



Master rad iz matematike
Modul: Statistika, aktuarska i finansijska matematika

Statističko zaključivanje u stohastičkim difuzijama na osnovu diskretizovanog uzorka

Predrag Pilipović

15. septembar 2021.

Matematički fakultet, Univerzitet u Beogradu

Mentor:

Dr Bojana Milošević, docent

Članovi komisije:

Dr Bojana Milošević, docent

Dr Marko Obradović, docent

Dr Marija Cuparić, asistent

Apstrakt

Ideja ovog master rada jeste da uvede u priču o statistici u stohastičkim diferencijalnim jednačinama i modelima zasnovanim na njima. Ukratko će biti predstavljen relativno nov pristup diskretizacije zasnovan na metodama deljenja. Ovaj metod je dobro poznat u numeričkoj matematici i ima značajnu primenu na polju običnih diferencijalnih jednačina, ali još uvek njegov potencijal nije iskorišćen u domenu stohastičkih diferencijalnih jednačina. Master rad će se pozvati na određene naučne radove koji pokazuju kako ovakva vrsta diskretizacije kod oscilatornih modela čuva sva bitna geometrijska svojstva. Pomenuto povlači i statistička svojstva, kao što je postojanost i asimptotska raspodela. Rad će, takođe, na određenoj klasi difuzija ilustrovati različite načine ocenjivanja parametara zasnovane na metodi maksimalne verodostojnosti. Takođe, nakon ocenjivanja parametara, moguće je uraditi dijagnostiku modela, odnosno testiranje saglasnosti modela sa podacima. Na konkretnom primeru biće predložen statistički test na osnovu reziduala modela i biće ispitana njegova svojstva. Sve simulacije i neophodne funkcije biće urađene u programskim jezicima R i Wolfram Mathematica.

Zahvalnost

Pre svega, zahvalio bih se profesorki Suzani Ditlevsen sa Univerziteta u Kopenhagenu koja me uvela u priču o stohastičkim difuzijama i pomogla da izgradim intuiciju i razumem kompleksnost ove oblasti.

Zatim, ogromnu zahvalnost dugujem profesorki Bojani Milošević koja je, pored toga što je moj zvanični mentor pri izradi ovog rada, bila uz mene sve ove godine na Matematičkom fakultetu, pa i u Matematičkoj gimnaziji. Ona je bila prva osoba koja mi je približila svet verovatnoće i statistike i koja mi je neprekidno pomagala pri usvajanju novih znanja, a još važnije, razvijanju intuicije. Kao i do sada, pri izradi ovog rada uvek je bila tu da pomogne i posavetuje i bez prigovora je prihvatila da se upustimo u kompleksnost teme koja je još uvek nedovoljno izučavana.

Zahvalio bih se i ostalim članovima komisije na izdvojenom vremenu i strpljenju prilikom čitanja ovog rada, kao i na svim korisnim sugestijama.

Posebnu zahvalnost dugujem Emiliji Petrinjac koja je izdvojila ogromnu količinu vremena kako bi sa punom pažnjom ispravila sve gramatičke, sintaksne i leksičke greške. Bez nje ovaj master rad ne bi bio čitljiv.

Za kraj, najveću zahvalnost dugujem svojoj porodici bez koje ne bih bio ovde, koja je uvek bila tu za mene da me bezuslovno podrži i trpi.

Predrag Pilipović

Sadržaj

Apstrakt	ii
Zahvalnost	iv
1 Uvod	1
1.1 Motivacija - modelovanje akcionog potencijala	2
1.1.1 Membranski potencijal	3
1.1.2 Hodžkin-Hakslijev i FicHju-Nagumov model	4
1.1.3 Stohastičnost u modelu	6
1.2 Opis problema i predloženi pristup	8
1.3 Sadržaj master rada	8
2 Matematički preduslov	10
2.1 Elementi običnih diferencijalnih jednačina	10
2.1.1 Tok vektorskog polja	11
2.1.2 Linearne diferencijalne jednačine i ekvilibrijumi	11
2.2 Elementi numeričke matematike	13
2.2.1 Ojlerova metoda diskretizacije	13
2.2.2 Numerički tok vektorskog polja	14
2.2.3 Red konvergencije	14
2.2.4 Metode deljenja vektorskog polja	15
2.3 Elementi verovatnoće i statistike	18
2.3.1 Konvergencije slučajnih veličina	18
2.3.2 Slučajni procesi	19
2.3.3 Braunovo kretanje	21
2.3.4 Totalna i kvadratna varijacija	21
2.3.5 Stohastičke diferencijalne jednačine	22
2.3.6 Stohastičke difuzije	24
2.3.7 Itoova formula	26
2.3.8 Ornštajn-Ulenbekov proces	27
2.3.9 Višedimenzionalne stohastičke difuzije	27
2.3.10 Ocene parametara	30
2.3.11 Testovi saglasnosti	30
2.3.12 Mera i moć test	33

2.3.13	Statističke metode simulacije	33
3	Diskretizacija stohastičkih diferencijalnih jednačina	34
3.1	Red srednje-kvadratne konvergencije	34
3.2	Metode diskretizacije	36
3.2.1	Metode Ojler-Marujame	36
3.2.2	Metode višeg reda	37
3.3	Struktura šuma: eliptičnost i hipoeliptičnost	39
3.4	Model i osobine	40
3.4.1	Metode deljenja za stohastičku diferencijalnu jednačinu	41
3.4.2	Srednje-kvadratna konvergencija za metode deljenja	42
3.4.3	Očuvanje hipoeliptičnosti	45
3.4.4	Očuvanje strukture modela	46
4	Ocenjivanje parametara	49
4.1	Funkcija verodostojnosti	49
4.2	Problemi hipoeliptičnog modela	50
4.3	Iterativni pristup: Ditlevsen-Samson	51
4.4	Ocenjivanje bez iteracija: Melnjikova	54
4.5	Metoda najmanjih kvadrata: Melnjikova	55
4.6	Nov predlog ocenjivanja parametara	56
4.7	Numerički eksperimenti	57
5	Testiranje saglasnosti podataka sa modelom	63
5.1	Negri-Nišijaminov test saglasnosti	63
5.2	Lijev test saglasnosti	65
5.3	Numerički eksperimenti	69
6	Zaključak	74
	Literatura	76
	Biografija	78

Uvod

U prethodnih nekoliko decenija razvoj informacionih tehnologija nadmašio je sva ljudska očekivanja. Danas je gotovo nezamislivo da ljudi nisu čuli za mašinsko učenje i veštačku inteligenciju. Sve više se ulaže u razvijanje modela dubokog učenja koji rade kao crne kutije i rešavaju probleme za koje do pre par godina nismo mogli ni da pretpostavimo da će biti rešavani od strane mašina. Pored svih pozitivnih strana ove istraživačke sfere postoji i očigledna mana, a to je da se statistička i programerska naučna zajednica sve više fokusira na modele mašinskog učenja, a sve manje na čiste statističke modele za koje još uvek ne postoji metodologija koja bi opisala protokol modelovanja. Jedna takva tema jesu upravo modeli zasnovani na stohastičkim diferencijalnim jednačinama koji opisuju neku pojavu ili proces u neprekidnom vremenu.

Kada je reč o neprekidnim funkcijama u matematici, postoji vrlo dobar matematički aparat koji se bavi ovim problemom. Svima su dobro poznati integrali, izvodi, funkcionali i diferencijalne jednačine. Sa druge strane, kada pričamo o neprekidnom vremenu, nailazimo na maglovitu rupu u statističkom delu matematike. Razlog za ovo je vrlo očigledan: svi podaci koje imamo i možemo dobiti i skladišiti su diskretizovani, čak i kada je proces koji posmatramo neprekidan po vremenu. Zbog toga modeli kao što su linearna regresija, ili generalizovana linearna regresija, imaju razvijen ceo metodički aparat - znamo kako da postavimo problem, šta je cilj i kako da donosimo smislene statističke zaključke o ovakvim modelima. Slično važi i za neke druge statističke grane kao što su vremenske serije ili analiza preživljavanja.

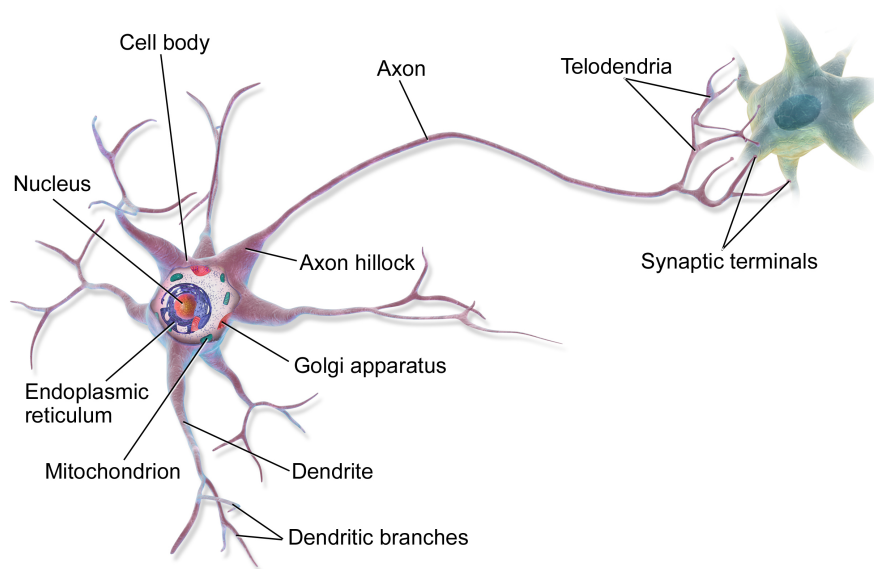
Kada pričamo o stohastičkim diferencijalnim jednačinama, vrlo često problem posmatramo sa čisto probablističkog pristupa, tj. bavimo se teorijom iza stohastičkih diferencijalnih jednačina, kao što su dobra definisanost same jednačine (problem sa diferencijabilnošću Braunovog kretanja), jedinstvenost i postojanje rešenja i, eventualno, kako analitički rešiti tu jednačinu. Takođe, pod uticajem numeričke matematike za obične diferencijalne jednačine, u velikoj meri istraženo je i diskretizovanje date stohastičke diferencijalne jednačine i vrlo često se mogu naći simulacije i njihova analiza. Međutim, kada je reč o modelovanju stvarnih procesa stohastičkom diferencijalnom jednačinom i kada je reč o statističkom zaključivanju, nailazimo na jako mali broj radova koji se uglavnom bave specifičnim modelom i neretko se problemu pristupa tzv. brut forsom (na eng. *brute force*), što znači da se ne traži pametan i optimalan način, već isključivo postoji želja da se dati podaci uklope u dati model i donesu zaključci koji su u uskoj vezi

sa konkretnim problemom. Zbog svega pomenutog cilj ovog master rada je da pokrene ideju o formiranju jedinstvene metodologije za rešavanje pomenutog statističkog pitanja. Biće predstavljene sve neophodne definicije i teoreme iz teorije verovatnoća, statistike, diferencijalnih jednačina, numeričke matematike, kao i neka nova zapažanja o tome kako doći do jedinstvenog protokola. Pre svega toga, u narednom odeljku biće uveden problem iz realnog života koji će nam kroz ceo rad služiti kao primer za prikazanu teoriju i simulacije.

1.1 Motivacija - modelovanje akcionog potencijala

Ideja ovog odeljka je da ukratko ukaže na to da postoje razne oblasti van granica matematičke apstrakcije koje primenjuju sve što će biti izloženo u ovom master radu. Izabrani primer se tiče modelovanja akcionog potencijala jednog neurona. Ovaj primer jeste najprostiji moguć iz oblasti neurologije, ali je takođe dosta zanimljiv i prilično se razlikuje od generičkih primera koji se koriste kada se izučavaju regresija ili vremenske serije. Kako bi rad bio potpun, ukratko će biti opisana biološka strana problema.

