $G(z) = z$ i reši diferencijalna jednačina dobija se $p_0(z) = C \exp\{-kz^2/2\}$. Na osnovu uslova normiranja $\int_{-\infty}^{\infty} p_0(z) dz = 1$, dalje, sledi $C = \sqrt{k/2\pi}$. Pošto je $p_0 \in P$, i p_0 zadovoljava uslov $\int_{-\infty}^{\infty} z^2 p_0(z) dz \leq \sigma^2$, na taj način najnepovoljnija gustina verovatnoće u okviru date klase je normalna, tj.

$$p_0(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{z^2}{2\sigma^2}\right\}. \quad (6.137)$$

Optimalna kriterijumska funkcija u (6.26) je

$$F_0(z) = -\log p_0(z) = \log(2\pi\sigma) + \frac{z^2}{2\sigma^2}, \quad (6.138)$$

S obzirom na usvojeno $G(\cdot)$ i (6.129) dobija se za Fišerovu količinu informacije

$$I(p_\xi) \geq \frac{1}{\int_{-\infty}^{\infty} z^2 p_0(z) dz} = \frac{1}{\sigma^2} = I(p_0) \quad (6.139)$$

dok je funkcija uticaja data izrazom

$$\psi_0(z) = -[\log p_0(z)]' = \frac{z}{\sigma^2} \quad (6.140)$$

c) P_3 - klasa kontaminiranih ili aproksimativno normalnih funkcija gustine verovatnoće

$$P_3 = \{p_\xi(z) : p_\xi(z) = (1-\varepsilon)p_1(z) + \varepsilon p_2(z), \quad 0 \leq \varepsilon < 1, \quad p_1 \text{ je normalna gustina } N(0, \sigma^2), \quad p_2(\cdot) \text{ proizvoljna simetrična gustina } \}. \quad (6.141)$$

Parametar ε naziva se stepen kontaminacije i označava verovatnoću pojavljivanja uzoraka generisanih funkcijom gustine raspodele verovatnoće $p_2(\cdot)$. Ukoliko je $\varepsilon = 0$ svi uzorci su normalno raspodeljeni, dok uslov $\varepsilon = 1$ karakteriše potpuno odsustvo apriorne informacije o raspodeli poremećaja. Razmatrana klasa predstavlja u literaturi najčešće korišćeni model sporadičnih poremećaja koji po svojim statističkim osobinama odstupaju od ostatka populacije (anomalna merenja, “bad data” ili “outlier”-i) [4,33,34,40,75,76,89]. Ukoliko u sistemu postoje retke, velike realizacije poremećaja kao statistički model poremećaja navedenog tipa usvaja se, obično, klasa (6.141) uz pretpostavku da je i $p_2(\cdot)$ gausovska gustina, nulte srednje vrednosti i varijanse $\sigma_2^2 \gg \sigma^2$. U literaturi posvećenoj obradi signala ovakav šum naziva se impulsnim [27,46,59,94]

Uslov (6.141) može se napisati i u obliku

$$p_\varepsilon(z) \geq (1-\varepsilon)p_1(z) \quad (6.142)$$

uz ograničenje

$$\int_{-\infty}^{\infty} p_\varepsilon(z) dz = 1. \quad (6.143)$$

Na taj način razmatrani zadatak se svodi na problem minimizacije Fišerove količine informacija $I(p)$ uz ograničenja (6.142) i (6.143). Rešenje ovog zadatka može se tražiti u obliku $p_0(z) = (1-\varepsilon)p_1(z)$ za $|z| \leq \Delta$ i $p_0(z) \geq (1-\varepsilon)p_1(z)$ za $|z| > \Delta$, gde je $\Delta > 0$ i $p_0(-z) = p_0(z)$. U tom slučaju najnepovoljnija gustina određuje se iz uslova

$$p_0(z) = \arg \min_{p_\varepsilon \in P_\Delta} \int_{-\infty}^{\infty} \frac{p_\varepsilon'^2(z)}{p_\varepsilon(z)} dz \quad (6.144)$$

uz ograničenja

$$\int_0^\Delta (1-\varepsilon)p_1(z) dz + \int_\Delta^\infty p_\varepsilon(z) dz = \frac{1}{2} \quad (6.145)$$

$$p_\varepsilon(\Delta) = (1-\varepsilon)p_1(\Delta). \quad (6.146)$$

Ako se, dalje, u (6.130) usvoji $G(z) = 1$ i diferencijalna jednačina reši na intervalu (Δ, ∞) dobija se $p_0(z) = k_1 \exp\{-k_2 z\}$. Nepoznate konstante k_1 i k_2 mogu se odrediti u funkciji parametra Δ iz (6.146) i uslova $p_0'(\Delta) = (1-\varepsilon)p_1'(\Delta)$, gde je $p_1(z) = N(0, \sigma^2)$. Na taj način se dobija

$$p_0(z) = \begin{cases} \frac{1-\varepsilon}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{z^2}{2\sigma^2}\right\} & |z| \leq \Delta \\ \frac{1-\varepsilon}{\sqrt{2\pi}\sigma} \exp\left\{\frac{\Delta^2}{2\sigma^2} - \frac{\Delta|z|}{\sigma^2}\right\} & |z| > \Delta \end{cases} \quad (6.147)$$

Funkcionalna veza između parametara Δ i ε definisana je uslovom normiranja (6.143), tj.

$$\int_{-\Delta}^{\Delta} p_1(z) dz + \frac{2\sigma^2}{\Delta} p_1(\Delta) = \frac{1}{1-\varepsilon}. \quad (6.148)$$

U tabeli 6.1 date su vrednosti parametra $k = \frac{\Delta}{\sigma}$, određene na osnovu (6.148), za različite vrednosti stepena kontaminacije ε . Interesantno je uočiti da se i za male vrednosti stepena kontaminacije ε dobijaju relativno male vrednosti parametra k .

ε	0	0.01	0.02	0.05	0.1	0.2	0.5	1
$k = \frac{\Delta}{\sigma}$	∞	2	1.7	1.4	1.1	0.9	0.4	0

TABELA 6.1 Funkcionalna veza statističkih parametara u ε -kontaminiranoj klasi normalnih funkcija gustine verovatnoće

S obzirom na (6.147), optimalna kriterijumska funkcija (6.26) je

$$F_0(z) = -\log p_0(z) = \begin{cases} \frac{\sqrt{2\pi}\sigma}{1-\varepsilon} + \frac{z^2}{2\sigma^2} & |z| \leq \Delta(\varepsilon) \\ \frac{\sqrt{2\pi}\sigma}{1-\varepsilon} - \frac{1}{2} \frac{\Delta^2}{\sigma^2} + \frac{|z|\Delta}{\sigma^2} & |z| > \Delta(\varepsilon) \end{cases} \quad (6.149)$$

odnosno, izvedeni estimator ekvivalentan je Windzorijanskoj sredini [4,76].

Fišerova količina informacija za najnepovoljniju funkciju gustine verovatnoće data je izrazom

$$I^{-1}(p_0) = \left[\int_{-\Delta}^{\Delta} \psi_0'(z) p_0(z) dz \right] = \left[2(1-\varepsilon) \operatorname{erf}\left(\frac{\Delta}{\sigma}\right) \right]^{-1} \quad (6.150)$$

gde je $\psi_0(z) = dF_0(z)/dz$ i funkcija greške

$$\operatorname{erf}(k) = \frac{1}{\sqrt{2\pi}} \int_0^k \exp\left\{-\frac{1}{2} z^2\right\} dz. \quad (6.151)$$

Izvod optimalne funkcije gubitaka, odnosno funkcija uticaja, data je izrazom

$$\psi_0(z) = -[\log p_0(z)] = \begin{cases} \frac{z}{\sigma^2} & |z| \leq \Delta \\ \frac{\Delta}{\sigma^2} \operatorname{sgn} z & |z| > \Delta \end{cases}, \quad \Delta = \Delta(\varepsilon) \quad (6.152)$$

gde je $\Delta(\varepsilon)$ definisano relacijom (6.148).

U praksi vrednosti indeksa kontaminacije ε i varijanse šuma σ^2 nisu nikada tačno poznate, tako da se relacija (6.148) ne može pouzdano koristiti za određivanje Δ . U sedmom poglavlju biće razmatrana adaptivna procedura estimacije kod kojih se statistički parametri ε i σ^2 određuju iz reziduala merenja, jednovremeno sa procesom estimacije parametara.

Primetimo na kraju, da se aproksimacija neklasičnog optimizacionog problema (6.27) suboptimalnom procedurom (6.116) može opravdati i na osnovu sledećeg razmatranja. Naime, pretpostavimo da je matrica pojačanja $\Gamma(i)$ u (6.51) definisana kao sledeća aproksimacija optimalne matrice pojačanja $\Gamma_o(i)$ u (6.54)

$$\Gamma^{-1}(i) = i\beta(\psi_0, p)B(i); \quad B(i) = i^{-1} \sum_{j=1}^i \nabla_{\theta} v(j, \theta^*) \nabla_{\theta}^T v(j, \theta^*) \quad (6.153)$$

Ovakva aproksimacija pretpostavlja da je i dovoljno veliko, kao i da procena parametara $\hat{\theta}(i)$ u (6.51) konvergira ka θ^* , odnosno da je $\hat{\theta}(i)$ blisko θ^* . Pod navedenim pretpostavkama gradijent kriterijumske funkcije $F(\cdot)$ u (6.2) može se aproksimirati Taylorovim redom na sledeći način

$$\begin{aligned} \nabla_{\theta} F(v(i+1, \hat{\theta}(i))) &\approx \nabla_{\theta} v(i+1, \theta^*) \cdot \\ &[\psi(v(i+1, \theta^*)) + \psi'(v(i+1, \theta^*)) \nabla_{\theta}^T v(i+1, \theta^*) \tilde{\theta}(i)] \end{aligned} \quad (6.154)$$

gde je $\psi(\cdot) = F'(\cdot)$ i $\tilde{\theta}(i) = \hat{\theta}(i) - \theta^*$. Dalje se uslovna kovarijaciona matrica inkrementa procena vektora parametara u dva sukcesivna trenutka vremena, označenih kao $(i+1)$ i (i) , kada su na raspolaganju merenja do trenutka i , može približno napisati kao

$$\begin{aligned} E\{i\Delta\theta(i+1)\Delta\theta^T(i+1)|i\} &\approx [i\beta B(i)]^{-1} \nabla_{\theta} v(i+1, \theta^*) \cdot \\ &[\beta \tilde{\theta}^T(i) \nabla_{\theta} v(i+1, \theta^*)]^2 + \alpha \nabla_{\theta}^T v(i+1, \theta^*) [\beta B(i)]^{-1} \end{aligned} \quad (6.155)$$