Signali kroz naše telo se prenose kroz nerve, a svaki nerv sačinjen je od nervnih ćelija - neurona. Na slici 1.1 je prikazan šematski prikaz neurona. Signal se preko sinapsi prenosi kao električni impuls od jednog neurona do drugog. Kada se naglo promeni voltaža u membranskom potencijalu, stvara se akcioni potencijal koji deluje kao okidač za prenos signala sa jednog neurona na drugi. Na nastanak akcionog potencijala utiče promena voltaže u membrani, kao i impuls dobijen od nekog prethodnog neurona. Da bismo bolje razumeli ove procese, potrebno je nešto bliže da sagledamo okolinu membrane neurona.

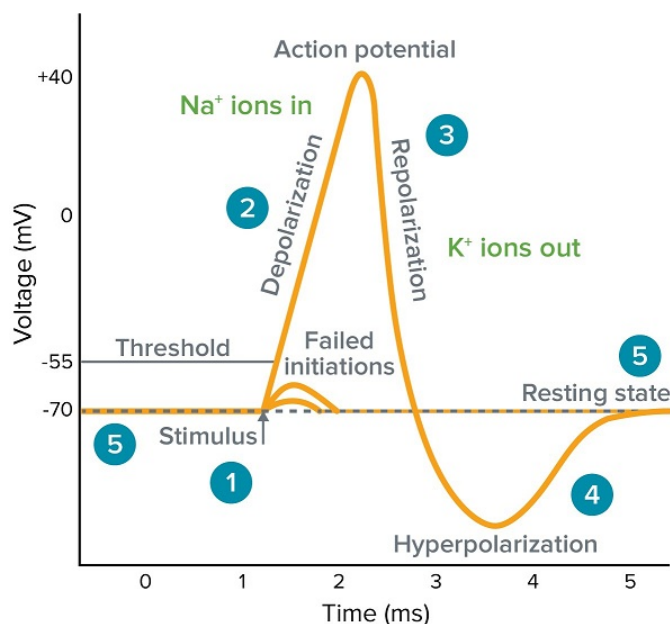


Slika 1.1: Šematski prikaz neurona. ¹

¹Slika preuzeta sa: https://en.wikipedia.org/wiki/File:Blausen_0657_MultipolarNeuron.png

1.1.1 Membranski potencijal

Membranski potencijal je razlika između električnog potencijala unutrašnjosti i spoljašnjosti ćelije. Ovaj potencijal se uobičajeno izražava u milivoltima. Transmembranski proteini, poznati kao jonske pumpe, aktivno guraju jone kroz membranu dok, s druge strane, jonski kanali dozvoljavaju jonima da se vraćaju odakle su prethodno došli. Jonske pumpe i jonski kanali u membrani ekvivalent su bateriji i otpornicima, te tako stvaraju napon između dve strane membrane. Važno je još napomenuti da su ključni joni za stvaranje akcionog potencijala u stvari pozitivno naelektrisani joni kalcijuma i natrijuma, Ka^+ i Na^+ . Ovo ne znači da su oni jedini joni, već da su najznačajniji.



Slika 1.2: Akcioni potencijal neurona. 1. Stimulus (ili nadražaj) pokreće brzu promenu napona i u trenutku prelaska praga, akcioni potencijal je ispaljen. 2. Depolarizacija je uzrokovana brzim povećanjem membranskog potencijala zbog otvaranja kanala za jone natrijuma. 3. Membranska repolarizacija nastaje brzim zatvaranjem kanala natrijuma, kao i dolaskom velike količine jona kalijuma kroz otvorene kanale za jone kalijuma. 4. Hiperpolarizacija je stanje posle repolarizacije kada je membranski potencijal smanjen ispod nivoa mirovanja, a uzrok tome je suvišan nivo jona kalijuma i zatvaranje kanala za jone natrijuma. 5. Stanje mirovanja se potom polako dostiže i membranski potencijal se vraća na nivo na kom je bio pre stimulusa.²

Neuroni u svom stanju mirovanja imaju membranski potencijal u relativno stabilnom stanju u granicama od $-70mV$ do $-80mV$. Ovaj potencijal se naziva i potencijal mirovanja. Otvaranje i zatvaranje jonskih kanala u slučaju da napon postane manje negativan može izazvati pomeranje iz ravnoteže koja se naziva depolarizacija. Dovoljno velika depolarizacija potom može dovesti do akcionog potencijala u kom se membranski potencijal menja vrlo brzo i jako značajno u jako kratkom vremenskom periodu (reda veličine od 1 do 100 milisekundi). Akcioni potencijal je ono što želimo da izučavamo.

²Slika preuzeta sa: <https://www.moleculardevices.com/applications/patch-clamp-electrophysiology/what-action-potential>

On je ključan za prenos impulsa sa jednog neurona na drugi i ključan za razumevanje fiziologije mozga i nervnog sistema. Na slici 1.2 prikazan je akcioni potencijal neurona.

Pre nego što nastavimo, važno je uvesti pojam obrnutog potencijala, iliti Nernstovog potencijala. Nernstov potencijal, poznat još i kao potencijal ekvilibrijuma, konstantna je vrednost za svaki tip jona i predstavlja napon na kom se potencijali na membrani potiru. Drugim rečima, na toj voltaži više ne postoji protok jona kroz membranu. Dakle, struja je nula i ne menja se kada je potencijal konkretnog jona dostigao Nernstovu vrednost. Obično se označava kao E_{jon} .

Sada, nakon što smo u određenoj meri upoznati sa pričom o akcionom potencijalu, kroz naredni pododeljak diskutovaćemo dva najpoznatija modela membranskog potencijala.

1.1.2 Hodžkin-Hakslijev i FicHju-Nagumov model

Hodžkin-Hakslijev model predložen je 1952. godine kao objašnjenje inicijalizacije i propagacije akcionog potencijala unutar gigantskog neurona lignje. Za ovaj model Alan Hodžkin i Endru Haksli su 1963. dobili Nobelovu nagradu koja se dodeljuje za psihologiju ili medicinu.

Ovde nećemo izvoditi model, već ćemo ga predstaviti gotovog i pokušati ukratko da objasnimo njegove komponente. Model čini četvorodimenzionalni nelinearni sistem diferencijalnih jednačina i glasi

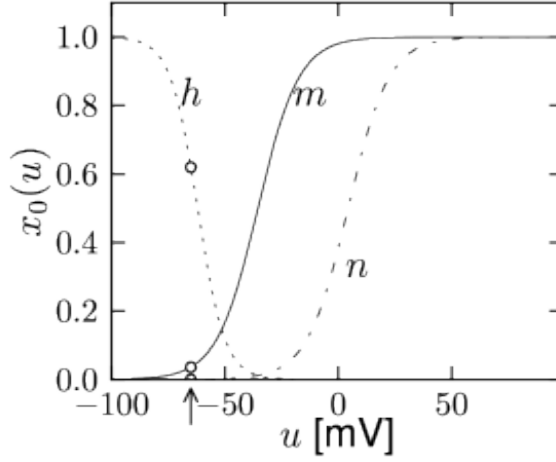
$$C \frac{dU}{dt} = -g_{Na}m^3h(U - E_{Na}) - g_Kn^4(U - E_K) - g_l(U - E_l) + I(t), \quad (1.1)$$

$$\frac{dm}{dt} = -\frac{m - m_0(U)}{\tau_m(U)}, \quad (1.2)$$

$$\frac{dn}{dt} = -\frac{n - n_0(U)}{\tau_n(U)}, \quad (1.3)$$

$$\frac{dh}{dt} = -\frac{h - h_0(U)}{\tau_h(U)}. \quad (1.4)$$

Glavna komponenta ovog modela je U što je napon koji želimo da modelujemo i kroz koji pratimo akcioni potencijal. Ostale diferencijalne jednačine opisuju kada su kanali otvoreni za protok odgovarajućih jona. Funkcije m i h opisuju aktivaciju, odnosno deaktivaciju protoka natrijuma, dok funkcija n opisuje aktivaciju protoka kalijuma. Funkcija $I(t)$ predstavlja spoljni stimulus ubačen u sistem i najčešće se modeluje kratkim pulsom ili konstantom. Ovo je veličina koju eksperimentalno zadajemo. Vrednosti E_{Na} , E_K i E_l predstavljaju Nernstove potencijale za protok jona natrijuma, kalijuma i ostalih jona (najčešće hlora), respektivno. Skraćenica l u oznaci, od engleske je reči *leak* i označava "curenje" ostalih jona kroz membranu. Parametri g_{Na} , g_K i g_l predstavljaju maksimalnu provodljivost kada su svi kanali otvoreni. Poslednje tri jednačine opisuju otvaranje i zatvaranje kanala i uzimaju vrednost između 0 i 1. Njih interpretiramo tako što se za fiksiranu vrednost napona U , promenljiva x približava ciljnoj vrednosti $x_0(U)$ brzinom $\tau_x(U)$, gde x predstavlja jednu od promenljivih m , n i h . Na slici 1.3 su prikazane ciljane funkcije n_0 , m_0 i h_0 u zavisnosti od napona U dobijene empirijski. Na isti način određena su i vremena $\tau_x(U)$.



Slika 1.3: Vrednosti ekvilibrijuma za promenljive n , m i h u zavisnosti od napona U . Strelica pokazuje na stanje mirovanja $U = -70\text{mV}$.³

Iako je nastao davno, ovaj model i danas je referenca za modelovanje neurona i akcionih potencijala. S obzirom na to da je model zamišljen da modeluje konkretan neuron, vrlo lako je prilagodljiv i može se proširiti na modelovanje celog snopa neurona, odnosno nerava, što je mnogo praktičnije u primeni. S druge strane, možemo primetiti da je model sam po sebi dosta kompleksan. Ima četiri dimenzije što automatski isključuje izučavanja faznih portreta, a o kompleksnosti ocenjivanja parametara ne moramo ni govoriti. Iz tog razloga ćemo, motivisani prethodnim modelom, definisati pojednostavljenu verziju modela u 2 dimenzije. Ovaj model, poznatiji kao FicHju-Nagumov model, biće glavni primer kroz ceo master rad na kome ćemo primenjivati svu izloženu teoriju.

FicHju-Nagumov model dobio je ime po R. FicHjuu koji je predložio model 1961. godine i J. Nagumou koji je ekvivalentan model predložio naredne godine. FicHju-Nagumov model se dobija uprošćavanjem Hodžkin-Hakslijevog modela, tako što se iskoristi činjenica da je brzina promene promenljive m jako velika u odnosu na ostale promenljive, pa se m može aproksimirati sa $m(U) = m_0(U)$. Druga aproksimacija se zasniva na činjenici da su brzine promene za n i h veoma slične, pa možemo jednom promenljivom V zameniti pomenute dve uvođenjem smene $V = b - h = an$. Time smo dobili dvodimenzionalni model oblika

$$\frac{dU}{dt} = \frac{1}{\tau_U}(F(U, V) + I), \quad (1.5)$$

$$\frac{dV}{dt} = \frac{1}{\tau_V}G(U, V), \quad (1.6)$$

gde su τ_U i τ_V parametri brzine, a F i G nelinearna vektorska polja kretanja dinamičkog sistema. Dobili smo uopšteni model, a konkretno FicHju-Nagumov model ima sledeći oblik

$$\frac{dU}{dt} = U - \frac{U^3}{3} - V + I, \quad (1.7)$$

$$\frac{dV}{dt} = \frac{1}{\tau}(U - a + bV), \quad (1.8)$$

³Slika preuzeta sa: <https://neurondynamics.epfl.ch/online/Ch2.S2.html>

gde su a, b i τ parametri modela, a I eksterni stimulus koji može biti poznat, ali i ne mora. Kao i do sad U predstavlja membranski napon, a nova promenljiva V se naziva još i promenljivom oporavka. Ona aproksimira kontrolu svih kanala koji dovode jone unutar neurona, a time i do akcionog potencijala, ali i kontroliše oporavljanje nakon ispaljivanja stimulusa. U ovom modelu je dovoljan samo jedan vremenski parametar τ koji kontroliše brzinu oporavka. Već smo videli da akcioni potencijal kada se desi, desi se jako brzo, dok oporavak traje jako dugo. Upravo to modelujemo parametrom τ koji nam govori da je brzina promene V drastično drugačija od brzine U .

U daljem radu, kako bismo se sinhronizovali sa određenim radovima radi poređenja rezultata, koristićemo varijaciju FicHju-Nagumovog modela. Dalje ćemo ukratko izdiskutovati odakle dolazi slučajnost, a potom ćemo završiti poglavlje opisom statističkog problema i kako ćemo mu pristupiti.

1.1.3 Stohastičnost u modelu

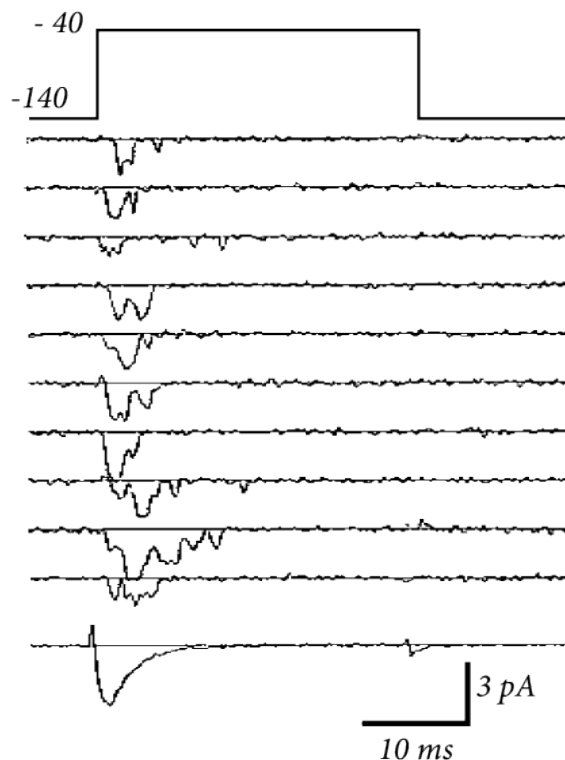
Zasad smo opisali proces koji se dešava svakog trenutka u mozgovima svih nas i uspeći smo da ga opišemo matematičkom jednačinom. Problem ovog modela, kao i većine determinističkih modela danas, jeste što ne uračunava slučajnost koja se dešava iz raznih razloga. Neki od najčešćih razloga zašto moramo uključiti slučajnost u deterministički model su:

1. Broj jonskih kanala u određenom delu membrane je konačan, ali razlikuje se od neurona do neurona, od dela membrane do njenog drugog dela. Takođe, otvaranje i zatvaranje tih kanala je slučajan proces.
2. Do jednog neurona stiže na hiljade drugih koji šalju svoje signale i impulse. Kada problem posmatramo iz ugla jednog neurona, ti impulsi dolaze nasumično.
3. Greška merenja je takođe jedan od razloga što moramo uključiti šum u deterministički model, pogotovo kada merimo nešto tako malo i delikatno kao što je neuron.

Na slici ?? je prikazano više ponavljanja istog eksperimenta koji indukuje određeni impuls u neuron. Možemo odmah primetiti da svaki eksperiment ima drugačiju trajektoriju, što implicira da je posmatrani proces slučajan.

Postoji više načina rešavanja ovog problema. Verovatno najčešći način koji su naučnici širom sveta koristili skoro u svakoj situaciji, jeste posmatrati prosek realizacija i pretvarati se da je taj prosek deterministički određen. Drugi način se bavi modelovanjem ispaljivanja akcionog potencijala Puasonovim procesom. U ovom scenariju putanja napona membrane je i dalje opisana determinističkom formulom, ali se modeluje verovatnoća da napon pređe prag i aktivira akcioni potencijal u svakom trenutku. Više o ovom načinu modelovanja može se naći u knjizi [5]. Ono što nas zanima, jeste da modelujemo dinamički sistem stohastičkom diferencijalnom jednačinom i bavimo se statistikom unutar okvira stohastičkih diferencijalnih jednačina. U ovom radu ćemo posmatrati okvir Itoovih stohastičkih diferencijalnih jednačina gde ćemo grešku modelovati Braunovim kretanjem, iliti Vinerovim procesom. Postoje i drugačije definicije stohastičkog integrala, ali se ovde

nećemo baviti njima. Takođe, postoje i druge raspodele za greške, a Braunovo kretanje se može uopštitit Levijevim procesima, što takođe izlazi iz okvira ovog master rada.



Slika 1.4: Stohastičko ponašanje napona membrane u istom predelu istog neurona snimljeno pri ponavljanju istog eksperimenta. Struja koja se šalje u vidu impulsa (gornji red) se primenjuje na mali deo jednog neurona, i taj eksperiment se primenjuje više puta (svaki red je realizacija jednog eksperimenta). Uprosečene realizacije su prikazane u poslednjem redu.⁴

Dakle, naš model će imati sledeći oblik

$$dU_t = \frac{1}{\tau_U} (F(U_t, V_t) + I) dt + \sigma_1 dW_t^1, \quad (1.9)$$

$$dV_t = \frac{1}{\tau_V} G(U_t, V_t) dt + \sigma_2 dW_t^2, \quad (1.10)$$

gde su σ_1 i σ_2 novi parametri modela koji određuju raspršenost, a W_t^1, W_t^2 dva nezavisna standardna Vinerova procesa. U kasnijem radu će ovi modeli biti formalno uvedeni.

Nakon što smo ukratko sagledali primer iz realnog sveta, njegovo fizičko manifestovanje i kako ga matematički modelujemo, spremni smo da uopštimo problem na veliku klasu primera i da vidimo kako možemo da ga rešimo iz statističkog ugla.

⁴Slika je preuzeta sa: <https://neurondynamics.epfl.ch/online/Ch2.S2.html>

1.2 Opis problema i predloženi pristup

Kao što smo videli u prethodnom primeru, problem iz stvarnog života se prilično jednostavno opisuje diferencijalnim jednačinama i dinamičkim sistemima. Ovakvih primera ima u svim oblastima od neurologije, preko ekonomije i finansija, do ekologije i klimatskih promena. Glavni problem kod ovakvih modela je što isključuju svaki vid slučajnosti. Isto kao i u prethodnom primeru, postoje različiti izvori za slučajnost, na primer slučajne greške ili nedovoljan broj informacija kao što su koliko ima jonskih kanala i kako se otvaraju. Svi ovi nedostaci modela i uprošćena verzija stvarnog procesa mogu da se poboljšaju uključivanjem slučajne komponente u jednačinu.

Glavni cilj svakog modelovanja se svodi na neki vid predviđanja. Na osnovu datih podataka, često želimo da vidimo kako će se stvari odvijati za one podatke koje još uvek ne posedujemo (npr. kada želimo da predvidimo budućnost ili ponašanje nekog neurona kog ne možemo da snimimo). Ako ne uračunamo slučajnost u model, ocene i predviđanja su često gruba i zavise od eksperimenta, te nemamo nikakvu predstavu o verodostojnosti, kao ni to koliko odstupamo od stvarnih vrednosti parametara modela i ne možemo da testiramo hipoteze. Zbog svega ovoga mnogo je bolje da model na samom početku postavimo kao stohastički. Iako u tom slučaju možda deluje da se kompleksnost udvostručuje, dobijeni rezultati su znatno informativniji.

Bilo da je reč o običnim ili stohastičkim diferencijalnim jednačinama, jasno je da problem pravi neprekidnost vremena. Niti možemo da merimo, niti da skladištimo neprekidno vreme, zbog čega je neophodno pomiriti se sa greškom diskretizacije. Kad smo već primorani da diskretizujemo, želimo da to uradimo što smislenije kako bi dodatna greška koju činimo bila što manja i da bi se sve osobine našeg neprekidnog modela prenele na diskretizovani model. U radu će biti pomenute neke vrste diskretizacije i njihove osobine. Videćemo da najpoznatija i najjednostavnija diskretizacija nije uvek najbolji izbor iako se i dalje koristi u najvećem broju slučajeva.

Krajni cilj ovog rada bi bio da se formiraju metodologija i protokol koji će koristiti napredne metode diskretizacije i koji će biti efikasni u statističkim zadacima kao što su ocene parametara, testiranje hipoteza i predviđanje tamo gde podaci nisu dostupni. Naravno, skup svega navedenog prevazilazi okvire jednog master rada, ali ćemo se potruditi da uvedemo čitaoca u priču, da objasnimo pojmove i pokažemo šta sve danas postoji u literaturi, kao i da predložimo neke od načina za približavanje cilju.

1.3 Sadržaj master rada

Master rad je izdelfjen na šest poglavlja. U prvom poglavlju videli smo šta je problem i kako planiramo da mu pristupimo. Prikazali smo primenu na realnom primeru. Ukratko smo sagledali fiziku iza aktivacije neurona i kako se može matematički modelovati. Takođe smo videli ograničenja determinističkog modela i kako je moguć prelazak na stohastički. U sledećem poglavlju ćemo se podsetiti svih važnih matematičkih pojmova i osobina u vezi sa slučajnim procesima, numeričkom matematikom, diferencijalnim jednačinama

i, konačno, stohastičkim diferencijalnim jednačinama. U trećem poglavlju ćemo uvesti neke od algoritama diskretizacije modela, videti neke njihove osobine, kao i ukazati na to zašto su neki bolji od drugih. U četvrtom poglavlju biće više reči o oceni parametara. Na primeru FicHju-Nagumovog modela poredićemo ocene dobijene različitim metodama. Na kraju, u petom poglavlju, zapoćemo diskusiju o testiranju hipoteza i pozabaviti se pitanjem testova saglasnosti sa modelom. Poslednje poglavlje je zaključak gde ćemo ukratko iskomentarisati dobijene rezultate.

Matematički preduslov

U ovom poglavlju ćemo se podsetiti nekih pojmova i uvesti poneki nov iz oblasti numeričke matematike, diferencijalnih jednačina, slučajnih procesa i stohastičkih diferencijalnih jednačina. Svi ovi pojmovi biće potrebni za bolje razumevanje materije, s obzirom na to da su znanja iz različitih oblasti potrebna kako bi se tema zaokružila.

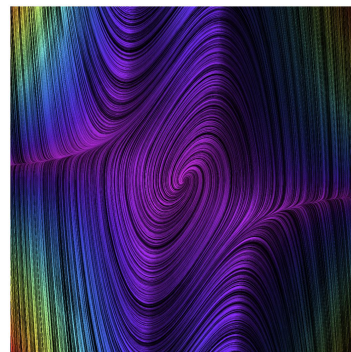
Najpre ćemo se ukratko podsetiti običnih diferencijalnih jednačina, a potom ćemo uvesti neke osnovne pojmove iz numeričke matematike u vezi sa njima (diferencijalnim jednačinama). Potom ćemo se podsetiti slučajnih procesa i Braunovog kretanja. Pomenućemo važnu osobinu Braunovog kretanja koja kaže da realizacija Braunovog kretanja nije nigde diferencijabilna sa verovatnoćom jedan. Ovo će biti važan koncept za uvođenje Itoovog integrala, kao i stohastičke diferencijalne jednačine zasnovane na Itoovom integralu. Zatim ćemo uvesti pojam stohastičke difuzije i pomenuti regulatorne uslove pod kojima postoji jedinstvena difuzija kao rešenje stohastičke diferencijalne jednačine. Na kraju poglavlja podsetićemo se osnovnih pojmova iz statistike.

2.1 Elementi običnih diferencijalnih jednačina

Podsetimo se definicije autonomne obične diferencijalne jednačine

$$\mathbf{y}'(t) = \mathbf{f}(\mathbf{y}), \quad \mathbf{y}, \mathbf{f} \in \mathbb{R}^d. \quad (2.1)$$

Odgovarajući prostor $\mathbb{R}^d$ nazivamo **faznim prostorom**. Rešenje prethodnog sistema u faznom prostoru nazivamo **trajektorijom**. Dodatno, **fazni portret** je kolektivni grafik trajektorija u faznom prostoru. Očigledno ne možemo nacrtati sve trajektorije, pa faznim portretom nazivamo uprošćeni grafik koji prikazuje nekoliko trajektorija. Na slici 2.1 je prikazan primer jednog faznog portreta.



Slika 2.1: Fazni portret

Dve trajektorije u faznom prostoru se ne smeju preseći jer bi to značilo da rešenje nije jedinstveno u tački preseka. Takođe, rešenje diferencijalne jednačine „putuje” krivom

trajektorije samo u jednom smeru, tj. ne može se okrenuti jer bi to ponovo bilo u suprotnosti sa jedinstvenošću rešenja.