Poslednji izraz se dobija kada se relacija (6.51) napiše u alternativnom obliku

$$\hat{\theta}(i+1) = \hat{\theta}(i) - \Gamma(i) \nabla_{\theta} F(v(i+1, \hat{\theta}(i))) \quad (6.156)$$

odakle se za inkrement procena vektora parametara dobija

$$\Delta\theta(i+1) = \hat{\theta}(i+1) - \hat{\theta}(i) = -\Gamma(i)\nabla_{\theta}F(v(i+1, \hat{\theta}(i))) \quad (6.157)$$

gde je $\Gamma(i)$ definisano izrazom (6.153) a gradijent nelinearne transformacije reziduala merenja izrazom (6.154). Nalaženjem transponovanog vektora $\Delta\theta^T(i+1)$ i množenjem sa $\Delta\theta(i+1)$, nakon primene operatora uslovnog matematičkog očekivanja $E\{\cdot|i\}$ na tako dobijeni proizvod, dobija se izraz (6.115), gde su α i β definisani izazima (6.34) i (6.35), respekti vno. Ako se, dalje, prepostavi da procene parametara konvergiraju na takav način da je

$$\left[\tilde{\theta}^T(i)\nabla_{\theta}v(i+1, \theta^*)\right]^2 < \varepsilon < \alpha\beta^2 \quad (6.158)$$

relacija (6.155) može se dalje aproksimirati izrazom

$$E\{i\Delta\theta(i+1)\Delta\theta^T(i+1)|i\} \approx i^{-1}B^{-1}(i)\nabla_{\theta}v(i+1, \theta^*)\nabla_{\theta}^Tv(i+1, \theta^*)B^{-1}(i)\alpha\beta^{-2} \quad (6.159)$$

Na taj način, ako nam je cilj da eliminišemo realizaciju impulsnog šuma u trenutku $(i+1)$, koja je generisana funkcijom gustine verovatnoće P_2 u (6.141), uslova kovarijacione matrice sa leve strane izraza (6.159) mora se učiniti što je moguće manjom, imajući u vidu postignutu tačnost estimacije parametara u prethodnoj iteraciji i zadate uslvoe stohastičkog okruženja. Pošto u izrazu (6.151) od nepoznate funkcije gustine verovatnoće p iz usvojene klase P zavisi samo skalarna veličina

$$\rho(\psi, p) = \frac{\alpha(\psi, p)}{\beta^2(\psi, p)} = \frac{E_p\{\psi^2(z)\}}{E_p\{\psi'(z)\}} \quad (6.160)$$

neophodno je da se ovaj faktor učini što je moguće manjim, imajući u vidu da realna funkcija gustine verovatnoće p^* nije poznata. Kao što je pokazao u ovom poglavlju, izbor nelinearne transformacije reziduala

$$\psi(\cdot) = \psi_0(\cdot) = -[\log p_0(\cdot)]; \quad p_0 = \arg \min_{p \in P} I(p) \quad (6.161)$$

gde je p_0 najnepovoljnija funkcija gustine verovatnoće koja minimizira Fišerovu količinu informacija $I(p)$, u okviru zadate klase P , tj. koja se dobija iz uslova (6.116), rezultuje u uslov mini-maksa

$$\min_{\psi} \rho(\psi, p_0) = \rho(\psi_0, p_0) = \max_{p \in P} \rho(\psi_0, p) \quad (6.162)$$

gde je P usvojena klasa funkcija raspodela verovatnoće, za koje postoje odgovarajuće funkcije gustine verovatnoće p , a $\psi(\cdot)$ je proizvoljna klasa neparnih funkcija uticaja, odakle se zaključuje da je

$$\rho(\psi_0, p) \leq \rho(\psi_0, p_0) = I(p_0) \leq \rho(\psi, p_0) \quad (6.163)$$

a to je istovremeno u važnosti i za uslovnu matricu kovarijanse u (6.159).

Dakle, izborom nelinearne transformacije reziduala u (6.51) na osnovu izraza (6.161), uz istovremeni izbor matrice pojačanja algoritma (6.51) na osnovu (6.153) obezbeđuje se da uslovna matrica kovarijanse inkrementa procene parametara dva sukcesivna vremenska trenutka za proizvoljnu funkciju gustine verovatnoće iz zadate klase neće preći granicu robusnosti

$$i^{-1}B^{-1}(i)\nabla_{\theta}v(i+1,\theta^*)\nabla_{\theta}^T v(i+1,\theta^*)I(p_0) \quad (6.164)$$

dok će za proizvoljnu neparnu nelinearnu funkciju $\psi(\cdot)$ u (6.51) i izbor $\Gamma(i)$ prema (6.153) uslovna matrica kovarijanse inkrementa procena parametara za najnepovoljniju funkciju gustine verovatnoće p_0 iz (6.161) biti ne manja od veličine definisane sa (6.164), koja predstavlja granicu robusnosti.

7. ADAPTIVNA ROBUSNA ESTIMACIJA PARAMETARA METODOM QQ-KRIVE

U ovom poglavlju analiziran je, pomalo specifičan pristup problemu estimacije parametara linearnog stohastičkog sistema, zasnovan na metodu QQ-krive. Naime, do sada je u literaturi QQ-krivom označavan vrlo jednostavan i efikasan način da se proverí valjanost pretpostavke o raspodeli sekvence realizacija nekog stohastičkog procesa [18,26]. U specijalnom slučaju, kada je pretpostavljena Gausovska raspodela, a prava raspodela stohastičkog procesa poseduje kontaminaciju koja je okarakterisana takozvanim zadebljalim repovima, odnosno prisustvom realizacija velikog intenziteta, tada se QQ-kriva deformiše na karakterističan, vrlo prepoznatljiv način. U ovom radu je pokazano da ukoliko se u dvodimenzionom prostoru formira skup tačaka kojima je jedna od koordinata vrednost realizacije stohastičkog procesa a druga koordinata je funkcija ranga dotične realizacije u sortiranoj sekvenci i pretpostavljene funkcije raspodele, tada će se, ukoliko je pretpostavka o raspodeli dobra, ove tačke, sa malim odstupanjem, naći na jednoj pravoj [19,20]. Na osnovu ovakve osobine QQ-krive moguće je proširiti njenu primenu. U prvom delu ovog poglavlja, pokazano je kako se ovaj metod može koristiti kao estimator srednje vrednosti i varijanse i izvedene su osobine takvog estimatora, dok je u drugom delu procedura primenjena na estimaciju parametara linearnog dinamičkog stohastičkog sistema i istovremenu estimaciju statistika prvog i drugog reda šuma merenja. Estimaciona procedura se zasniva na skupu podataka iz pokretnog prozora i na taj način joj je inherentno pridružen atribut adaptivnosti. Sa druge strane, korišćenjem rezultata teorije min-maksne robusnosti [50,51,52], QQ-metod estimacije se vrlo jednostavno može učiniti robusnim u smislu neosetljivosti na loša merenja koja su posledica zadebljanih repova funkcije gustine verovatnoće šuma merenja. Dobijeni rezultat je primenjen na problem estimacije nepoznatih statističkih parametara šuma u cilju definisanja funkcije verodostojnosti i projektovanja adaptivnog rekurzivnog estimatora nepoznatih parametara sistema.

7.1. QQ-METOD ESTIMACIJE STATISTIČKIH PARAMETARA

Posmatrajmo, u širem smislu stacionaran, stohastički proces čijih nam je n realizacija poznato $\{z_i\}$, $i=1,\dots,n$. Neka je tom procesu pridružena funkcija raspodele verovatnoće $F(z)$ sa odgovarajućom funkcijom gustine verovatnoće $f(z)$. Sortiranjem ovih realizacija $\{z_i\}$ dobija se neopadajuća sekvenca $\{y_i\}$, $i=1,\dots,n$ takva da je $y_i \leq y_j$ za $i < j$. Uslovnu verovatnoću da je neka od realizacija y i -tog ranga u sortiranoj sekvenci $\{y_i\}$ jednostavno je odrediti primenom Bernulijevog eksperimenta [41,69,71]:

$$P(i/y) = \binom{n-1}{i-1} F^{i-1}(y) (1-F(y))^{n-1} \quad (7.1)$$

Zanimljivo je pogledati čemu su onda jednaki uslovno matematičko očekivanje i uslovna varijansa slučajne promenljive ranga i , $m_{i/y}$ i $\sigma_{i/y}^2$, respektivno:

$$\begin{aligned} E\{i/y\} &= m_{i/y} = \sum_{i=1}^n i P(i/y) \\ &= \sum_{i=1}^n i \binom{n-1}{i-1} F^{i-1}(y) (1-F(y))^{n-1} \\ &= (1-F(y))^{n-1} \sum_{i=1}^n \frac{i(n-1)!}{(i-1)!(n-i)!} \left(\frac{F(y)}{1-F(y)} \right)^{i-1} \end{aligned} \quad (7.2)$$

Uvodeći smenu:

$$\alpha(i) = \left(\frac{F(y)}{1-F(y)} \right) \quad (7.3)$$

dobija se:

$$\begin{aligned} m_{i/y} &= (1-F(y))^{n-1} \frac{d}{d\alpha} \sum_{i=1}^n \frac{(n-1)!}{(i-1)!(n-i)!} \int i \alpha^{i-1} d\alpha \\ &= (1-F(y))^{n-1} \frac{d}{d\alpha} \sum_{i=1}^n \binom{n-1}{i-1} \alpha^i \\ &= (1-F(y))^{n-1} \frac{d}{d\alpha} \sum_{i=0}^{n-1} \binom{n-1}{i} \alpha^{i+1} \\ &= (1-F(y))^{n-1} \frac{d}{d\alpha} (\alpha(1+\alpha)^{n-1}) \\ &= (1-F(y))^{n-1} ((1+\alpha)^{n-1} + (n-1)\alpha(1+\alpha)^{n-2}) \end{aligned} \quad (7.4)$$

smenjujući (7.3) u (7.4) dobija se:

$$m_{i/y} = E\{i/y\} = 1 + (n-1)F(y) \quad (7.5)$$

Analognim postupkom može se odrediti i uslovna varijansa:

$$\sigma_{i/y}^2 = \text{var}\{i/y\} = E\{i^2/y\} - E^2\{i/y\}$$

$$= \sum_{i=1}^n \binom{n-1}{i-1} F^{i-1}(y) [1-F(y)]^{n-i} - m_{i/y}^2 \quad (7.6)$$

Dalje se može pisati:

$$\begin{aligned} \text{var}\{i/y\} &= (1-F(y))^{n-1} \sum_{i=1}^n i^2 \binom{n-1}{i-1} \alpha^{i-1} - m_{i/y}^2 \\ &= (1-F(y))^{n-1} \frac{d}{d\alpha} \sum_{i=1}^n i \binom{n-1}{i-1} \alpha^i - m_{i/y}^2 \\ &= (1-F(y))^{n-1} \frac{d}{d\alpha} \left[\alpha \sum_{i=1}^n i \binom{n-1}{i-1} \alpha^{i-1} \right] - m_{i/y}^2 \\ &= (1-F(y))^{n-1} \frac{d}{d\alpha} \left[\alpha \frac{d}{d\alpha} \sum_{i=1}^n i \binom{n-1}{i-1} \alpha^i \right] - m_{i/y}^2 \\ &= (1-F(y))^{n-1} \frac{d}{d\alpha} \left[\alpha \frac{d}{d\alpha} \sum_{i=0}^{n-1} i \binom{n-1}{i} \alpha^{i+1} \right] - m_{i/y}^2 \\ &= (1-F(y))^{n-1} \frac{d}{d\alpha} \left[\alpha \frac{d}{d\alpha} (\alpha(1+\alpha)^{n-1}) \right] - m_{i/y}^2 \end{aligned} \quad (7.7)$$