Funkcija $\mathbf{f} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ definiše **vektorsko polje** na vektorskom prostoru, tj. svaka tačka $\mathbf{y} \in \mathbb{R}^d$ odgovara vektoru $\mathbf{f}(\mathbf{y}) \in \mathbb{R}^d$. Vektorsko polje $\mathbf{f}$ je svuda tangentno na sva rešenja diferencijalne jednačine 2.1. Trajektorija se još naziva i **orbita** ili **integralna kriva** u faznom prostoru.

2.1.1 Tok vektorskog polja

Ako na diferencijalnu jednačinu (2.1) dodamo i inicijalnu vrednost $\mathbf{y}(0) = \mathbf{y}_0$, dobijamo **Košijev problem**. Za Košijev problem definišemo **tok vektorskog polja**, skraćeno samo tok, kao preslikavanje

$$\Phi_\tau(\mathbf{y}_0) = \mathbf{y}(\tau), \quad (2.2)$$

gde je $\mathbf{y}(t)$ rešenje Košijevog problema. Preslikavanje $\Phi_\tau : \mathbb{R}^d \rightarrow \mathbb{R}^d$ preslikava inicijalnu vrednost u poslednju vrednost posle trenutka τ .

Posmatrajmo diferencijalnu jednačinu (2.1) sa inicijalnim uslovom $\mathbf{y}(0) = \mathbf{y}_0$ i označimo sa $\mathbf{y}_1$ rešenje u trenutku t_1 , tj. $\mathbf{y}(t_1) = \mathbf{y}_1$. Sada posmatrajmo diferencijalnu jednačinu (2.1) sa inicijalnim uslovom $\mathbf{y}(0) = \mathbf{y}_1$ i označimo sa $\mathbf{y}_2$ rešenje u trenutku t_2 , odnosno $\mathbf{y}(t_2) = \mathbf{y}_2$. Zbog invarijantnosti autonomnih sistema na translaciju vremena, jasno je da je $\mathbf{y}_2$ rešenje koje bismo dobili da smo rešavali diferencijalnu jednačinu (2.1) sa uslovom $\mathbf{y}(0) = \mathbf{y}_0$, posle vremena $t_1 + t_2$. Drugim rečima, važi

$$\Phi_{t_2} \circ \Phi_{t_1} = \Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2}, \quad t_1, t_2 > 0. \quad (2.3)$$

Dakle, kompozicija dva toka je takođe tok istog vektorskog polja. Kada bismo mogli da rešimo diferencijalnu jednačinu za sva pozitivna i negativna vremena, onda bismo pisali

$$\Phi_t \circ \Phi_{-t} = \Phi_0, \quad (2.4)$$

ali Φ_0 je rešenje diferencijalne jednačine na intervalu nulte dužine, pa je to zapravo samo identitet. Dakle, važi

$$\Phi_{-t} = \Phi_t^{-1}. \quad (2.5)$$

Dakle, familija svih tokova $\{\Phi_t \mid t \in \mathbb{R}\}$ čini grupu u odnosu na operaciju kompozicije.

2.1.2 Linearne diferencijalne jednačine i ekvilibrijumi

Linearna diferencijalna jednačina

$$x' = \lambda x + c, \quad x(0) = x_0, \quad (2.6)$$

ima jedinstveno rešenje

$$x(t) = e^{\lambda t} x_0 + \frac{e^{\lambda t} - 1}{\lambda} c, \quad (2.7)$$

gde su $\lambda, c \in \mathbb{R}$, a $x : \mathbb{R} \rightarrow \mathbb{R}$. **Ekvilibrijum** ili **fiksna tačka** sistema se dobija kada rešimo jednačinu $x' = 0$, tj. $x = -\frac{c}{\lambda}$. Za ekvilibrijum kažemo da je **stabilan** ako je $\lambda < 0$, a **nestabilan** ako je $\lambda > 0$.

Višedimenzionalna linearna diferencijalna jednačina

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{b}, \quad \mathbf{y}(0) = \mathbf{y}_0, \quad (2.8)$$

gde je $\mathbf{A} \in \mathbb{R}^{d \times d}$, $\mathbf{b} \in \mathbb{R}^d$ i $\mathbf{y} : \mathbb{R}^d \rightarrow \mathbb{R}^d$, ima jedinstveno rešenje

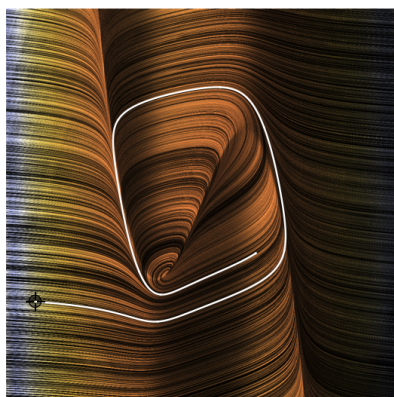
$$\mathbf{y}(t) = e^{\mathbf{A}t}\mathbf{y}_0 + \mathbf{A}^{-1}(e^{\mathbf{A}t} - \mathbf{I})\mathbf{b}, \quad (2.9)$$

kad je $\mathbf{A}$ invertibilna matrica. U slučaju da nije, inverz je zapisan kao konvergentan red. Važno je prisetiti se eksponenta matrice koji je definisan za kvadratne matrice $\mathbf{A}$ kao

$$e^{\mathbf{A}} = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{A}^k, \quad (2.10)$$

gde je $\mathbf{A}^0 = \mathbf{I}$ jedinična matrica. Prethodni red konvergira za sve kvadratne matrice $\mathbf{A}$. Dalje neće biti reči o eksponentu matrice, iako će se kasnije u radu koristiti neke od njegovih osobina. Za više informacija o eksponentu matrice konsultovati bilo koji udžbenik iz dinamičkih sistema.

U višedimenzionalnim linearnim jednačinama ekvilibrijum se traži na isti način i dostiže u tački $\mathbf{y} = \mathbf{A}^{-1}\mathbf{b}$, kada je matrica $\mathbf{A}$ invertibilna. Stabilnost ekvilibrijuma određuju sopstvene vrednosti matrice $\mathbf{A}$ i to: ako su sve sopstvene vrednosti (ili realni delovi sopstvenih vrednosti) negativne, ekvilibrijum je stabilan; a ako su sve pozitivne, ekvilibrijum je nestabilan. Ako ekvilibrijum nije ni stabilan ni nestabilan, onda je sedlo. U zavisnosti od toga da li su sopstvene vrednosti realne ili kompleksne, vrši se dalja karakterizacija ekvilibrijuma, što za nas neće biti važno.



Slika 2.2: Fazni portret za Fichju-Nagumov model sa parametrima $\varepsilon = 0.1$, $\gamma = 1.5$, $s = 1$ i $\alpha = 1.3$. Bela linija predstavlja jednu trajektoriju.

U slučaju nelinearne diferencijalne jednačine određuje se ekvilibrijum i onda se izvršava linearizacija Tejlrovim razvojem kako bi se odredila stabilnost. Ovaj metod radi skoro uvek, osim u pojedinim slučajevima (kada je ekvilibrijum centar), mada navedeno neće biti od važnosti za dalji rad.

Fichju-Nagumov model koji ćemo koristiti u radu zadat je sledećim sistemom nelinearnih diferencijalnih jednačina

$$\begin{cases} U'(t) &= \frac{1}{\varepsilon}(U(t) - U(t)^3 - V(t) + s), \\ V'(t) &= \gamma U(t) - V(t) + \alpha. \end{cases} \quad (2.11)$$

Neznatno smo izmenili jednačine u odnosu na originalan model radi poređenja sa nekim drugim radovima koji koriste ovu verziju modela. Parametar ε kontroliše razliku u brzini, dok je parametar s impuls koji kontroliše

istraživač. Često se pretpostavlja da je $s = 0$. U zavisnosti od izbora parametara $\varepsilon, \gamma, \alpha$ i s , ovaj model ima jedan, dva ili tri ekvilibrijuma i njihova stabilnost zavisi isključivo od vrednosti parametara. Na slici 2.2 prikazan je fazni portret za jedan FicHju-Nagumov model. Primetimo da ovaj model ima nestabilan ekvilibrijum koji je bela trajektorija zaobišla. Umesto toga, stvara se **granični krug**, tj. stabilna zatvorena linija ka kojoj trajektorija konvergira.

2.2 Elementi numeričke matematike

Bilo da je reč o običnim diferencijalnim jednačinama ili o stohastičkoj verziji, potreban je jak numerički aparat. Pretpostavimo da želimo da rešimo sledeću običnu d -dimenzionalnu diferencijalnu jednačinu

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}), \quad \mathbf{y}(t_0) = \mathbf{y}_0, \quad t \in D = [t_0, t_0 + T]. \quad (2.12)$$

Možemo reći da je osnovni zadatak numeričke aproksimacije da može da aproksimira model u konačnom broju operacija. Proces zamene neprekidnog rešenja konačnim, naziva se **diskretizacijom**. To radimo tako što zamenimo interval D tačkama $t_n = t_0 + nh$, $n = 0, 1, \dots, N$, kojih ima N . Vrednost T predstavlja **dužinu intervala**, a $h = T/N$ je **dužina koraka**. Slično, menjamo neprekidno rešenje $\mathbf{y}(t)$ na intervalu D numeričkim aproksimacijama

$$\tilde{\mathbf{y}}_n \approx \mathbf{y}(t_n), \quad n = 0, 1, 2, \dots, N. \quad (2.13)$$

Algoritmi kojima ćemo se baviti aproksimiraju $\tilde{\mathbf{y}}_{n+1}$ rekurzivno pomoću $\tilde{\mathbf{y}}_n$, h i eksplisitne forme vektorskog polja $\mathbf{f}$. Ovakve metode su zasnovane na aproksimaciji jednog koraka unapred, zbog čega se i nazivaju metode **jednog koraka**. Postoje metode koje idu i unazad ili koriste više od jednog koraka. Jedna od najpoznatijih klasa metoda su **metode Runge-Kutovog tipa**, ali o njima neće biti reči u ovom radu. Pomenućemo samo najjednostavniju metodu iz ove klase o kojoj će u daljem radu i biti reči.

2.2.1 Ojlerova metoda diskretizacije

Ovo je najstarija i najjednostavnija metoda diskretizacije običnih diferencijalnih jednačina. **Ojlerova metoda** data je sledećom rekurzivnom jednačinom

$$\tilde{\mathbf{y}}_{n+1} = \tilde{\mathbf{y}}_n + h\mathbf{f}(\tilde{\mathbf{y}}_n). \quad (2.14)$$

Ojlerova metoda se može interpretirati na više načina, a neki od njih su:

- (1) direktna ekstrapolacija lokalnog nagiba kroz tačku $(t_n, \tilde{\mathbf{y}}_n)$;
- (2) Tejlorov razvoj prvog reda;
- (3) definicija izvoda preko limesa, ali bez granične vrednosti.

Postoji i stohastički pandan ove metode o kojem će biti reč nešto kasnije. Sada ćemo definisati numerički tok, prokomentarisati kako se meri greška metode i na kraju uvesti metode na osnovu deljenja vektorskog polja koje će biti ključne za master rad.

2.2.2 Numerički tok vektorskog polja

Iz definicije toka vektorskog polja (2.2) vidimo da za dve uzastopne tačke t_n i t_{n+1} na razdaljini h važi

$$\mathbf{y}(t_{n+1}) = \Phi_h(\mathbf{y}(t_n)). \quad (2.15)$$

Analogno, Ojlerovu metodu možemo videti kao iteraciju preslikavanja

$$\Psi_h(\mathbf{y}) = \mathbf{y} + h\mathbf{f}(\mathbf{y}), \quad (2.16)$$

koja aproksimira tok Φ_h . Za preslikavanje Φ_h kažemo da je numerički tok i definišemo ga kao

$$\tilde{\mathbf{y}}_{n+1} = \Psi_h(\tilde{\mathbf{y}}_n). \quad (2.17)$$

Za razliku od običnog toka vektorskog polja Φ_h , numerički tok Ψ_h ne formira grupu i možemo odmah videti da

$$\Psi_h \circ \Psi_h \neq \Psi_{2h}. \quad (2.18)$$

Ovo je glavna razlika između običnog i numeričkog toka vektorskog polja.