Ponovo smenom relacija (7.3) i (7.5) u (7.7) dobija se konačno:

$$\sigma_{i/y}^2 = E\{i^2/y\} - E^2\{i/y\} = (n-1)F(y)(1-F(y)) \quad (7.8)$$

Treba primetiti da je vrednost varijanse $\sigma_{i/y}^2$ mala za realizacije y generisane sa krajeva raspodele $F(y)$, jer je ili vrednost $F(y)$ ili $1-F(y)$ mala. Sa druge strane, varijansa $\sigma_{i/y}^2$ dostiže maksimum za vrednost $F(y)=0.5$ i taj maksimum iznosi $(n-1)/4$. To znači da se $\sigma_{i/y}^2$ može ograničiti odgovarajućim izborom broja podataka n . Dalje, mala vrednost $\sigma_{i/y}^2$ označava da je rang i slučajne promenljive y u sortiranoj sekvenci $\{y_i\}$ vrlo bliska srednjoj vrednosti $m_{i/y}$, odnosno $i \approx m_{i/y}$, što pak na osnovu relacije (7.5) znači da je relacija između ranga i i $F(y)$ aproksimativno linearna. Pri tome, treba primetiti da je linearnost utoliko izraženija ukoliko je vrednost broja podataka n manja i ukoliko je vrednost funkcije $F(y)$ udaljenija od 0.5. Skup tačaka oblika $(F^{-1}(\rho_i), y_i)$ naziva se QQ-krivom, gde je [27]. Usvajajući aproksimaciju $m_{i/y} \sim i$ u relaciji (7.5) dobija se eksplicitni izraz za QQ-krivu

$$y_i = F^{-1}(\rho_i) \approx F^{-1}(r_i); \quad \rho_i = ((i-1))/(n-1) \approx r_i = (i-0.5)/n \quad (7.9)$$

gde je sa $F^{-1}(\cdot)$ označena inverzna funkcija raspodele funkcije $F(\cdot)$. Vrednosti r_i i ρ_i su vrlo bliske, ali uobičajeno je da se koristi r_i kako bi se i za prvu i početnu realizaciju u sortiranoj sekvenci odredile konačne vrednosti $F^{-1}(r_i)$ u slučaju najčešće korišćenih raspodela, kao što su Gausovska, Košijeva i Laplasova. Ukoliko tako formiran skup tačaka QQ-krive (7.9) karakteriše linearnu zavisnost, tada je skup tačaka $\{y_i\}$, odnosno $\{z_i\}$ generisan shodno raspodeli $F(\cdot)$. Ukoliko se sa m označi srednja vrednost a sa σ standardna devijacija ove raspodele, tada se izraz za tačke QQ-krive može napisati u obliku forme linearne regresije:

$$y_i = m + \sigma F_n^{-1}(r_i) = m + \sigma \tilde{r}_i \quad (7.10)$$

Sa $F_n(\cdot)$ je označena standardna funkcija raspodele nulte srednje vrednosti i jedinične varijanse koja je pridružena slučajnim promenljivima $\tilde{r}_i = (y_i - m)/\sigma$. Estimacija srednje vrednosti i standardne devijacije se tada može izvršiti primenom metode najmanjih kvadrata, koja je opisana u četvrtom poglavlju, odnosno

$$\begin{bmatrix} \hat{m} \\ \hat{\sigma} \end{bmatrix} = \left(\sum^t \sum \right)^{-1} \sum^t Y \quad (7.11)$$

gde su $2 \times n$ dimenziona matrica $\sum^t$ i n dimenzioni vektor Y :

$$\sum^t = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \tilde{r}_1 & \tilde{r}_2 & \dots & \tilde{r}_n \end{bmatrix}; \quad Y^t = [y_1 y_2 \dots y_n] \quad (7.12)$$

Uzimajući u obzir relacije (7.11) i (7.12) izraz za procenu srednje vrednosti se može napisati u sledećoj formi:

$$\hat{m} = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n \tilde{r}_i^2 - \sum_{i=1}^n \tilde{r}_i \sum_{i=1}^n y_i \tilde{r}_i}{n \sum_{i=1}^n \tilde{r}_i^2 - \left[\sum_{i=1}^n \tilde{r}_i \right]^2} \quad (7.13)$$

tako da pomeraj ove procene iznosi

$$b(\hat{m}) = E\{\hat{m}\} - m = \frac{\sum_{i=1}^n \tilde{r}_i \sum_{i=1}^n \tilde{r}_i (m - m_{y/i})}{n \sum_{i=1}^n \tilde{r}_i^2 - \left[\sum_{i=1}^n \tilde{r}_i \right]^2} \quad (7.14)$$

U slučaju simetrične raspodele $F(\cdot)$, možemo pisati:

$$\sum_{i=1}^n \tilde{r}_i = \frac{1}{\sigma} \sum_{i=1}^n \left[F^{-1}\left(\frac{i-1}{n-1}\right) - m \right] = -\frac{nm}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^n F^{-1}\left(\frac{i-1}{n-1}\right) \quad (7.15)$$

Dalje je

$$\begin{aligned} \sum_{i=1}^n F^{-1}\left(\frac{i-1}{n-1}\right) &= \sum_{i=1}^{\lfloor n/2 \rfloor} F^{-1}\left(\frac{i-1}{n-1}\right) + \sum_{i=\lfloor n/2 \rfloor+1}^{\lfloor (n+1)/2 \rfloor} F^{-1}\left(\frac{i-1}{n-1}\right) + \sum_{i=\lfloor (n+1)/2 \rfloor+1}^n F^{-1}\left(\frac{i-1}{n-1}\right) \\ &= \sum_{i=1}^{\lfloor n/2 \rfloor} F^{-1}\left(\frac{i-1}{n-1}\right) + \sum_{i=\lfloor n/2 \rfloor+1}^{\lfloor (n+1)/2 \rfloor} F^{-1}\left(\frac{i-1}{n-1}\right) + \sum_{j=1}^{\lfloor n/2 \rfloor} F^{-1}\left(\frac{n-j+1-1}{n-1}\right) \\ &= \sum_{i=1}^{\lfloor n/2 \rfloor} \left\{ F^{-1}\left(\frac{i-1}{n-1}\right) + F^{-1}\left(1 - \frac{i-1}{n-1}\right) \right\} = \sum_{i=\lfloor n/2 \rfloor+1}^{\lfloor (n+1)/2 \rfloor} F^{-1}(0.5) \end{aligned} \quad (7.16)$$

gde $\lfloor \cdot \rfloor$ označava ceo deo, a kako je za simetričnu raspodelu $F(\cdot)$

$$F^{-1}(x) - m = m - F^{-1}(1-x) \quad (7.17)$$

zaključuje se

$$\sum_{i=1}^n \tilde{r}_i = -\frac{nm}{\sigma} + \frac{1}{\sigma} \left\{ \sum_{i=1}^{\lfloor n/2 \rfloor} \left[F^{-1}\left(\frac{i-1}{n-1}\right) - F^{-1}\left(1 - \frac{i-1}{n-1}\right) \right] + \sum_{i=\lfloor n/2 \rfloor+1}^{\lfloor (n+1)/2 \rfloor} F^{-1}(0.5) \right\} = 0 \quad (7.18)$$

Na osnovu relacija (7.14) i (7.18) zaključuje se $b(\hat{m}) = 0$, što znači da je procena srednje vrednosti nepomerena u slučaju simetričnih raspodela. U slučaju nesimetričnih raspodela, generalno govoreći, estimator (7.13) poseduje pomeraj, dok u slučaju simetričnih raspodela, razlika između procene i prave srednje vrednosti iznosi:

$$\hat{m} - m = \frac{\sum_{i=1}^n (y_i - m) \sum_{i=1}^n \tilde{r}_i^2}{n \sum_{i=1}^n \tilde{r}_i^2} \quad (7.19)$$

na osnovu čega varijansa greške estimacije iznosi:

$$\sigma_{\hat{m}}^2 = E\left\{(\hat{m} - m)^2\right\} = \frac{\sigma^2}{n} \quad (7.20)$$

gde je $\sigma^2 = E\left\{(y - m)^2\right\}$. Na osnovu poslednje relacije se zaključuje da estimator srednje vrednosti (7.13) u slučaju simetričnih raspodela poseduje osobinu konzistentnosti.

Sa druge strane, na osnovu (7.11) i (7.12) estimator standardne devijacije se može napisati u sledećem obliku:

$$\begin{aligned}
 \hat{\sigma} &= \frac{n \sum_{i=1}^n \tilde{r}_i E\{y_i\} - \sum_{i=1}^n \tilde{r}_i \sum_{i=1}^n E\{y_i\}}{n \sum_{i=1}^n \tilde{r}_i^2 - \left(\sum_{i=1}^n \tilde{r}_i \right)^2} = \frac{n \sum_{i=1}^n \tilde{r}_i m_{y/i} - \sum_{i=1}^n \tilde{r}_i \sum_{i=1}^n m_{y/i}}{n \sum_{i=1}^n \tilde{r}_i^2 - \left(\sum_{i=1}^n \tilde{r}_i \right)^2} = \frac{n \sum_{i=1}^n \tilde{r}_i m_{y/i} - \sum_{i=1}^n \tilde{r}_i nm}{n \sum_{i=1}^n \tilde{r}_i^2 - \left(\sum_{i=1}^n \tilde{r}_i \right)^2} \\
 &= \frac{n \sum_{i=1}^n \tilde{r}_i (m_{y/i} - m)}{n \sum_{i=1}^n \tilde{r}_i^2 - \left(\sum_{i=1}^n \tilde{r}_i \right)^2} \quad (7.22)
 \end{aligned}$$

Pri tome je upotrebljena oznaka $E\{y_i\} = m_{y/i}$ i iskorišćena činjenica da je $\sum_{i=1}^n m_{y/i} = nm$. Na osnovu relacija (7.9) i (7.10), deterministički član $\tilde{r}_i$ iznosi

$$\tilde{r}_i = \frac{1}{\sigma} \left[F_n^{-1} \left(\frac{i-1}{n-1} \right) - m \right] \quad (7.23)$$

Smenjujući relaciju (7.23) u (7.22) i primenjujući operator matematičkog očekivanja dobija se:

$$E\{\hat{\sigma}\} = \sigma \frac{n \sum_{i=1}^n \left[F_n^{-1} \left(\frac{i-1}{n-1} \right) - m \right] (m_{y/i} - m)}{n \sum_{i=1}^n \left[F_n^{-1} \left(\frac{i-1}{n-1} \right) - m \right]^2 - \left\{ \sum_{i=1}^n \left[F_n^{-1} \left(\frac{i-1}{n-1} \right) - m \right] \right\}^2} \quad (7.24)$$

Pod pretpostavkom simetrične funkcije raspodele F , poslednja relacija postaje:

$$E\{\hat{\sigma}\} = \frac{n \sum_{i=1}^n \tilde{r}_i (m_{y/i} - m)}{n \sum_{i=1}^n \tilde{r}_i^2} = \sigma \frac{\sum_{i=1}^n \left[F^{-1} \left(\frac{i-1}{n-1} \right) - m \right] [m_{y/i} - m]}{\sum_{i=1}^n \left[F^{-1} \left(\frac{i-1}{n-1} \right) - m \right]^2} \quad (7.25)$$

U želji da ispitamo osobine ovog estimatora posmatrajmo relaciju (7.5). Ova relacija aproksimativno važi za svaku pojedinačnu vrednost argumenta y , pa stoga važi i za $y = m_{y/j}$, odnosno:

$$m_{i/m_{y/j}} = 1 + (n-1)F(m_{y/j}) \quad (7.26)$$

ili drugačije napisano:

$$m_{y/j} = F^{-1} \left(\frac{m_{i/m_{y/j}} - 1}{n-1} \right) \quad (7.27)$$

U slučaju velikog broja podataka n nad kojima se obrada vrši, može se napisati

$$m_{i/m_{y/j}} \approx j \Leftrightarrow m_{y/j} = F^{-1}\left(\frac{j-1}{n-1}\right) \quad (7.28)$$

na osnovu čega se zaključuje da je

$$E\{\hat{\sigma}\} \approx \sigma \quad (7.29)$$

odnosno da je estimator standardne devijacije aproksimativno nepomeren. Generalno govoreći, pomeraj estimatora (7.21) nije jednak nuli, odnosno $b(\hat{\sigma}) = E\{\hat{\sigma}\} - \sigma \neq 0$. Međutim, relativni pomeraj $b(\hat{\sigma})/\sigma$ zavisi isključivo od tipa normalizovane funkcije raspodele $F_n(\cdot)$ i od broja podataka n [17,19,20].

7.2 FORMULACIJA PROBLEMA PARAMETARSKE ESTIMACIJE

U radu je dalje razmatran problem identifikacije linearnog dinamičkog sistema estimacijom parametara na osnovu merenja ulaznog i izlaznog signala. Stoga, posmatrajmo apstraktni linearni, dinamički, stacionarni, vremenski diskretni sistem, koji se može modelirati diferencnom jednačinom sa konstantnim parametrima:

$$y(i) = -\sum_{k=1}^n a_k y(i-k) + \sum_{k=1}^m b_k u(i-k) + \xi(i) \quad (7.30)$$

pri čemu su $y(i) \in R^1$, $u(i) \in R^1$, $\xi(i) \in R^1$ redom sekvenca merenja izlaza sistema, merljivi signal na ulazu i stohastički ulaz, dok konstante $a_i, i=1, \dots, n$ i $b_j, j=1, \dots, m$ predstavljaju parametre sistema. Pretpostavimo još da je sekvenca stohastičkog ulaza $\{\xi(i)\}$ stohastički proces generisan u okviru sub-sigma algebre $\{F_i\}$, i da su konstante m i n unapred poznate ($0 < m \leq n$). Funkcija gustine verovatnoće realizacija procesa $\xi(i)$ nije u potpunosti poznata, pri čemu je apriorno znanje o ovoj raspodeli vezano za klasu P mogućih simetričnih raspodela sa nultom srednjom vrednošću, kao što je pretpostavljeno u šestom poglavlju.

Uvodeći operator pomeranja unazad $q^{-k}y(i) = y(i-k)$, relacija (7.30) se može napisati u sledećoj ARX polinomijalnoj formi

$$A(q^{-1})y(i) = B(q^{-1})u(i) + \xi(i) \quad (7.31)$$

gde su

$$A(q^{-1}) = 1 + \sum_{k=1}^n a_k q^{-k}, \quad B(q^{-1}) = \sum_{k=1}^m b_k q^{-k} \quad (7.32)$$

takozvani karakteristični i kontrolni polinomi. Relacija se može napisati i u linearnoj regresionoj formi

$$y(i) = Z^T(i) \Theta + \xi(i) \quad (7.33)$$

gde regresioni vektor $Z^T(i) = [-y(i-1), \dots, -y(i-n), u(i-1), \dots, u(i-m)]$ predstavlja vektor odbiraka ulaznog i izlaznog signala a $\Theta^T = [a_1 \dots a_n \ b_1 \dots b_m]$ predstavlja vektor konstantnih parametara sistema. Problem rekurzivne identifikacije sistema opisanog relacijom (7.33) je zapravo problem procene nepoznatih parametara sadržanih u vektoru Θ u realnom vremenu, a na osnovu merenja signala na ulazu i izlazu sistema, kao što je definisano u trećem poglavlju. Na osnovu nepotpunog apriornog znanja o prirodi stohastičkog ulaza (poremećaja) može se konstruisati min-max optimalni robusni algoritam identifikacije, koji minimizira usvojeni indeks performanse $V(T, p)$ koji predstavlja asimptotsku kovarijacionu matricu greške estimacije, pri čemu je izabrana forma estimatora T iz klase mogućih estimatora t a za neku funkciju gustine verovatnoće poremećaja p iz klase P [50,51,52]. Ovaj problem detaljno je razmatran u šestom poglavlju. Kao što je pokazano u ovom poglavlju, mi biramo formu estimatora $T \in \tau$, funkcija gustine verovatnoće poremećaja $p \in P$ je određena prirodom sistema i prirodom merenja dok $V(T, p)$ predstavlja konačni rezultat takve strategije. Ovakav izbor ima sedlastu tačku (T_0, p_0) ukoliko estimator T_0 i funkcija gustina verovatnoće p_0 zadovoljavaju

$$\min_{T \in \tau} \max_{p \in P} V(T, p) = V(T_0, p_0) = \max_{p \in P} \min_{T \in \tau} V(T, p) \quad (7.34)$$

Kao što je navedeno u šestom poglavlju, ovakav estimator T_0 se označava kao min-max optimalni robusni estimator, dok se p_0 naziva najnepoželjnijom funkcijom gustine verovatnoće.

Polazeći od metodologije izložene u šestom poglavlju, ukoliko je klasa estimatora τ izabrana u skupu rekurzivnih estimatora tipa stohastičkog gradijenta, tada se rekurzivna relacija kojom se estimiraju parametri vektora Q može napisati u formi

$$\hat{\Theta}(i) = \hat{\Theta}(i-1) + \Gamma(i) Z(i) \psi(v(i, \hat{\Theta}(i-1))) \quad (7.35)$$

gde je $v(i, \hat{\Theta}) = y(i) - \hat{\Theta}^T Z(i)$ greška predikcije ili rezidual, dok je min-max optimalni robusni estimator T_0 definisan sledećom relacijom

$$\psi(\cdot) = \psi_0(\cdot) = -[\log p_0(\cdot)]' \quad (7.36)$$

$$\Gamma(i) = \Gamma_0(i) = [i E \{ \psi_0'(\cdot) \} E \{ Z(i) Z^T(i) \}]^{-1} \quad (7.37)$$

Najnepoželjnija funkcija gustine verovatnoće $p_0(\cdot)$ je ona funkcija koja minimizira Cramer-Rao-vu granicu u okviru posmatrane klase raspodela P . Na taj način, $\psi_0(\cdot)$ u relaciji (7.36) predstavlja funkciju tipa maksimalne verodostojnosti (ML) koja odgovara specifičnoj raspodeli p_0 , a estimaciona procedura (7.36) se svodi na ML tip identifikacije. Međutim, problem je odrediti najnepoželjniju funkciju gustine verovatnoće $p_0(\cdot)$, jer se u tom cilju mora rešiti varijacioni problem koji nije klasičnog tipa i koji se uglavnom rešava numeričkim metodama. U šestom poglavlju je pokazano da se ovaj problem može rešiti analitički samo u aproksimativno, pri čemu se problem onda svodi na minimizaciju Fisher-ove količine informacija $I(p) = E\{(p'/p)^2\}$ u okviru zadate klase raspodela P . Različiti primeri za klase P i odgovarajuće funkcije gustine verovatnoće $p_0(\cdot)$ koje minimiziraju Fisher-ovu količinu informacija $I(p)$ u okviru klase P , navedeni su u šestom poglavlju. Međutim, praktični izbor težinske matrice $\Gamma_0(\cdot)$ u relaciji (7.37) nije moguć bez poznavanja stvarne raspodele stohastičkog ulaza. Stoga se u okviru robusne estimacije zahteva takav estimator koji će biti efikasan u slučaju kada je raspodela poremećaja normalna, ali istovremeno treba da ima dobre osobine i u slučaju kada raspodela poremećaja značajno odstupa od normalne u smislu posedovanja takozvanih 'zadebljelih repova' koji rezultuju ekstremno velikim vrednostima u sekvenci merenja izlaza. Otuda se, u cilju dizajna praktično primenjivog rekurzivnog robusnog estimatora, mora ići dalje u aproksimaciji min-max optimalnog rešenja (7.35)-(7.39). Jedan od mogućih pristupa se zasniva na algoritmu stohastičkog gradijenta (7.35) i ovaj je metod prikazan u sledećem odeljku rada. Rekurzivni estimator dat relacijom (7.35) koji za konkretnu funkciju ψ ne koristi ML tip funkcije dat relacijom (7.36), se naziva aproksimativnim rekurzivnim estimatorom maksimalne verodostojnosti ili skraćeno M-estimatorom [40].

7.3. ROBUSNI REKURZIVNI ALGORITAM TIP STOHAISTIČKOG GRADIJENTA

U cilju realizacije min-max optimalnog robusnog rekurzivnog algoritma (7.35)-(7.37), neophodno je izvršiti izbor nelinearne funkcije $\psi(\cdot)$ sa osnovnim ciljem da na neki način prepozna neregularna merenja, takozvane outlier-e, i da smanjuje njihov uticaj na kvalitet identifikacije parametara sistema. Nelinearna funkcija, koja se u literaturi posvećenoj robusnoj estimaciji naziva i funkcijom uticaja [**] mora biti izabrana tako da obezbedi visoku efikasnost u slučaju kada je raspodela mernog šuma nominalna, odnosno Gaussovska, ali da ovu osobinu sačuva i u slučaju prisustva neregularnih merenja, outliera, iz prethodno definisane i strateški izabrane klase raspodela. Pogodno je da, između ostalog, ova funkcija ima i osobine ograničenosti i neprekidnosti. Ograničenost funkcije $\psi(\cdot)$ obezbeđuje da bilo koja pojedinačna opservacija ne može imati neograničen uticaj na kvalitet estimacije, dok neprekidnost funkcije omogućava da se efekti zaokruživanja i kvantizacije mogu

zanemariti. Dalje, funkcija $\psi(z)$ mora biti linearna za male vrednosti argumenta, te se u literaturi uobičajeno usvaja funkcija oblika [40]

$$\psi(z) = \min\left(\frac{z}{\sigma^2}, \frac{k}{\sigma^2}\right) \text{sgn}(z) \quad (7.38)$$

gde je sa σ označena standardna devijacija mernog šuma, a parametar k predstavlja konstantu koju treba izabrati sa ciljem da se postigne željena efikasnost algoritma u slučaju nominalnog modela mernog šuma. Kako je, dalje, vrednost varijanse σ^2 uglavnom nepoznata, neophodno ju je na neki način proceniti. Popularna i često korišćena jednostavna forma robusne procene varijanse σ^2 jeste mediana [39]:

$$d = \text{median}\{|y(i) - \text{median}\{y(i)\}|\} / 0.6745 \quad (7.39)$$

Ostali, primenjivi tipovi nelinearne funkcije $\psi(\cdot)$ se mogu naći u literaturi [4,33,39,40,75,76,89].