2.2.3 Red konvergencije

Kao što je prikazano, numerička metoda je aproksimacija tačnog toka diferencijalne jednačine. Uzevši je u obzir kao takvu, uočavamo da uvek postoji greška metode. Definišemo **globalnu grešku** posle n koraka kao razliku između diskretizovane aproksimacije i tačnog rešenja, tj.

$$\mathbf{e}_n = \tilde{\mathbf{y}}_n - \mathbf{y}(t_n). \quad (2.19)$$

Za aproksimaciju želimo da ima što manju grešku, pa bi bilo poželjno da za svaki korak simulacije ima ograničenu globalnu grešku, tj.

$$\max_{n=0,\dots,N} \|\mathbf{e}_n\| \leq \delta, \quad (2.20)$$

za unapred izabran prag tolerancije δ . Za dato vektorsko polje $\mathbf{f}$, inicijalnu vrednost $\mathbf{y}_0$ i vremenski interval T , imamo samo jedan slobodan parametar - dužinu koraka $h = T/N$ koja može da se menja tako da se osigura da globalna greška ne pređe prag tolerancije.

Ako je vektorsko polje $\mathbf{f}$ Lipšicovo, za metodu kažemo da **konvergira** ako za svako T

$$\lim_{h \rightarrow 0} \max_{n=0,\dots,N} \|\mathbf{e}_n\| = 0. \quad (2.21)$$

Definišemo **lokalnu grešku** numeričke metode diskretizacije kao razliku između numeričkog i tačnog toka, tj.

$$\text{le}(\mathbf{y}, h) = \Psi_h(\mathbf{y}) - \Phi_h(\mathbf{y}). \quad (2.22)$$

Lokalna greška meri grešku posle jednog koraka dužine h . Ako postoji konstanta C koja zavisi od $\mathbf{y}(t)$ i njegovih izvoda, ali ne i od h , tako da važi

$$\|\text{le}(\mathbf{y}, h)\| \leq Ch^{p+1}, \quad (2.23)$$

za neko $p \geq 1$, onda kažemo da je metoda **postojana**. Prethodna nejednakost se može zapisati i pomoću notacije veliko o, tj.

$$\|\mathbf{le}(\mathbf{y}, h)\| = \mathcal{O}(h^{p+1}). \quad (2.24)$$

Globalnu grešku možemo zapisati kao razliku numeričkog i tačnog toka na sledeći način

$$\begin{aligned} \mathbf{e}_{n+1} &= \tilde{\mathbf{y}}_{n+1} - \mathbf{y}(t_{n+1}) \\ &= \Psi_h(\tilde{\mathbf{y}}_n) - \Phi_h(\mathbf{y}(t_n)) \\ &= \Psi_h^n(\mathbf{y}_0) - \Phi_{nh}(\mathbf{y}_0), \end{aligned} \quad (2.25)$$

gde smo iskoristili oznaku

$$\Psi_h^n = \Psi_h \circ \Psi_h^{n-1}, \quad \Psi_h^0 = id. \quad (2.26)$$

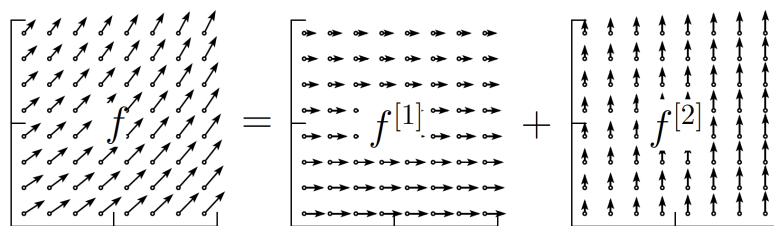
Sledeća teorema je dosta jaka i govori nam da je dovoljno da je metoda jednog koraka postojana da bi konvergirala. Ova teorema je uprošćena verzija Laksove teoreme ekvivalencije [10]. Ona će takođe biti korisna za uvođenje reda konvergencije.

Teorema 2.2.1. *Za diferencijalnu jednačinu (2.12) sa Lipšicovim vektorskim poljem $\mathbf{f}$ i postojanom numeričkom metodom jednog koraka Ψ_h , globalna greška zadovoljava*

$$\max_{n=0,\dots,N} \|\Psi_h^n(\mathbf{y}_0) - \Phi_{nh}(\mathbf{y}_0)\| = \mathcal{O}(h^p). \quad (2.27)$$

Prethodna teorema nam daje da je dovoljno da proverimo postojanost numeričke metode da bismo znali da ona konvergira. Na osnovu toga za postojanu numeričku metodu za koju važi relacija (2.24) ili (2.27) kažemo da konvergira sa **redom konvergencije** p . Ojlerov metod ima red konvergencije 1.

2.2.4 Metode deljenja vektorskog polja



Slika 2.3: Primer deljenja vektorskog polja¹

Ove metode su verovatno najvažnije za razvoj algoritama aproksimacije u kontekstu diferencijalnih jednačina. Pretpostavimo da želimo da rešimo diferencijalnu jednačinu

$$\mathbf{z}' = \mathbf{f}(\mathbf{z}) = \mathbf{f}^{[1]}(\mathbf{z}) + \mathbf{f}^{[2]}(\mathbf{z}), \quad (2.28)$$

¹Slika preuzeta iz knjige [6].

gde obe diferencijalne jednačine

$$\mathbf{z}' = \mathbf{f}^{[1]}(\mathbf{z}), \quad (2.29)$$

$$\mathbf{z}' = \mathbf{f}^{[2]}(\mathbf{z}) \quad (2.30)$$

moгу da se reše analitički. Primetimo da to ne znači da je početna jednačina analitički rešiva. Na slici 2.3 prikazan je generički primer kako možemo podeliti vektorsko polje.

Za svaku tačku u faznom prostoru vektorsko polje može biti podeljeno na dve komponente $\mathbf{f}^{[1]}$ i $\mathbf{f}^{[2]}$. Ovaj metod može biti konstruisan tako što prvo napravimo jedan korak u pravcu prvog vektorskog polja prateći njegovu integralnu krivu, a potom napravimo sledeći korak u pravcu drugog polja i odgovarajuće integralne krive za svaki korak dužine h . Preciznije, možemo krenuti od tačke $\mathbf{z}_n$ i rešiti sledeći Košijev problem

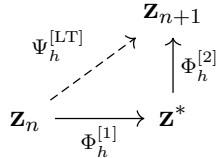
$$\mathbf{z}' = \mathbf{f}^{[2]}(\mathbf{z}), \quad \mathbf{z}(0) = \tilde{\mathbf{z}}_n, \quad (2.31)$$

na intervalu dužine h . Sada smo u tački $\mathbf{z}^* = \mathbf{z}^{[2]}(h, \tilde{\mathbf{z}}_n)$, gde $\mathbf{z}^{[2]}$ predstavlja tačno rešenje diferencijalne jednačine (2.30). Sada krećemo iz tačke $\mathbf{z}^*$ i rešavamo sledeći Košijev problem

$$\mathbf{z}' = \mathbf{f}^{[1]}(\mathbf{z}), \quad \mathbf{z}(0) = \mathbf{z}^*, \quad (2.32)$$

na intervalu dužine h . Time dobijamo sledeću tačku, $\mathbf{z}_{n+1} = \mathbf{z}^{[1]}(h, \mathbf{z}^*)$. Ako označimo sa $\Phi_h^{[1]}$ i $\Phi_h^{[2]}$ tačne tokove vektorskih polja $\mathbf{f}^{[1]}$ i $\mathbf{f}^{[2]}$, redom, onda se prethodni proces može zapisati kao

$$\mathbf{z}_{n+1} = \Psi_h^{LT}(\tilde{\mathbf{z}}_n) = \Phi_h^{[1]}(\Phi_h^{[2]}(\tilde{\mathbf{z}}_n)) = (\Phi_h^{[1]} \circ \Phi_h^{[2]})(\tilde{\mathbf{z}}_n). \quad (2.33)$$



Slika 2.4: Prikaz kompozicije Li-Troterove metode deljenja

Opisani metod se naziva **metodom deljenja vektorskog polja**, preciznije **Li-Troterov metod deljenja**. Na slici 2.4 prikazan je dijagramski prikaz Li-Troterovog deljenja. Moguće je zameniti mesta $\Phi_h^{[1]}$ i $\Phi_h^{[2]}$, ali iz konvencijalnih razloga to će označavati novo deljenje kada su zamenjena mesta i vektorskih polja $\mathbf{f}^{[1]}$ i $\mathbf{f}^{[2]}$.

Prema [1], neke od najznačajnijih prednosti ovih metoda su

- laki su za implementaciju;
- najčešće su eksplicitne metode (ne zahtevaju poznavanje budućih tačaka za ocenu prošlih);
- efikasni su u pogledu memorijske efikasnosti;

- i možda najvažnije: čuvaju strukturne osobine tačnog rešenja što ih čini superiornijim u odnosu na druge numeričke metode, pogotovo kada se radi o dužim vremenskim intervalima. Neki od primera ovih strukturnih osobina su: simpleksičnost, čuvanje zapremine, vremenska simetričnost i čuvanje prvih integrala.

Pitanje koje se nameće jeste koji je red konvergencije ovih metoda. Sledeća teorema nam daje odgovor na isto.

Teorema 2.2.2. *Li-Troterova metoda deljenja (2.33) za diferencijalnu jednačinu (2.28) ima lokalnu grešku*

$$le(\mathbf{z}, h) = \frac{h^2}{2} [\mathbf{f}^{[1]}, \mathbf{f}^{[2]}] + \mathcal{O}(h^3), \quad (2.34)$$

gde $[\mathbf{f}, \mathbf{g}]$ označava **komutator** vektorskih polja $\mathbf{f}$ i $\mathbf{g}$, što je još jedno vektorsko polje dato sa

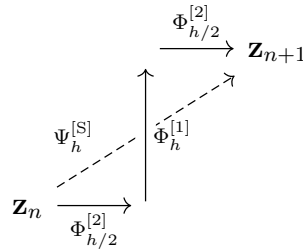
$$[\mathbf{f}, \mathbf{g}] = (\nabla \mathbf{f})\mathbf{g} - (\nabla \mathbf{g})\mathbf{f}, \quad (2.35)$$

gde je $\nabla \mathbf{f}$ Jakobijan vektorskog polja $\mathbf{f}$.

Ako vektorska polja $\mathbf{f}^{[1]}$ i $\mathbf{f}^{[2]}$ komutiraju, tj. ako je $[\mathbf{f}^{[1]}, \mathbf{f}^{[2]}] = 0$, onda metoda deljenja daje tačno rešenje. Međutim, ovakvi slučajevi nisu previše zanimljivi.

Prethodna teorema nam daje da je red konvergencije Li-Troterove metode deljenja 1, isti kao i Ojlerova metoda. Međutim, ova metoda je i dalje bolja od Ojlerove jer čuva strukturne osobine. Sa druge strane, metode deljenja mogu da povećaju red konvergencije tako što se u kompoziciju tokova uvede još tokova oblika $\Phi_{a_{ij}h}^{[i]}$. Jedan primer takvog deljenja je poznat kao **Strangovo deljenje** gde je numerički tok vektorskog polja dat sa

$$\Psi_h^{[S]} = \Phi_{h/2}^{[2]} \circ \Phi_h^{[1]} \circ \Phi_{h/2}^{[2]}. \quad (2.36)$$



Slika 2.5: Prikaz kompozicije Strangove metode deljenja

Može se pokazati da Strangova metoda deljenja ima red konvergencije 2. Na slici 2.5 prikazana je Strangova metoda deljenja.

U slučaju da analitičko rešenje za diferencijalne jednačine posle deljenja vektorskog polja ne postoji, moramo pribeći numeričkom aproksimiranju tih rešenja. Kada komponujemo numeričke tokove umesto tačnih tokova, te metode nazivamo **metodama kompozicije**. O ovome neće biti dalje reči u master radu jer ćemo pretpostavljati da vektorsko polje možemo podeliti tako da su sve dobijene diferencijalne jednačine analitički rešive.