Opisani rekursivni robusni estimator ima sledeća svojstva: Kao što je pokazano u šestom poglavlju, ukoliko se za nelinearnu funkciju uticaja izabere sledeća forma:

$$\psi(\cdot) = \psi^*(\cdot) = -[\log(p^*(\cdot))]' ; \quad p^* = \arg \min_{p \in P} I(p) \quad (7.40)$$

ovakav izbor ne zadovoljava optimalnost u min-max smislu (7.34), već on minimizira uslovnu kovarijacionu matricu $E\left\{\left[\hat{\Theta}(i) - \hat{\Theta}(i-1)\right]\left[\hat{\Theta}(i) - \hat{\Theta}(i-1)\right]^T \middle| F_{i-1}\right\}$ za slučaj najgore raspodele p^* u (7.40). Dalje, na osnovu rezultata izloženih u šestom poglavlju, ideja uvođenja nelinearne funkcije uticaja $\psi(\cdot)$ u formi relacije (7.38) odgovara obliku funkcije $\psi^*(\cdot)$ iz relacije (7.40) za slučaj kada klasa P definiše takozvanu ε – kontaminiranu familiju raspodela, definisanu kao:

$$P = P_\varepsilon = \left\{p \middle| p = (1-\varepsilon)N(0, \sigma^2) + \varepsilon h\right\}, \quad \varepsilon \in [0,1] \quad (7.41)$$

gde je $h(\cdot)$ simetrična funkcija gustine verovatnoće nulte srednje vrednosti, dok $N(0, \sigma^2)$ predstavlja Gaussovu raspodelu nulte srednje vrednosti i varijanse σ^2 . Najnepoželjnija raspodela p^* iz ove klase, sa takozvanim 'otežalim repovima' rezultuje u $I^{-1}(p^*) = 2(1-\varepsilon)\text{erf}(k)\sigma^2$, gde je sa erf označena funkcija greške (videti poglavlje šest).

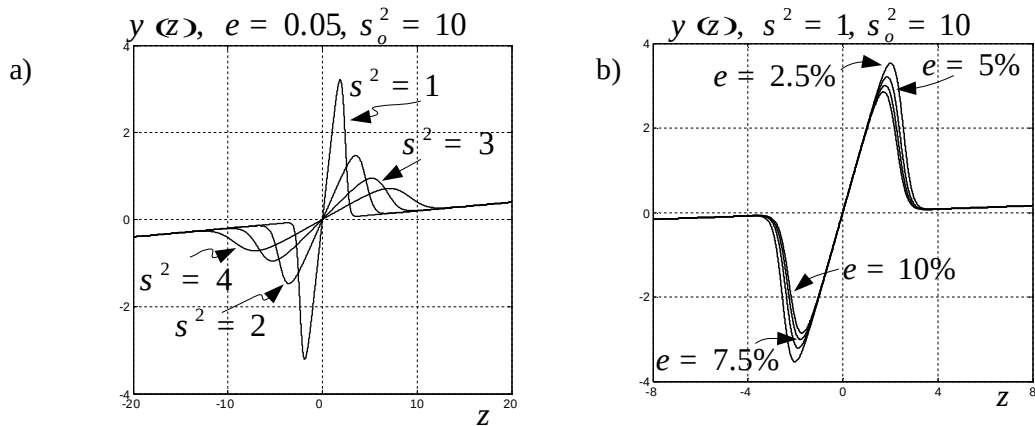
7.4. ADAPTACIJA ROBUSNOG REKURZIVNI ALGORITAM TIP A STOHAŠTIČKOG GRADIJENTA

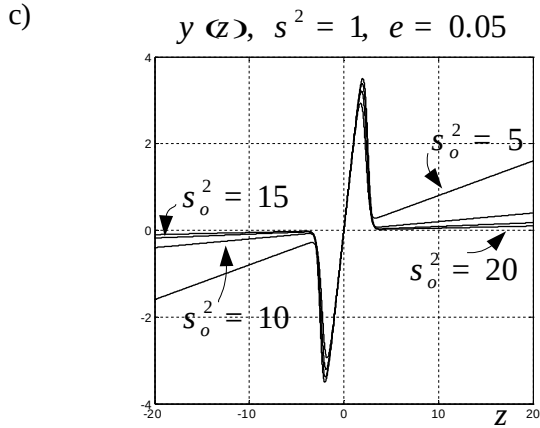
ε – kontaminirana familija raspodela (7.41) je od izuzetnog značaja za oblast robusne estimacije jer ona modelira veliki broj različitih aplikacija u kojima je sporadična pojava neregularnih merenja velikih intenziteta prisutna [4]. Šta više, u takvim situacijama je svrsishodno za kontaminantu $h(\cdot)$ iz relacije (7.41) usvojiti Gaussovu raspodelu nulte srednje vrednosti i varijanse koja je značajno veća od nominalne [17-20]. Shodno tome, usvaja se model funkcije gustine verovatnoće mernog šuma u sledećoj formi:

$$p = (1 - \varepsilon)N(0, \sigma^2) + \varepsilon N(0, \sigma_0^2) \quad (7.42)$$

koja je okarakterisana sa tri nepoznata parametra: intenzitetom kontaminacije $\varepsilon \in [0,1]$, nominalnom varijansom šuma regularnih merenja σ^2 i varijansom outlier-a $\sigma_0^2 \gg \sigma^2$. U nedostatku apriornog znanja o ovim parametrima, uobičajeno se usvaja robusni estimator (7.35)-(7.37) sa funkcijom uticaja (7.38), pri čemu se procena varijanse zamenjuje medijanom (7.39) a parametar k se bira tako da iznosi $1.42\sigma_d$ u cilju obezbeđivanja 95% efikasnosti estimatora. Ovakav izbor je krajnje opravdan u nedostatku apriornih informacija o parametrima raspodele mernog šuma, međutim ima i značajnih nedostataka. Prvi nedostatak jeste da je nemoguće dokazati svojstva estimatora (7.39), pa se samim tim i željena efikasnost dovodi u pitanje. Drugi značajan nedostatak jeste taj da je intuitivno jasno da funkcija uticaja mora biti podložna promenama intenziteta kontaminacije i varijansi outlier-a, što se izborom funkcije (7.38) ne obezbeđuje.

U cilju ilustracije značaja parametara ε , σ^2 i σ_0^2 u modelu šuma (7.42), na izbor funkcije uticaja, pogledajmo na slikama 7.1.a,b i c kako bi teorijski izgledale ove funkcije primenom metoda maksimalne verodostojnosti ($\psi(\cdot) = [-\ln(p)]$). Na ovim slikama su skicirane idealne ML funkcije uticaja pri čemu je na slici 7.1.a. usvojeno $\varepsilon = 5\%$, $\sigma_0^2 = 10$ dok varijansa nominalnog modela uzima različite vrednosti $\sigma^2 \in \{1,2,3,4\}$. Na slici 7.1.b su varijanse nominalnog modela i outlier-a konstantne $\sigma^2 = 1$, $\sigma_0^2 = 10$, dok se intenzitet kontaminacije menja $\varepsilon[\%] \in \{2.5, 5, 7.5, 10\}$. Na slici 7.1.c su varijansa nominalnog modela $\sigma^2 = 1$ i intenzitet kontaminacije $\varepsilon = 5\%$ konstantni, dok se varijansa outlier-a menja $\sigma_0^2 \in \{5, 10, 15, 20\}$.





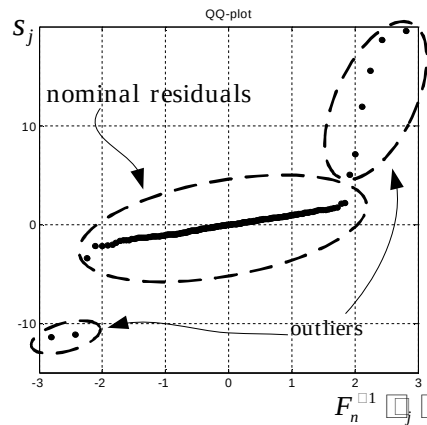
Slika 7.1: ML funkcije uticaja za različite vrednosti parametar ε , σ^2 i σ_0^2 u modelu šuma (7.42): a) za različite vrednosti varijanse σ^2 , b) za različite vrednosti intenziteta kontaminacije ε , c) za različite vrednosti varijanse σ_0^2

Na osnovu prikazanih funkcija uticaja mogu se izvesti sledeći zaključci. Svaka od prikazanih formi funkcija ima izrazita tri regiona. Prvi region jeste oblast za male vrednosti argumenta (sa obe strane koordinatnog početka) u kojima je funkcija uticaja gotovo idealno linearna. Ovakav oblik funkcije je u potpunom skladu sa teorijskim rezultatom koji kaže da za male vrednosti reziduala prisustvo outlier-a nije detektovano, te je nominalni model važeći, a shodno njegovoj Gaussovskeoj prirodi, linearna forma ML funkcije uticaja optimizira varijansu greške estimacije. Za nagib linearnog segmenta u ovom regionu odgovorna je isključivo varijansa nominalnog modela σ^2 . Treći region je zadužen za prisustvo outlier-a u strukturi mernog signala i on je takođe linearan, ali u opsezima intenzivnih reziduala velikih, pozitivnih ili negativnih, vrednosti. Na osnovu prikazanih slika je očigledno da je za nagib linearnih segmenata u ovom, trećem, regionu, odgovorna varijansa outlier-a σ_0^2 . Takođe se uočava da je sa povećanjem ove varijanse i nagib linearnog segmenta u trećem regionu sve manji, što je u skladu sa intencijom funkcije uticaja da takva, loša, merenja potisne. Konačno, postoji i druga, prelazna oblast između prve i treće, koja u najvećoj meri zavisi od intenziteta kontaminacije ε i varijanse nominalnog modela σ^2 , dok je prilično neosetljiva na varijansu outlier-a σ_0^2 . Ovo je oblast u kojoj se sa gotovo jednakim verovatnoćama mogu pojaviti i regularna i loša merenja, otuda funkcija uticaja u ovoj oblasti ima negativni gradijent, jer se sa povećavanjem vrednosti reziduala verovatnoća regularnog merenja neprekidno smanjuje.

Imajući napred navedeno u vidu, postaje jasno da bi estimacija parametara ε , σ^2 i σ_0^2 u velikoj meri pomogla u pravilnom izboru funkcije uticaja, što bi konačno rezultovalo kvalitetnim estimatorom parametara sistema koji se identifikuje. Pogodan teorijski okvir za estimaciju ovih parametara jeste QQ-kriva [4] koja se uobičajeno u literaturi koristi kao mera valjanosti pretpostavke da raspodela slučajne promenljive odgovara usvojenoj. Pretpostavimo da nam je na raspolaganju sekvenca poslednjih N reziduala:

$$v(k, \hat{\Theta}) = y(k) - \hat{\Theta}^T Z(k), \quad k = i - N + 1, \dots, i \quad (7.43)$$

gde je sa i označen tekući trenutak, $y(k)$ je merenje u k -tom trenutku, dok su $\hat{\theta}$ i $Z(k)$ definisani relacijama (7.33) i (7.36). Ukoliko za tako dobijenu sekvencu reziduala formiramo QQ-krivu (ilustrativni primer je prikazan na slici 2), gde se na apscisu nanose vrednosti $F_n^{-1}(\rho_j)$, $j=1, \dots, N$ gde je sa $F_n^{-1}(\cdot)$ označena inverzna funkcija raspodele za Gausovsku raspodelu nulte srednje vrednosti i jedinične varijanse a $\rho_j = (j-1)/(N-1)$, a na ordinatu vrednosti s_j , $j = 1, \dots, N$ koje predstavljaju sortirane, u rastući niz, vrednosti reziduala $v(k, \hat{\theta})$, $k = i - N + 1, \dots, i$, dobićemo grafički prikaz u kome su sadržane informacije o sva tri tražena a nepoznata parametra $(\varepsilon, \sigma^2, \sigma_0^2)$. Primetimo da ovaj postupak zahteva da se originalni sortirani reziduali centriraju u normalizuju, kako bi imali nultu srednju vrednost i jediničnu varijansu, čime se ne menja rang reziduala u sortiranoj sekvenci.



Slika 7.2: Primer QQ-krive za sekvencu reziduala shodno kontaminiranoj normalnoj raspodeli šuma merenja (7.42)

Sa slike 7.2 se vidi da će najveći broj odbiraka reziduala (teorijski njih $(1-\varepsilon)N$) biti skoncentrisano duž linearnog segmenta, dok će retki outlier-i (teorijski njih εN) značajno odstupati od ove linearne zavisnosti. Mogući postupak za određivanje gornje i donje granice QQ-krive u okviru kojih se merenja mogu smatrati regularnim opisan je sledećom optimizacionom procedurom [19,20]:

$$\begin{aligned} b_i &= \arg \min_b J_i(b); \quad J_i(b) = \left| b - F_n^{-1}(-p + F_n(b/i)) \right|; \\ a_i &= F_n^{-1}(F_n(b_i/i) - p/i); \quad i = 1, \dots, N \end{aligned} \quad (7.44)$$

pri čemu se parametar p usvaja eksperimentalno u opsegu $[0.7, 0.8]$. Naime, parametri α_i i β_i u proceduri (7.44) izabrani su tako da je masa verovatnoće u intervalu (α_i, β_i) jednaka unapred specificiranoj vrednosti p , koja je tako odabrana da se obezbedi efikasna robusnost, odnosno da algoritam estimacije poseduje visoku efikasnost pri nominalnoj normalnoj funkciji gustine verovatnoće u (7.42), ali da takođe efikasno eliminiše sporadične realizacije velikog intenziteta koje su

generisane kontaminirajućom funkcijom gustine verovatnoće u (7.42). Na taj način, parametri α_i i β_i moraju da zadovoljavaju uslov:

$$\int_{\alpha_i}^{\beta_i} f(y/i) dy = F(\beta_i/i) - F(\alpha_i/i) = p \quad (7.45)$$

podrazumevajući da je $\alpha_i < m_{y/i} < \beta_i$, pri čemu $F(\cdot/i)$ označava odgovarajuću aposteriornu funkciju raspodele verovatnoće, a $f(\cdot/i)$ njenu funkciju gustine verovatnoće. Relacija (7.45) ne definiše jednoznačno dva nepoznata parametra, te je neophodno da se uvede dodatna relacija u obliku nekog dodatnog kriterijuma. Jedan mogući prirodan izbor je da se odredi rešenje relacije (7.45) pod uslovom da je minimalno rastojanje

$$d = d(\alpha_i, \beta_i) = |\alpha_i - \beta_i| \quad (7.46)$$

Na taj način, regularne observacije koje su generisane nominalnom normalnom funkcijom gustine verovatnoće u modelu (7.42) biće linearno transformisane nakon prolaska kroz funkciju verodostojnosti u estimatoru parametara. Ovakav postupak formalno je zapisan u obliku relacije (7.44). Primetimo da je verovatnoća pogrešne odluke da se outlier proglasi za regularan podatak srazmerno sa rastojanjem (7.46), se te na ovaj način minimizira verovatnoća lažnog alarma. Postavljeni minimizacioni problem je nelinearan i može se rešiti numerički, na primer korišćenjem Newton-Raphosonovog postupka koji je opisan u četvrtom poglavlju

$$\hat{\beta}_i(k+1) = \hat{\beta}_i(k) - [J_i''(\hat{\beta}_i(k))]^{-1} [J_i'(\hat{\beta}_i(k))] \quad (7.47)$$

za $k = 1, 2, \dots$ i polazeći od nekog početnog događanja $\hat{\beta}_i(0)$. Prebrojavanjem broja tačaka N_o na QQ-krivoj koje ne zadovoljavaju uslov

$$\alpha_i < s_i < \beta_i, \quad i = 1, \dots, N \quad (7.48)$$

i deljenjem sa ukupnim brojem podataka N , dobija se procena intenziteta kontaminacije ε :

$$\hat{\varepsilon} = \frac{N_o}{N} \quad (7.49)$$

Shodno tome, skup svih sortiranih reziduala s_j , $j = 1, \dots, N$ možemo podeliti u dva podskupa s_j^r , $j = 1, \dots, N - N_o$ koji su regularni i zadovolji su uslov (7.48) i podskup s_j^o , $j = 1, \dots, N_o$ koji su očigledno pod uticajem outlier-a i nisu zadovoljili uslov (7.49). Takođe je pokazano da u (7.10) je nagib linearnog segmenta koji modelira najveći broj tačaka QQ-krive zapravo procena standardne devijacije nominalnog modela. Do ove procene se može doći primenom metoda najmanjih kvadrata (7.11), odnosno:

$$\begin{bmatrix} \hat{m} \\ \hat{\sigma} \end{bmatrix} = (\sum^T \sum)^{-1} \sum^T S$$

$$\sum^T = \begin{bmatrix} 1 & 1 & \dots & 1 \\ F_n^{-1}(\rho_1) & F_n^{-1}(\rho_2) & \dots & F_n^{-1}(\rho_{N-N_0}) \end{bmatrix}; \quad S^T = [s_1^r \quad s_2^r \quad \dots \quad s_{N-N_0}^r] \quad (7.50)$$

gde je sa $\hat{m}$ označena procena srednje vrednosti nominalnog modela i po pretpostavci bi trebalo da bude jednaka nuli, sa $\hat{\sigma}$ je označena procena standardne devijacije nominalnog modela i u poglavlju 7.1 je pokazano da je ovakva procena standardne devijacije nominalnog modela pomerena, pri čemu relativni pomeraj zavisi od broja merenja u sekvenci N i da je za $N > 100$ relativni pomeraj manji od par procenata. Konačno, izdvajanjem onih reziduala koji su pod nedvosmislenim uticajem outlier-a i procenom njihove varijanse dobija se traženi parametar σ_0^2 :

$$\hat{\sigma}_0^2 = \frac{1}{N_0} \sum_{j=1}^{N_0} \left(s_j^0 - \frac{1}{N_0} \sum_{k=1}^{N_0} s_k^0 \right)^2 \quad (7.51)$$

Alternativno se vatijansa (7.51) može proceniti primenom algoritma (7.50) na podatke klasifikovane u (7.48) kao neregularne, tj. outlier.

Periodičnom procenom ova tri parametra i njihovim uvrštavanjem u ML funkciju uticaja

$$\psi(\cdot) = -[\log(p(\cdot))]'; \quad p = (1-\varepsilon)N(0, \sigma^2) + \varepsilon N(0, \sigma_o^2) \quad (7.52)$$

robustna rekurzivna identifikacija sistema (7.35-7.38) postaje adaptivna, čime se značajno povećava kvalitet identifikacije, što će biti ilustrovano sledećim numeričkim primerom.

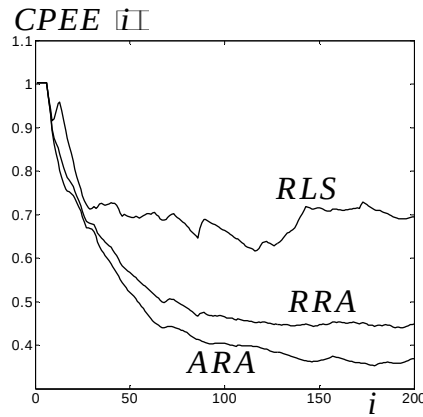
7.5. REZULTATI SIMULACIJE

U cilju analize robusnosti i efikasnosti algoritma, izvršene je sledeći eksperiment. Posmatrajmo model sistema četvrtog reda koji je dat u formi relacije (7.33): $Z^T(i) = [-y(i-1) \dots -y(i-4) u(i-1)]$, gde je vektor parametara $\Theta^T = [-1.48 \quad 1.28 \quad -1.09 \quad 0.37 \quad 2]$. Sekvenca $\{u(i)\}$ je generisana kao bela sekvenca sa Gaussovskom raspodelom nulte srednje vrednosti i jedinične varijanse, dok je pomećaj $\xi(i)$ generisan kao ε -kontaminirana funkcija gustine verovatnoće shodno relaciji (7.42), sa intenzitetom $\varepsilon = 10\%$, nominalnom varijansom $\sigma^2 = 1$ i varijansom outliera $\sigma_o^2 = 10$. Testirani su sledeći algoritmi: 1) rekurzivni algoritam najmanjih kvadrata označen kao RLS; 2) rekurzivni robusni algoritam definisan relacijama (7.35)-(7.38) označen kao RRA i konačno 3) adaptivni robusni rekurzivni algoritam definisan relacijama (7.35)-(7.37) i funkcijama adaptacije na osnovu

relacija (7.48)-(7.51), označen kao ARA. Kao mera poredjenja algoritama uvedena je kumulativna relativna greška estimacije parametara:

$$CPEE(i) = \frac{1}{i} \sum_{k=1}^i \sum_{l=1}^5 \left| \frac{\hat{\Theta}_l(k) - \Theta_l}{\Theta_l} \right| \quad (7.53)$$

gde je sa Θ_l , $l = 1, \dots, 5$ označen l -ti element u vektoru nepoznatih parametara Q , dok je sa $\hat{Q}_l(k)$ označena procena ovog parametra u k -toj iteraciji. Na slici 7.3 su prikazane vrednosti ovih kriterijuma za navedene algoritme.



Slika 7.3: Vrednosti kumulativnog kriterijuma estimacije parametara CPEE za različite algoritme

Slika 7.3 ilustruje efikasnost predloženog adaptivnog robusnog algoritma u uslovima prisustva impulsnog šuma u mernom signalu. Uočljivo je da ovaj algoritam manje osetljiv na sporadične pojave outlier-a što rezultuje ukupnom manjom kumulativnom greškom. Takođe je uočljivo i da je RRA algoritam manje osetljiv na ovakvu vrstu mernog šuma u odnosu na klasični RLS algoritam, ali s obzirom da ne poseduje odgovarajuće informacije o statistikama kontaminirane funkcije gustine verovatnoće mernog šuma, parametar k u relaciji (7.38) je fiksiran, te je stoga i manje efikasan u odnosu na predloženi adaptivni robusni (ARA) algoritam. Predloženi ARA algoritam pokazuje visoku efikasnost estimacije za vrednosti intenziteta kontaminacije $\varepsilon < 20\%$, nezavisno od varijanse outlier-a σ_o^2 . U situacijama kada ovaj parametar postane veći od 0.2, pretpostavka o sporadičnoj pojavi više outlier-a više nije održiva, te se stoga i procene parametara degradiraju.

8. ZAKLJUČAK

U doktorskom radu pod naslovom „Analiza jedne klase rekurzivnih stohastičkih algoritama Gauss-Newtonovog tipa sa primenama u parametarskoj identifikaciji FIR i IIR digitalnih sistema“ razmatrana je parametarska identifikacija modela sistema određene strukture, koja predstavlja linearni vremenski diskretni i vremenski invarijantan stohastički sistem. Inače, kao što je navedeno u uvodu i detaljno razmatrano u prvom poglavlju, modelovanje predstavlja integralni deo nauke i tehnike, koji obuhvata sve oblasti ljudske aktivnosti. Sam pojam modelovanja povezan je direktno sa pojmom procesa ili sistema, pri čemu sistem predstavlja subjektivni pojam i definisan je usvojenim granicama koje obuhvataju samo one osobine realnog procesa koje su od interesa za datu primenu. Na taj način, glavni zadatak modelovanja je da osvetli osnovne osobine realnog procesa koji se posmatra i da ih prevede na neki apstraktan jezik, pri čemu je uobičajeno da se u tu svrhu koristi jezik matematike. Iako ne postoji generalna teorija modeliranja, u literaturi se mogu uočiti dva pristupa sintezi modela sistema. Prvi od njih je deduktivni, koji dovodi do kvalitativnog modela, kod koga su na osnovu poznatih idealizovanih fizičkih zakona definisane matematičke relacije između relevantnih fizičkih varijabli, a ukoliko su i vrednosti parametara u modelu egzaktno poznate takav model se u engleskoj literaturi naziva “white box” model, odnosno na srpskom jeziku model bele kutije. Tipičan primer “white box” modela predstavljaju električna kola i sistemi.