U radu [1] je rečeno da metode deljenja ili komponovanja uključuju tri koraka:

- (i) biranje skupa vektorskih polja $\mathbf{f}^{[i]}$ tako da $\sum_i \mathbf{f}^{[i]} = \mathbf{f}$;
- (ii) rešavanje jednačina $\mathbf{z}' = \mathbf{f}^{[i]}(\mathbf{z})$ analitičkom ili numeričkom metodom;
- (iii) kombinovanje dobijenih rešenja tako da se konstruiše numeričko rešenje početnog problema.

Pretpostavimo da tok polazne diferencijalne jednačine formira određenu grupu difeomorfizama. Ako je $\mathbf{f}$ podeljena prema koraku (i), svi tokovi novonastalih diferencijalnih jednačina formiraju istu grupu difeomorfizama, a njihova rešenja mogu biti analitički dobijena (korak (ii)), tada pri komponovanju tih rešenja prema koraku (iii) dobijamo numeričku aproksimaciju rešenja polazne jednačine koja nasleđuje geometrijske osobine tačnog rešenja. Ova razmatranja (sa određenim modifikacijama) važe i kada polazno vektorsko polje formira semigrupu ili simetrični prostor.

Što se tiče koraka (i) i (ii), dodaćemo još par komentara. Prvo, za neke klase običnih diferencijalnih jednačina deljenje vektorskog polja može da se uradi sistematično za svako $\mathbf{f}$, dok u nekim drugim slučajevima nije poznato da li postoji opšti obrazac deljenja, pa se deljenje vrši po dogovoru od slučaja do slučaja. Drugo, nekada je poznata standardna procedura deljenja za dato vektorsko polje $\mathbf{f}$, ali postoje druge metode deljenja koje su više efikasne. Treće, ako prvobitna jednačina poseduje određena geometrijska svojstva koja želimo da sačuvamo, neke metode deljenja će sačuvati jedna, a druge neka druga svojstva. Nije uvek moguće sačuvati sva svojstva jednom metodom.

Ovo je bilo sve što će nam biti potrebno za dalju analizu problema iz ugla numeričke matematike i diferencijalnih jednačina. Sada ćemo se podsetiti određenih stvari iz oblasti verovatnoće i statistike kako bismo bili spremni da se vratimo problemu rada.

2.3 Elementi verovatnoće i statistike

U ovom poglavlju ćemo se priseliti osnovnih stvari u vezi sa slučajnim procesima, pogotovo u vezi sa Braunovim kretanjem. To će nam poslužiti za uvod u stohastičke diferencijalne jednačine. Tu ćemo pomenuti linearnu, tj. najjednostavniju stohastičku diferencijalnu jednačinu. Ona će nam biti glavna komponenta u deljenju vektorskog polja stohastičke diferencijalne jednačine. Na kraju ćemo se priseliti nekih statističkih pojmova, kao što su ocene, postojanost i testiranje hipoteza. Za početak, podsetimo se pojedinih elementarnosti iz teorije verovatnoća, kao što su konvergenције slučajnih veličina.

Neka je od sada pa nadalje $(\Omega, \mathcal{F}, \mathbb{P})$ **prostor verovatnoća** na kom ćemo raditi.

2.3.1 Konvergenције slučajnih veličina

Neka je $\{X_n\}_{n \in \mathbb{N}}$ niz slučajnih veličina sa odgovarajućim funkcijama raspodele $\{F_n\}_{n \in \mathbb{N}}$. Ako slučajna veličina X ima funkciju raspodele F i ako za svako $x \in \mathbb{R}$ važi

$$\lim_{n \rightarrow \infty} F_n(x) = F(x), \quad (2.37)$$

onda kažemo da niz X_n **konvergira u raspodeli** ka X i pišemo $d - \lim_{n \rightarrow \infty} X_n = X$ ili $X_n \xrightarrow{d} X$. Ova konvergencija znači jedino da raspodele konvergiraju, ali zapravo ne govori ništa o slučajnim veličinama. Zbog toga nam treba jača definicija od prethodne. Kažemo da niz X_n **konvergira u verovatnoći** ka slučajnoj veličini X ako za svako $\varepsilon > 0$ važi

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0. \quad (2.38)$$

Konvergenciju u verovatnoći označavamo sa $p - \lim_{n \rightarrow \infty} X_n = X$ ili $X_n \xrightarrow{p} X$. Ova vrsta konvergencije je tzv. tačka po tačka jer označava konvergenciju verovatnoća. Ipak, i dalje je ovo jači tip konvergencije i implicira verovatnoću u raspodeli. Dva najjača tipa konvergencije su **skoro sigurna konvergencija** i **konvergencija u srednjem**, poznata i kao **L^2 -konvergencija**. Obe ove konvergencije impliciraju konvergenciju u verovatnoći. Skoro sigurna konvergencija, zapisana kao $X_n \xrightarrow{s.s.} X$, kaže da X_n konvergira ka X ako važi

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1. \quad (2.39)$$

Konvergencija u srednjem kaže da X_n konvergira u srednjem ka X ako važi

$$\mathbb{E}[X_n - X]^2 = 0, \quad (2.40)$$

a označava se sa $L^2 - \lim_{n \rightarrow \infty} X_n = X$ ili $X_n \xrightarrow{L^2} X$. Sve prethodne konvergencije se lako mogu uopštiti i u višedimenzionalom slučaju.

2.3.2 Slučajni procesi

Realni slučajni proces $\{X_\gamma \mid \gamma \in \Gamma\}$ je familija slučajnih veličina, uzima realne vrednosti i definisan je na skupu $\Gamma \times \Omega$. Dakle, slučajne veličine ove familije možemo posmatrati kao funkcije

$$\begin{aligned} X : \Gamma \times \Omega &\rightarrow \mathbb{R} \\ (\gamma, \omega) &\rightarrow X(\gamma, \omega). \end{aligned}$$

Ako je $\Gamma = \mathbb{N}$, onda imamo **slučajni proces sa diskretnim vremenom**, a ako je $\Gamma \subseteq \mathbb{R}$, onda kažemo da je **slučajni proces sa neprekidnim vremenom**. Zainteresovani smo za ove druge, preciznije kada je $\Gamma = [0, +\infty)$ i uvek razmišljamo o Γ kao o vremenskom intervalu.

Kada govorimo o ovim procesima, najčešće ćemo ih označavati sa $X = \{X_t \mid t \geq 0\}$, a ponekad da bismo izbegli nagomilvanje indeksa, označavaćemo ih sa $X(t)$ umesto X_t . Najčešće ćemo i oznaku celog procesa skratiti na X_t ili $X(t)$ i to onda kada ne bude bilo zabune da govorimo o čitavom procesu, a ne samo o jednoj slučajnoj veličini u trenutku t . Za fiksiranu vrednost ω , skup $\{X(t, \omega) \mid t \geq 0\}$ se naziva **putanjom** ili **trajektorijom** procesa i predstavlja jednu realizaciju procesa. Za fiksirano t , skup $\{X(t, \omega) \mid \omega \in \Omega\}$ predstavlja sva moguća stanja procesa u trenutku t .

Filtracija $\{\mathcal{F}_t \mid t \geq 0\}$ predstavlja rastući skup pod- σ -algebri od $\mathcal{F}$ indeksiran sa $t \geq 0$. Drugim rečima, važi

$$(\forall 0 \leq s < t) \quad \mathcal{F}_s \subset \mathcal{F}_t, \quad \mathcal{F}_0 = \{\emptyset, \Omega\}. \quad (2.41)$$

Svaki proces $X = \{X_s \mid s \geq 0\}$, za svako $t \geq 0$, može biti pridružen najmanjoj σ -algebri generisanoj procesom X do trenutka t , tj. $\mathcal{F}_t = \sigma(X_s; 0 \leq s \leq t)$. To znači da je $\mathcal{F}_t$ najmanja pod- σ -algebra od $\mathcal{F}$ koja čini slučajne veličine $X(s, \omega)$ merljivim za svako $0 \leq s \leq t$, odnosno to je najmanji skup podskupova od Ω koji omogućava da dodelimo verovatnoće svim događajima procesa X do vremena t . Ova filtracija se naziva još i **prirodnom filtracijom** procesa X .

Za slučajni proces $X = \{X_t \mid t \geq 0\}$ i filtraciju $\{\mathcal{F}_t \mid t \geq 0\}$ (ne nužno generisanu slučajnim procesom X), kažemo da je proces X adaptiran na filtraciju $\{\mathcal{F}_t \mid t \geq 0\}$ ako za svako $t \geq 0$, X_t je $\mathcal{F}_t$ -merljivo.

Očekivana vrednost i disperzija slučajnog procesa X definišu se kao

$$\mathbb{E}X_t = \int_{\Omega} X(t, \omega) d\mathbb{P}(\omega), \quad t \in [0, T], \quad (2.42)$$

$$\mathbb{D}X_t = \mathbb{E}[X_t - \mathbb{E}X_t]^2, \quad t \in [0, T]. \quad (2.43)$$

Prethodne veličine su dobro definisane. Definišemo i **funkciju kovarijacije** procesa X u vremenima s i t kao

$$\text{cov}(X_s, X_t) = \mathbb{E}[(X_s - \mathbb{E}X_s)(X_t - \mathbb{E}X_t)]. \quad (2.44)$$

Slučajna veličina $X_t - X_s$ naziva se **priraštaj** slučajnog procesa X od trenutka s do trenutka t , $s < t$.

Slučajna veličina $Z = \mathbb{E}[Y \mid \mathcal{F}]$ naziva se **uslovno očekivanje** od Y u odnosu na σ -algebru $\mathcal{F}$ ako su ispunjena sledeća dva svojstva:

- (1) Y je $\mathcal{F}$ -merljivo,
- (2) $\mathbb{E}[Y I_A] = \mathbb{E}[X I_A]$, za svako $A \in \mathcal{F}$,

gde je I_A indikator sa verovatnoćom jedan ako je $\omega \in A$, odnosno nula u suprotnom. Uсловно očekivanje, ako postoji, jedinstveno je.

Slučajni proces $X = \{X_t \mid t \geq 0\}$ je **martingal** u odnosu na filtraciju $\{\mathcal{F}_t \mid t \geq 0\}$ ako važi

- (1) $\mathbb{E}|X_t| < \infty$, za svako $t \geq 0$,
- (2) X je adaptiran na filtraciju $\{\mathcal{F}_t \mid t \geq 0\}$,
- (3) $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$, za svako $0 \leq s < t < \infty$.

Poslednji uslov znači da je X_s najbolji prediktor za X_t ako nam je poznato $\mathcal{F}_t$. Ako je jednakost u svojstvu (3) u prethodnoj definiciji zamenjena znakom „ $\geq$ ”, onda proces X nazivamo **submartingal**, a ako je znak „ $\leq$ ”, onda je X **supermartingal**. Iz svojstva uslovnog očekivanja imamo da je

$$\mathbb{E}X_s = \mathbb{E}[\mathbb{E}[X_t | \mathcal{F}_s]] = \mathbb{E}X_t, \quad (2.45)$$

što znači da martingali imaju konstantno očekivanje za svako $t \geq 0$.

2.3.3 Braunovo kretanje

Definišemo **Braunovo kretanje** ili **standardni Vinerov proces** kao $W = \{W_t | t \geq 0\}$ tako da važi

- (1) $W_0 = 0$ skoro sigurno,
- (2) W ima nezavisne priraštaje, tj. slučajne veličine $W_{t_2} - W_{t_1}$ i $W_{s_2} - W_{s_1}$ su nezavisne za svako $t_1 \leq t_2 \leq s_1 \leq s_2$,
- (3) W je Gausov proces, tj. $W_t - W_s \sim \mathcal{N}(0, t - s)$, za $0 \leq s < t$.

Iz prethodne definicije lako možemo izračunati funkciju kovarijacije za Braunovo kretanje i ona je data sa

$$\text{cov}(W_s, W_t) = \min(s, t) = s \wedge t. \quad (2.46)$$

Za dati korak $h > 0$ vrlo lako možemo simulirati trajektoriju Braunovog kretanja na intervalu $[0, T]$ koristeći da je

$$W_{t+h} - W_t \sim \mathcal{N}(0, h) \sim \sqrt{h}\mathcal{N}(0, 1). \quad (2.47)$$

Još jedna varijacija Braunovog kretanja koja će nam koristiti jeste **Braunov most**. Braunov most B_t je slučajni proces sa neprekidnim vremenom čija raspodela je zadata kao uslovna raspodela Braunovog kretanja W_t , pod uslovom da je $W_T = 0$, tako da je proces zakucan za početak u početnom i krajnjem trenutku, $t = 0$ i $t = T$. Formalno,

$$B_t = (W_t | W_T = 0), \quad t \in [0, T]. \quad (2.48)$$

Očekivanje Braunovog mosta je 0, a disperzija je $\frac{t(T-t)}{T}$, što znači da je najveća raspršenost na sredini mosta, a bez raspršenosti na krajevima. Funkcija kovarijacije je data sa $\text{cov}(B_s, B_t) = \min(s, t) - \frac{st}{T}$. Priraštaji ovog procesa nisu nezavisni.

2.3.4 Totalna i kvadratna varijacija

Jedno od najvažnijih svojstava Braunovog kretanja jeste da ima nigde diferencijabilne trajektorije. Ovo možemo videti preko totalne varijacije procesa W koja je definisana kao granična vrednost u verovatnoći

$$V_t(W) = p - \lim_{\|\Pi_n\| \rightarrow 0} \sum_{k=0}^{n-1} |W(t_{k+1}) - W(t_k)|, \quad (2.49)$$

gde je Π_n particija intervala $[0, t]$ u n manjih podintervala $0 = t_0 < t_1 < \dots < t_n = t$. Procesi sa diferencijabilnim trajektorijama imaju ograničenu totalnu varijaciju. Međutim, za Braunovo kretanje W važi $V_t(W) = +\infty$. To znači da W ima nigde diferencijabilne trajektorije.

Drugi indikator nigde diferencijabilnih trajektorija jeste i kvadratna varijacija Braunovog kretanja koja nije 0. U slučaju da su trajektorije diferencijabilne, kvadratna varijacija bi morala biti nula. Nju definišemo kao

$$[W, W]_t = p - \lim_{\|\Pi_n\| \rightarrow 0} \sum_{k=0}^{n-1} |W(t_{k+1}) - W(t_k)|^2. \quad (2.50)$$

Za Braunovo kretanje W važi da je

$$[W, W]_t = t, \quad t \geq 0. \quad (2.51)$$

2.3.5 Stohastičke diferencijalne jednačine

Kako bismo što bolje razumeli definiciju stohastičke diferencijalne jednačine i Itoovog integrala, prvo predstavljamo heuristiku iza notacije. Krećemo od veličine X_t koja slučajno varira kroz vreme $t \geq 0$. Razmotrimo sada varijaciju $\Delta X_t = X_{t+\Delta t} - X_t$ na malom intervalu $[t, t + \Delta t)$. Brzina promene veličine X je dinamika definisana kao odnos $\Delta X_t / \Delta t$ i možemo je razložiti na

$$\frac{\Delta X_t}{\Delta t} = \text{deterministički deo} + \text{stohastički deo}.$$

Možemo pretpostaviti, za početak, da je deterministički deo konstantna veličina μ , a za stohastički deo pretpostavimo da je povezan sa varijacijom nekog izvora šuma. Označimo sa $\Delta W_t = W_{t+\Delta t} - W_t$ varijaciju procesa šuma i pretpostavimo da je proporcijalan konstanti $\sigma > 0$. Prirodna hipoteza bi bila da pretpostavimo Gausovsko ponašanje šuma, što implicira da bi W trebalo biti Braunovo kretanje. Time smo brzinu promene od X opisali sledećom jednačinom

$$\frac{\Delta X_t}{\Delta t} = \mu + \sigma \frac{\Delta W_t}{\Delta t}. \quad (2.52)$$

Ako sada razmatramo infinitezimalne intervale, tj. pustimo da $\Delta \rightarrow 0$, prethodna jednačina dobija oblik

$$X'_t = \mu + \sigma W'_t. \quad (2.53)$$

Međutim, znamo da ovo nema smisla jer ne možemo diferencirati Braunovo kretanje W . Alternativa bi bila zapisati prethodnu jednačinu u obliku diferencijalne forme

$$dX_t = \mu dt + \sigma dW_t. \quad (2.54)$$

Ovo i dalje nema nikakvo matematičko značenje jer znamo da totalna varijacija dW_t nije konačna. Da bismo dali prethodnoj jednačini značenje, moramo načiniti još jedan korak i prebaciti se na integralnu formu

$$X_t = X_0 + \int_0^t \mu du + \int_0^t \sigma dW_u, \quad (2.55)$$

gde smo pretpostavili da veličina X ima početnu vrednost X_0 u trenutku $t = 0$. Prethodna jednačina uvodi **stohastički** ili **Itoov integral** slučajne veličine X

$$I(X) = \int_0^t X_t dW_t \quad (2.56)$$

u odnosu na Braunovo kretanje. Sada je poželjno da konstruišemo ovaj integral i damo intuiciju prethodnom razmatranju. Prirodno, jednačinu (2.55) bismo rešili kao

$$X_t = X_0 + \mu t + \sigma W_t, \quad (2.57)$$

gde smo pretpostavili da je

$$\int_0^t \sigma dW_u = \sigma(W_t - W_0) = \sigma W_t. \quad (2.58)$$

Naravno, obe vrednosti μ i σ mogu biti funkcije od t ili X_t , pa nam je potrebna uopštena definicija za $I(Y)$. U jednostavnom slučaju, kada je X deo-po-deo konstantan proces, možemo iskoristiti isto što i malopre da rešimo jednačinu (2.55). U suprotnom, kada X nije prost proces, definišemo integral $I(X)$ kao limes integrala $I(X^{(n)})$, gde su $X^{(n)}$ **prosti procesi** definisani sa

$$X^{(n)}(t, \omega) = X(t_j, \omega), \quad t_j \leq t \leq t_{j+1}, \quad (2.59)$$

za $t_j \in \Pi_n$, gde je Π_n particija intervala $[0, T]$, tako da maksimalna dužina podintervala $\|\Pi_n\|$ teži 0, kad $n \rightarrow \infty$. Tada definišemo $I(X^{(n)})$ kao

$$\begin{aligned} I(X^{(n)}) &= \sum_{j=0}^{n-1} X^{(n)}(t_j) (W(t_{j+1}) - W(t_j)) \\ &= \sum_{j=0}^{n-1} X(t_j) (W(t_{j+1}) - W(t_j)). \end{aligned} \quad (2.60)$$

Na prethodnu jednakost ne možemo primeniti limes jer znamo da W nema ograničenu varijaciju. Sa druge strane, znamo da Braunovo kretanje ima osobinu $\mathbb{E}[W_t - W_s]^2 = t - s$, za $t > s$, pa važi

$$\mathbb{E}[I(X^{(n)})]^2 = \sum_{j=0}^{n-1} \mathbb{E}[X^{(n)}(t_j)]^2 (t_{j+1} - t_j). \quad (2.61)$$

Može se pokazati da $X^{(n)}$ konvergira u srednjem ka X , pa zajedno sa prethodnom jednačinom možemo pokazati da

$$I(X^{(n)}) \xrightarrow{L^2} I(X). \quad (2.62)$$

Iz prethodne, grube, konstrukcije Itoovog integrala možemo izvući neke od osobina koje slučajna veličina X mora da zadovoljava. Za početak, X mora biti adaptiran na prirodnu filtraciju Braunovog kretanja, tj. X_t mora biti $\mathcal{F}_t$ -merljiv, za svako $t \geq 0$. Dodatno,

ponašanje podintegralne funkcije mora da kompenzuje „čudno” ponašanje Braunovog kretanja. To znači da mora da važi

$$\int_0^t \mathbb{E}[X_u]^2 dU < \infty. \quad (2.63)$$

Neka od svojstava Itoovog integrala su:

1. Ako je X Ito integrabilan, onda je

$$\mathbb{E} \left[\int_0^T X_t dW_t \right] = 0, \quad (2.64)$$

$$\mathbb{D} \left(\int_0^T X_t dW_t \right) = \int_0^T \mathbb{E} X_t^2 dt \quad (\text{Itoova izometrija}). \quad (2.65)$$

2. (**Linearnost**) Ako su X i Y dva Ito integrabilna procesa, a a i b dve konstante, onda je

$$\int_0^T (aX_t + bY_t) dW_t = a \int_0^T X_t dW_t + b \int_0^T Y_t dW_t. \quad (2.66)$$

3. Ako je X Ito integrabilan, onda je proces

$$M_t = M_0 + \int_0^t X_s dW_s \quad (2.67)$$

martingal sa očekivanjem M_0 , gde je M_0 konstanta.

2.3.6 Stohastičke difuzije

Itoov proces $\{X_t \mid t \geq 0\}$ je slučajni proces koji može da se napiše u formi

$$X_t = X_0 + \int_0^t G_u du + \int_0^t H_u dW_u, \quad (2.68)$$

gde su $G, H : [0, T] \times \Omega \rightarrow \mathbb{R}$ dve **progresivno merljive** funkcije (tj. $\mathcal{B}[0, T] \otimes \mathcal{F}_t$ -merljive, gde je $\mathcal{B}[0, T]$ Borelova σ -algebra), takve da je

$$\int_0^T |G_t| dt < \infty \text{ s.s.} \quad (2.69)$$

$$\int_0^T H_t^2 dt < \infty \text{ s.s.} \quad (2.70)$$

Proces difuzije ili **stohastička difuzija** ili **slučajna difuzija**, proces je koji rešava sledeću stohastičku diferencijalnu jednačinu

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t, \quad (2.71)$$

sa početnim uslovom X_0 . Kao i do sad, prethodnu jednačinu interpretiramo u Itoovom integralnom smislu kao

$$X_t = X_0 + \int_0^t \mu(u, X_u) du + \int_0^t \sigma(u, X_u) dW_u. \quad (2.72)$$

Početna vrednost X_0 može biti slučajna ili ne. Ako je slučajna veličina, onda mora biti nezavisna od σ -algebre generisane Braunovim kretanjem W i mora da važi $\mathbb{E}X_0^2 < \infty$. Funkcije μ i σ nazivaju se funkcijama **drifta** i **difuzije**, redom. Kroz ceo test, čak i kada nije naznačeno, pretpostavljamo da je

$$\int_0^T \sup_{|x| \leq R} \{|\mu(t, x)| + \sigma^2(t, x)\} dt < \infty \text{ s.s.} \quad (2.73)$$

za sve $T, R \in [0, \infty)$. Prethodno se pretpostavlja da bi proces (2.72) bio Itoov proces.

Sledeća teorema nam daje uslove za postojanje i jedinstvenost stohastičke difuzije.

Teorema 2.3.1. *Neka stohastička diferencijalna jednačina (2.71) zadovoljava sledeća dva uslova*

- (1) (**Globalna Lipšicovost**) *Za sve $x, y \in \mathbb{R}$, i sve $t \in [0, T]$, postoji konstanta $K < \infty$ takva da je*

$$|\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| < K|x - y|. \quad (2.74)$$

- (2) (**Linearni rast**) *Za sve $x \in \mathbb{R}$, i sve $t \in [0, T]$, postoji konstanta $C < \infty$ takva da je*

$$|\mu(x, t)| + |\sigma(t, x)| < C(1 + |x|^2). \quad (2.75)$$

*Tada postoji jedinstveno, neprekidno i adaptirano **jako rešenje** X stohastičke diferencijalne jednačine (2.71) sa početnim uslovom X_0 , takvo da važi*

$$\mathbb{E} \left[\int_0^T X_t^2 \right] < \infty. \quad (2.76)$$

Prethodni rezultat nam kaže da je rešenje X jakog tipa. To u stvari implicira da je trajektorija rešenja jedinstvena, odnosno da je

$$\mathbb{P} \left(\sup_{t \geq 0} |X^{(1)} - X^{(2)}| = 0 \right) = 1, \quad (2.77)$$

gde su $X^{(1)}$ i $X^{(2)}$ jaka rešenja sa istim početnim uslovom. Takođe je moguće dobiti i **slabo rešenje** pod oslabljenim uslovima. U mnogo slučajeva u statistici uslovi za slabo rešenje su dovoljni jer to zapravo znači da dva slaba rešenja $X^{(1)}$ i $X^{(2)}$ nemaju identične trajektorije (čak ne moraju biti definisani ni na istim prostorima verovatnoća), ali da su njihove raspodele iste, što je dovoljno za statističko zaključivanje. Naravno, jako rešenje jeste i slabo rešenje, dok suprotno ne važi.