Drugi pristup je induktivni, a polazi od unapred usvajane strukture modela, koja obično predstavlja linearan vremenski invarijantan diskretni sistem. Pošto konačan broj parametara u ovakvoj strukturi nema nikakvo fizičko tumačenje, ovakav model naziva se na engleskom jeziku “black box” model, odnosno na srpskom jeziku model crne kutije. Važan korak u ovakvom pristupu predstavlja estimacija parametara, na osnovu raspoloživih realnih merenja na posmatranom sistemu, a ovaj zadatak se u literaturi

naziva parametarska identifikacija. Ovakav pristup je vrlo popularan u prirodnim i tehničkim naukama, s obzirom da linearni kontinualni i diskretni modeli navedenog tipa vrlo dobro aproksimiraju ponašanje tipičnih industrijskih i tehnoloških procesa oko neke izabrane radne tačke.

U okviru teze dat je pregled trenutnog stanja u istraživanjima i primenama po pitanju parametarske identifikacije linearnih dinamičkih modela signala i sistema.

U tezi je predložena upotreba metoda parametarske identifikacije za usvojene reprezentacije modela sistema različite fizičke prirode, a koje predstavljaju digitalne vremenski invarijantne linearne stohastičke sisteme sa konačnim impulsnim odzivom (engleski Finite Impulse Response ili skraćeno FIR) i beskonačnim impulsnim odzivom (engleski Infinite Impulse Response ili skraćeno IIR). Kao generalni pristup za sintezu algoritama parametarske identifikacije u usvojenoj FIR ili IIR strukturi modela korišćen je metod greške predikcije. Zadatak parametarske identifikacije je postavljen kao problem optimizacije bez ograničavanja pogodno odabrane kriterijumske funkcije. U optimizacioni kriterijum je uključena i raspoloživa apriorna informacija o statističkim karakteristikama greške predikcije, posmatrane kao stohastička promenljiva, u obliku odgovarajuće klase raspodela poremećaja kojoj pripada i njegova stvarna ali nepoznata funkcija raspodele verovatnoće, a ovakva informacija je preslikana u formu nelinearne transformacije greške predikcije. U suštini, nelinearna transformacija reziduala merenja ili greške predikcije u ovakvom pristupu predstavlja funkciju verodostojnosti, koja je definisana u odnosu na najnepovoljniju funkciju gustine verovatnoće u okviru usvojene klase, u smislu minimalne Fišerove količine informacije. S obzirom da je tako dobijeni optimizacioni kriterijum nelinearna funkcija greške predikcije, kao sredstvo za više koračnu numeričku minimizaciju ovog kriterijuma korišćen je Gauss- Newtonov rekurzivni stohastički metod. Pokazano je da se ovaj metod pod određenim uslovima svodi na rekurzivni algoritam tipa stohastičkog gradijenta. Matrica pojačanja ovakvog algoritma određena je iz uslova minimuma totalne asimptotske varijanse greške estimacije, čime je obezbeđena maksimalno moguća asimptotska brzina konvergencije pod zadatim uslovima. Na taj način, predloženi pristup rezultuje u optimalne algoritme parametarske identifikacije na zadatoj klasi raspodela poremećaja. Predloženim postupkom je izbegnuta i proizvoljnost i subjektivnost pri izboru indeksa performanse i

algoritma parametarske identifikacije, a takođe su formalizovane i heurističke intervencije koje su uobičajene pri statističkoj analizi mernih podataka. Posebni značaj u navedenom pristupu ima teorijska analiza konvergencije, konzistentnosti i asimptotske brzine konvergencije procena parametara ka njihovim optimalnim vrednostima, koje odgovaraju minimumu usvojenog kriterijuma. Na bazi analize konvergencije došlo se do egzaktnih teorijskih uslova pod kojima procene parametara, koje su generisane rekurzivnim stohastičkim algoritmom, konvergiraju. Za razliku od postupaka simulacije, koji ne omogućava ekstenziju dobijenih rezultata na nove slučajeve koji nisu obuhvaćeni simulacijom, ovakav analitički rezultat ima univerzalnu vrednost i od korisnika se zahteva samo da proveri da li su izvedeni teorijski uslovi za konvergenciju algoritma parametarske identifikacije ispunjeni u razmatranom slučaju. Nedostatak predložene teorijske analize je što ima asimptotski karakter, te važi na uzorku neograničene dužine. U praksi ovi rezultati važe na konačnom ali velikom uzorku, pri čemu teorijska analiza ne daje odgovor koliko veliki uzorak treba da bude. Odgovor na ovo pitanje, kao i sagledavanje osobina algoritama parametarske identifikacije na uzorku konačne, ili čak male, dužine može se dobiti samo putem simulacije. S toga se jedino kombinacijom teorijske i eksperimentalne analize, putem simulacije, mogu u potpunosti sagledati osobine predloženih algoritama u uslovima bliskim realnosti, što je u ovom radu i sprovedeno.

Posebna pažnja je, takođe, posvećena praktičnim aspektima estimacije statističkih karakteristika slučajnih poremećaja u sistemu, koje nikad u praksi nisu tačno poznate. Za specijalan, ali za praksu vrlo važan, slučaj normalno raspodeljene merne populacije, koja je kontaminirana sa određenim procentom nekonzistentnih opservacija, koje se nazivaju greške merenja ili u anglosaksonskoj literaturi outlieri, predložena je procedura za estimaciju nepoznatog stepena kontaminacije i varijanse šuma merenja jednovremeno sa postupkom estimacije parametara. Estimacija nepoznatih statistika šuma merenja zasnovana je na prilazu QQ krive, koji je primenjen na memorisani rezidual merenja u okviru klizećeg prozora pogodne dužine. Za estimaciju parametara korišćeni su prethodno razmatrani rekurzivni stohastički algoritmi, čime je dobijena nova efikasna procedura za parametarsku identifikaciju. Ovim je omogućena sinteza nove klase efikasnih i praktično primenljivih rekurzivnih algoritama za estimaciju nepoznatih

parametara u usvojenoj FIR ili IIR strukturi modela sistema, kada je sistem podvrgnut dejstvu adaptivnog šuma čije statističke osobine nisu potpuno poznate. Ovako dobijeni modeli mogu naći široku primenu u raznorodnim zadacima analize tehnoloških procesa, njihove simulacije, kao i projektovanja sistema za obradu i prenos signala i upravljanje tehnološkim u industrijskim procesima. Drugim rečima, jedan od osnovnih pravaca daljih istraživanja predstavlja primenu izvedene nove klase algoritama parametarske identifikacije na konkretne probleme u različitim oblastima nauke i tehnologije, pre svega informaciono komunikacionih i upravljačkih tehnologija i složenih sistema.

Primetimo na kraju, da se osnovni naučno-istraživački rezultati u predloženoj doktorskoj disertaciji mogu svrstati u sledeće kategorije:

1. Sistematizacija brojnih postojećih rezultata u oblasti parametarske identifikacije modela signala i sistema.
2. Sistematizacija postojećih rezultata u oblasti robusne statistike, sa analizom mogućnosti njihove primene u zadacima parametarske identifikacije dinamičkih modela signala i sistema.
3. Sinteza originalnog rekurzivnog stohastičkog algoritma Gauss-Newtonovog tipa za estimaciju nepoznatih parametara u usvojenoj FIR ili IIR strukturi modela sistema, kada je sistem podvrgnut dejstvu adaptivnog šuma čije statističke osobine nisu potpuno poznate, ali je na raspolaganju izvesna apriorna informacija o ovom šumu u smislu date klase raspodele kojoj i stvarna raspodela šuma pripada.
4. Teorijska analiza konvergencije, konzistencije i asimptotske brzine konvergencije algoritma, na osnovu koje su izvedeni egzaktni analitički rezultati koji ukazuju na uslove konvergencije procena parametara ka njihovim optimalnim vrednostima, u smislu usvojenog kriterijuma i omogućavaju optimalan izbor matrice pojačanja algoritma, u smislu maksimalno moguće asimptotske brzine konvergencije u zadatim uslovima stohastičkog okruženja.
5. Sinteza praktično primeljivog adaptivnog algoritma parametarske identifikacije na bazi simultane estimacije parametara usvojene FIR ili IIR strukture modela i nepoznatih statističkih parametara normalno

raspodeljenog aditivnog šuma, na izlazu sistema, kontaminiranog retkim sporadičnim realizacijama velikog intenziteta, tzv. outlierima (ovakim mernim šumom je efekat delovanja stvarnog nepoznatog poremećaja u sistemu preslikan na njegov izlaz).

- 6.** Analiza praktične primenljivosti predloženih algoritama parametarske identifikacije na bazi simulacije u uslovima bliskim realnosti.

LITERATURA:

- [1] Aljančić S., *Uvod u Realnu i Funkcionalnu Analizu*, Građevinska Knjiga, 1979.
- [2] Anđelić T., *Matrice*, Zavod za Izdavanje Udžbenika, 1970.
- [3] Banjac Z.D., B.D. Kovačević, M.D. Veinović, M.M. Milosavljević, Robust adaptive filtering with variable forgetting factor, *WSEAS Transaction on Circuits and Systems*, Vol. 3, No.2, 2004, pp. 223-229.
- [4] Barnett V.D., T. V. Lewis, *Outliers in Statistical Data*, John Wiley, 1978.
- [5] Bellman R., *Introduction to Matrix Analysis*, McGraw-Hill, 1960.
- [6] Biernes H., *Robust Methods and Asymptotic Theory in Nonlinear Econometrics*, Springer Verlag, 1981.
- [7] Bosch P.P.J. van den, A.C. van der Klauw, *Modeling, Identification and Simulation of Dynamical Systems*, CRC Press, 1994.
- [8] Bird J., *Higher Engineering Mathematics*, Elsevier, 2006.
- [9] Bunke O., H. Bunke, *Nonlinear Regression, Functional Relations and Robust Methods*, John Wiley, 1987.
- [10] Buzi S., H.V. Poor, Channel estimation and multiuser detection in long-code DS/CDMA systems, *IEEE Journal on Selected Areas in Communications*, Vol. 19, No. 8, 2006, pp. 1476-1487.
- [11] Candy J.V., *Model-Based Signal Processing*, John Wiley, 2006.
- [12] Chan S., Y. Zou, A recursive least M-estimate algorithm for robust adaptive filtering in impulsive noise: fast algorithm and convergence performance analysis, *IEEE Transaction on Signal Processing*, Vol. 52, No. 4, 2004, pp. 975-991.
- [14] Dai H., N.K. Sinha, Robust recursive least-squares method with modified weights for bilinear system identification, *Proc. IEE*, Vol. 136, No. 6, 1988, pp. 122-126.
- [15] Đurović Ž.M., B.D. Kovačević, Robust estimation with unknown noise statistics, *IEEE Transaction on Automatic Control*, Vol. 44, No.6, 1999, pp. 1292-1296.
- [16] Đurović Ž.M., B.D. Kovačević, *Diskretni Signali i Sistemi*, Akademska Misao, 2004.
- [17] Đurović Ž.M., B.D. Kovačević, QQ-plot approach to robust Kalman filtering, *International Journal of Control*, Vol. 61, No. 4, 1995, pp. 837-857.
- [18] Đurović Ž.M., B.D. Kovačević, Bandwidth selection for kernel density estimation based on QQ-plot, *WSEAS Transaction on Circuits and Systems*, Vol. 3, No.3, 2004, pp. 679-687.
- [19] Đurović Ž.M., I.P. Kovačević, B.D. Kovačević, A new approach to robustified adaptive estimation of channels parameters, *Proceedings of the 12-th WSEAS International Conference on Systems: New Aspects of Systems*, Heraklion, Greece, 2008.
- [20] Đurović Ž.M., I.P. Kovačević, B.D. Kovačević, An adaptive channels parameters estimation using QQ-plot approach, *WSEAS Transaction on Circuits and Systems*, 2008 (accepted for publication).
- [21] Filipović V.Ž., B.D. Kovačević, On robustified minimum variance controller, *International Journal of Control*, Vol. 65, No. 1, 1996, pp. 117-129.
- [22] Filipović V.Ž., B.D. Kovačević, On robust AML identification algorithms, *Automatica*, Vol. 30, No. 11, 1994, pp. 1775-1778.
- [23] Eykhoff P., *System Identification: Parameter and State Estimation*, John Wiley, 1974.
- [24] Gajić Z., *Linear Dynamic Systems and Signals*, Prentice Hall, 2002.
- [25] Gelb A., *Applied Optimal Estimation*, MIT Press, 1974.
- [26] Gnanadesikan R., *Methods for Statistical Analysis of Multivariate Data*, John Wiley, 1977.
- [27] Gonzales J.G., J.L. Paredes, G.R. Arce, Zero-order statistics: a mathematical framework for the processing and characterization of very impulsive signals, *IEEE Transactions on Signal Processing*, Vol. 54, No. 10, 2006, pp. 3839-3851.
- [28] Goodwin G.C., K.S. Sin, *Adaptive Filtering, Prediction and Control*, Prentice Hall, 1984.
- [29] Goodwin G.C., R.L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, 1977.
- [30] Goodwin G.C., Identification: Experiment Design, In Sing M. (Ed.), *Encyclopedia of Systems and Control*, Pergamon Press, 1987.

- [31] Grewal M.S., A.P. Andrews, *Kalman Filtering: Theory and Practice with Matlab*, John Wiley, 2001.
- [32] Hall P., C.C. Heyde, *Martingale Limit Theory and Its Application*, Academic Press, 1980.
- [33] Hampel F., E. Ronchetti, P. Rousseuw, W. Stahel, *Robust Statistics*, John Wiley, 1986.
- [34] Handsdin E., F. C. Schweppe, J. Kohlas, A. Fiechter, Bad data analysis for power system state estimation, *IEEE Transaction on Power Systems and Apparatures*, Vol. 94, No. 3, 1975, pp. 329-337.
- [35] Hannan E.J., M. Deistler, *The Statistical Theory of Linear Systems*, John Wiley, 1988.
- [36] Hannan E.J., Recursive estimation based on ARMA models, *Annal of Statistics*, Vol. 8, 1980, pp. 762-777.
- [37] Hassibi B., B.M. Hochwald, How much training is needed in multiple-antenna wireless links, *IEEE Transactions on Information Theory*, Vol. 49, No. 4, 2003, pp. 951-963.
- [38] Haykin S., *Adaptive Filter Theory*, Prentice Hall, 1996.
- [39] Hogg R., *Robustness in Statistics*, Academic Press, 1979.
- [40] Huber P.J., *Robust Statistics*, John Wiley, 1981.
- [41] Jazwinski A.H., *Stochastic Processes and Filtering Theory*, Academic Press, 1970.
- [42] Kassam S.A., V.H. Poor, Robust Techniques in Signal Processing, *Proceedings IEEE*, Vol. 73, No3, 1985, pp. 433-481.
- [43] Kashiap R.I., A.R. Rao, *Dynamic Stochastic Models from Empirical Data*, Academic Press, 1976.
- [44] Knopp K., *Infinite Sequences and Series*, Dover, 1956.
- [45] Kovačević B.D., V.Ž. Filipović, Robust real time identification of linear systems with correlated noise, *International Journal of Control*, Vol. 48, No. 3, 1988, pp. 993-1010.
- [46] Kovačević B.D., Ž. M. Đurović, S. T. Glavaški, On robust Kalman filtering, *International Journal of Control*, Vol. 56, No. 3, 1992, pp. 547-562.
- [47] Kovačević B.D., V.Ž. Filipović, Robust recursive identification of linear systems using prior information, *Control Theory and Advanced Technology*, Vol. 10, No. 1, 1994, pp. 39-55.
- [48] Kovačević B.D., Ž.M. Djurović, S.S. Stanković, *Signali i Sistemi*, Akademska Misao, 2007.
- [49] Kovačević B.D., Ž.M. Djurović, *Fundamentals of Stochastic Signals, Systems and Estimation Theory with Worked Examples*, Springer Verlag, 2008.
- [50] Kovačević I.P., B.D. Kovačević, Ž.M. Djurović, A convergence theorem for a class of stochastic gradient type algorithms with application to robust system identification, *Proceeding of International Conference on Circuits-Systems-Communications-Computers*, 2006, Athens, Greece.
- [51] Kovačević I.P., B.D. Kovačević, Ž.M. Djurović, On strong consistency of a class of robust stochastic gradient system identification algorithms, *WSEAS Transaction on Circuits and Systems*, Vol. 5, No. 8, 2006, pp. 1244-1254.
- [52] Kovačević I.P., B.D. Kovačević, Ž.M. Djurović, On strong consistency of a class of recursive stochastic Newton-Raphson type algorithms with application to robust linear dynamic system identification, *Facta Universitatis*, Vol. 21, No. 1, 2008, pp. 1-21.
- [53] Kovačević I.P., *Analiza jedne klase algoritama za estimaciju parametara adaptivnih digitalnih filtara*, Magistarski rad, Univerzitet u Prištini-Fakultet Tehničkih nauka u Kosovskoj Mitrovici, 2004.
- [54] Kovačević I.P., A.M. Savić, *Matlab priručnik za laboratorijske vežbe iz matematike*, Visoka škola elektrotehnike i računarstva, 2007.
- [55] Kovačević I.P., A.M. Savić, Primena programskog paketa Matlab u izvođenju nastave matematike, *Zbornik konferencije Informacione Tehnologije*, Žabljak, Crna Gora, 2004.
- [56] Kreyszig E., *Advanced Engineers Mathematics*, John Wiley, 2006.
- [57] Kuncicky D.C., *Matlab Programming*, Prentice Hall, 2004.
- [58] Lazarević I., *Višedimenzionalna Matematička Analiza*, Orion, 2006.
- [59] Lee C.H., On robust linear prediction, *IEEE Transaction on Acoustics, Speech and Signal Processing*, Vol. 36, No. 2, 1988, pp. 642-650.
- [60] Lewis F.R., *Optimal Estimation with an Introduction to Stochastic Control Theory*, John Wiley, 1986.

- [61] Li H., H.V. Poor, Performance analysis of semiblind channel estimation in long-code DS-CDMA systems, *IEEE Transactions on Signal Processing*, Vol. 54, No. 9, 2006, pp. 3383-3399.
- [62] Lipster R., A. Shirayev, *Martingale Theory*, Moscow, 1986.
- [63] Ljung L., T. Soderstrom, *Theory and practice of recursive identification*, MIT Press, 1983.
- [64] Ljung L., *System Identification :Theory for the User*, Prentice Hall, 1987.
- [65] Lutovac M.D., D.V. Tošić, B.L. Evans, *Filter Design for Signal Processing Using Matlab and Mathematica*, Prentice Hall, 2001.
- [66] Martin R.D., D.J. Thomson, Robust resistant spectrum estimation, *Proceedings IEEE*, Vol. 70, No. 9, 1982, pp. 1097-1014.
- [67] Mareels I.M.Y., R.R. Bitmead, M. Gevers, C.R. Johnson, R.L. Kosut, M.A. Poubelle, How exciting can a signal really be, *Systems and Control Letters*, No. 8, 1987, pp. 197-204.
- [68] Mendel J.M., *Discrete Techniques of Parameter Estimation*, Marcel Dekker, 1973.
- [69] Miller S.L., D.G. Childers, *Probability and Random Processes with Applications to Signal Processing and Communications*, Academic Press, 2004.
- [70] Neveu J., *Discrete Parameter Martingales*, North Holland, 1975.
- [71] Papoulis A., *Probability, Random Variables and Stochastic Processes*, McGraw Hill, 2002.
- [72] Poljak B.T., Ya. Z. Tsyppkin, Robust identification, *Automatica*, Vol. 16, No.2, 1980, pp. 53-63.
- [73] Poor H.V., *An Introduction to Signal Detection and Estimation*, Pringer Verlag, 1994.
- [74] Recktenwald G., *Numerical Methods with Matlab: Implementation and Application*, Prentice Hall, 2000.
- [75] Rey W.J., *Introduction to Robust and Quasi-Robust Statistical Methods*, Springer Verlag, 1983.
- [76] Rey W.J., *Robust Statistical Methods*, Springer Verlag, 1977.
- [77] Rusbagi J.S., *Optimizing Methods in Statistics*, Academic Press, 1971.
- [78] Rousseeuw P.J., A.M. Leroy, *Robust Regression and Outlier Detection*, John Wiley, 1987.
- [79] Royden H.L., *Real Analysis*, The Macmillan Company, 1968.
- [80] Rugb W.J., *Linear System Theory*, Prentice Hall, 1993.
- [81] Soderstrom T., *Discrete Time Stochastic Systems*, Springer Verlag, 2002.
- [82] Solo V., The second order properties of a time series, *Annal of Statistics*, Vol. 9, 1981, pp. 307-317.
- [83] Shirayev A.N., *Probability*, Moscow, 1989.
- [84] Stanković S.S., B.D. Kovačević, Analysis of robust stochastic approximation algorithms for process identification, *Automatica*, Vol. 22, No. 4, 1986, pp. 483-488.
- [85] Stefanović V., *Diskretne Mreže i Procesiranje Signala*, Publikacije Univerziteta u Nišu, 2004.
- [86] Stengel R.F., *Stochastic Optimal Control*, John Wiley, 1986.
- [87] Svobodny T., *Mathematical Modeling for Industry and Engineers*, Prentice Hall, 2004.
- [88] Tsyppkin Ya. Z., *Adaptation and learning in automatic systems*, Academic Press, 1971.
- [89] Tsyppkin Ya. Z., *Foundations of informational identification theory*, Nauka, Moscow, 1984.
- [90] Tukey W.J., *Exploratory Data Analysis*, Addison Wesley, 1977.
- [91] Vidyasagar M., *Nonlinear Systems Analysis*, Prentice Hall, 1978.
- [92] White R.E., *Elements of Matrix Modeling and Computing with Matlab*, Chapman & Hall, 2007.
- [93] Willems J.L., *Stability Theory of Dynamical Systems*, Nelson, Walton-on-Tnemes, 1970.
- [94] Williamson G.A., P.M. Clarkson, W.A. Sethares, Performance characteristics of median LMS adaptive filter, *IEEE Transaction on Signal Processing*, Vol. 41, No.2, 1993, pp. 667-680.
- [95] Zeng H.H., L. Tong, Blind channel estimation using the second order statistics algorithm, *IEEE Transaction on Signal Processing*, Vol. 45, No.8, 1997, pp. 1919-1930.