Možemo primetiti da su prethodni uslovi za jako rešenje previše restriktivni i često neispunjivi. Postoje oslabljeni uslovi koji garantuju postojanje i jedinstvenost jakog rešenja, ali o njima neće biti reči u ovom radu jer zahtevaju dodatne definicije koje nisu previše važne za isti. Za više informacija o ovim uslovima pogledati u knjizi [8]. U nastavku ćemo predstaviti još i uslove za slabo rešenje.

U suštini smo zainteresovani za autonomnu verziju stohastičke diferencijalne jednačine (2.71), gde obe funkcije μ i σ ne zavise od vremena t . Drugim rečima, zanimaju nas diferencijalne jednačine oblika

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t. \quad (2.78)$$

Sledeća teorema govori o tome kada ovakve stohastičke diferencijalne jednačine imaju jedinstveno slabo rešenje.

Teorema 2.3.2. *Stohastička diferencijalna jednačina (2.78) ima jedinstveno slabo rešenje ako je μ lokalno ograničeno, σ neprekidna funkcija i ako postoji konstanta A takva da je*

$$x\mu(x) + \sigma^2(x) \leq A(1 + x^2). \quad (2.79)$$

2.3.7 Itoova formula

Važan alat u stohastičkoj analizi, koji je često koristan i u simulacijama, jeste **Itoova formula** ili **Itoova lema**. Ova formula se može shvatiti kao stohastička verzija Tejlorovog razvoja funkcije $f(X)$ zaključno sa drugim redom, gde je X stohastička difuzija. Itoova lema kaže da za dvostruko diferencijabilnu funkciju $f(t, x)$, po oba argumenta, važi

$$f(t, X_t) = f(0, X_0) + \int_0^t f_t(u, X_u)du + \int_0^t f_x(u, X_u)dX_u + \frac{1}{2} \int_0^t f_{xx}(u, X_u)(dX_u)^2, \quad (2.80)$$

gde smo koristili sledeće oznake za parcijalne izvode

$$f_t(t, x) = \frac{\partial f}{\partial t}(t, x), \quad (2.81)$$

$$f_x(t, x) = \frac{\partial f}{\partial x}(t, x), \quad (2.82)$$

$$f_{xx}(t, x) = \frac{\partial^2 f}{\partial x^2}(t, x). \quad (2.83)$$

Itoova formula (2.80) može se zapisati i u diferencijalnom obliku

$$df(t, X_t) = f_t(t, X_t)dt + f_x(t, X_t)dX_t + \frac{1}{2}f_{xx}f_{xx}(t, X_t)(dX_t)^2. \quad (2.84)$$

Uprkos naizgled jednostavnom obliku Itoove formule, u opštem slučaju je teško razviti član $(dX_t)^2$ bez dodatnog znanja o difuziji. Treba imati na umu da su članovi $dt dW_t$ i $(dt)^2$ reda $\mathcal{O}((dt)^2)$, što znači da za Itoovu formulu kada razvijamo $(dX_t)^2$, sve članove u kojima su diferencijali $dt dW_t$ ili $(dt)^2$ možemo zanemariti. Takođe, iz osobina Braunovog kretanja znamo da se $(dW_t)^2$ ponaša kao dt , pa u Itoovoj formuli možemo svuda zameniti $(dW_t)^2$ sa dt .

2.3.8 Ornštajn-Ulenbekov proces

Ornštajn-Ulenbekov proces je jedinstveno jako rešenje sledeće stohastičke diferencijalne jednačine

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 dW_t, \quad X_0 = x_0, \quad (2.85)$$

gde su $\theta_1, \theta_2 \in \mathbb{R}$ i $\theta_3 > 0$ parametri. Za $\theta_2 > 0$, Ornštajn-Ulenbekov proces **se vraća ka proseku**, što znači da proces teži da oscilira oko nekog fiksnog stanja. Još jedno interesantno svojstvo ovog procesa jeste da, za razliku od Braunovog kretanja, ovaj proces ima konačnu disperziju za sve $t \geq 0$. Još jedna česta parametrizacija ovog procesa data je sledećom jednačinom

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, \quad X_0 = x_0. \quad (2.86)$$

Parametre ovog procesa interpretiramo na sledeći način. Parametar $\sigma > 0$ predstavlja volatilnost procesa, μ predstavlja krajnji ekvilibrijum procesa, a θ brzinu dostizanja tog ekvilibrijuma. Primenom Itoove leme možemo pokazati da je krajnje rešenje dato sa

$$X_t = \mu + (x_0 - \mu)e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-u)} dW_u. \quad (2.87)$$

Za $t \geq 0$, Ornštajn-Ulenbekov proces ima Gausovu raspodelu prelaska $p(x_{t+h} | x_t)$, tj. gustina raspodele od X_{t+h} pod uslovom da je $X_t = x_t$ normalna raspodela sa očekivanjem i disperzijom

$$\mathbb{E}[X_{t+h} | X_t = x_t] = \mu + (x_t - \mu)e^{-\theta h}, \quad (2.88)$$

$$\mathbb{D}[X_{t+h} | X_t = x_t] = \sigma^2 \frac{1 - e^{-2\theta h}}{2\theta}. \quad (2.89)$$

Pod uslovom da je $X_0 = x_0$ konstanta, uslovna funkcija kovarijacije data je sa

$$\text{cov}(X_s, X_t | X_0 = x_0) = \frac{\sigma^2}{2\theta} (e^{-\theta|t-s|} - e^{-\theta(t+s)}). \quad (2.90)$$

Ornštajn-Ulenbekov proces je primer Gausovog procesa koji ima ograničenu disperziju i stacionarnu raspodelu za razliku od Braunovog kretanja. Takođe, postoji razlika i u driftu: za Braunovo kretanje on je konstantan, dok kod Ornštajn-Ulenbekovog procesa zavisi od trenutnog stanja procesa.

Ovaj proces je rešenje linearne stohastičke diferencijalne jednačine koju možemo u potpunosti da rešimo, isto kao i u slučaju linearne obične diferencijalne jednačine. Zbog toga će ovo biti važan proces u daljem pokušaju da napravimo matematički alat za ocenu parametara na velikoj klasi modela zasnovanih na stohastičkim difuzijama. Pre nego što se podsetimo nekih osnovnih pojmova iz statistike, uvedimo, dodatno, sve prethodno ključne pojmove u više dimenzija.

2.3.9 Višedimenzionalne stohastičke difuzije

Neka je $\mathbf{X} = \{\mathbf{X}_t | t \geq 0\}$ d -dimenzionalni slučajni proces koji uzima vrednosti u $\mathbb{R}^d$, definisan na prostoru verovatnoća $(\Omega, \mathcal{F}, \mathbb{P})$. Za slučajne veličine $\mathbf{X}_t$ mora da važi da su

$\mathcal{L}^d$ -merljive, za svako $t \geq 0$, gde su sada $\mathcal{L}^d$ Lebegovi podskupovi od $\mathbb{R}^d$. Ovo se razlikuje od jednodimenzionalnog slučaja kada su bili dovoljni Borelovi podskupovi.

Najjednostavniji d -dimenzionalni slučajni proces sa nula driftom $\boldsymbol{\mu} = \mathbf{0}$ i jediničnom matricom difuzije $\boldsymbol{\Sigma} = \mathbf{I}$, jeste **d -dimenzionalno Braunovo kretanje** $\mathbf{W}_t = (W_t^1, W_t^2, \dots, W_t^d)$, čije su koordinate međusobno nezavisna Braunova kretanja. Sledeća teorema nam daje karakterizaciju ovog vektorskog procesa.

Teorema 2.3.3 (Dubova teorema). *Vektorski proces $\{\mathbf{X}_t = (X_t^1, X_t^2, \dots, X_t^d) \mid t \geq 0\}$ adaptiran na filtraciju $\{\mathcal{F}_t \mid t \geq 0\}$ je d -dimenzionalno Braunovo kretanje ako i samo ako, za sve $i, j = 1, 2, \dots, d$ i sve $0 \leq s \leq t$ važi*

- (1) $X_0^i = 0$, skoro sigurno,
- (2) $\mathbb{E}[X_t^i - X_s^i] = 0$,
- (3) $\mathbb{E}[(X_t^i - X_s^i)(X_t^j - X_s^j)] = \delta_i^j(t - s)$,

gde je δ_i^j **Kronekerov delta simbol** definisan sa $\delta_i^j = 1$, za $i = j$, odnosno $\delta_i^j = 0$ u suprotnom.

Neka je $\mathbf{W}$ m -dimenzionalno Braunovo kretanje. Posmatrajmo d -dimenzionalnu vektorsku funkciju $\boldsymbol{\mu} : [0, T] \times \Omega \rightarrow \mathbb{R}^d$, za čije koordinate važi $\sqrt{|\mu^k|} \in \mathcal{L}_T^2$, za sve $k = 1, 2, \dots, d$ i $d \times m$ matricnu funkciju $\boldsymbol{\Sigma} : [0, T] \times \Omega \rightarrow \mathbb{R}^{d \times m}$, za čije koordinate važi $\Sigma^{i,j} \in \mathcal{L}_T^2$, za $i = 1, 2, \dots, d$ i $j = 1, 2, \dots, m$. U analogiji sa realnim slučajem, označićemo sa $\boldsymbol{\mu}_t$ i $\boldsymbol{\Sigma}_t$ vektorsku i matricnu slučajnu veličinu u trenutku t . Tada simbolički označavamo **d -dimenzionalnu stohastičku diferencijalnu jednačinu** sa

$$d\mathbf{X}_t = \boldsymbol{\mu}_t dt + \boldsymbol{\Sigma}_t d\mathbf{W}_t. \quad (2.91)$$

Prethodnu jednačinu interpretiramo preko vektorske integralne forme

$$\mathbf{X}_t = \mathbf{X}_s + \int_s^t \boldsymbol{\mu}_u du + \int_s^t \boldsymbol{\Sigma}_u d\mathbf{W}_u, \quad 0 \leq s \leq t \leq T, \quad (2.92)$$

koju interpretiramo koordinatno, tj.

$$X_t^k = X_s^k + \int_s^t \mu_u^k du + \sum_{j=1}^m \int_s^t \Sigma_u^{k,j} dW_u^j, \quad k = 1, 2, \dots, d. \quad (2.93)$$

Za unapred definisano $\mathcal{F}_0$ -merljivo $\mathbf{X}_0$ rešenje prethodne stohastičke diferencijalne jednačine $\mathbf{X} = \{\mathbf{X}_t = (X_t^1, X_t^2, \dots, X_t^d) \mid t \geq 0\}$ naziva se **višedimenzionalna stohastička difuzija**. Vektorsku funkciju $\boldsymbol{\mu}$ nazivamo **vektorom drifta** ili **pomeraja**, a $d \times d$ kvadratnu matricu $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$ **matricom difuzije**.

Neka funkcija $U : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ ima neprekidne parcijalne izvode $\frac{\partial U}{\partial t}, \frac{\partial U}{\partial x_t^k}, \frac{\partial^2 U}{\partial x_t^k \partial x_t^i}$, za $i, k = 1, 2, \dots, d$. Definišimo realni proces $\{Y_t \mid 0 \leq t \leq T\}$ sa

$$Y_t = U(t, \mathbf{X}_t) = U(t, X_t^1, X_t^2, \dots, X_t^d),$$

gde je $\mathbf{X}_t$ višedimenzionalna stohastička difuzija koja rešava jednačinu (2.91). Tada je **višekomponentna Itoova formula** data sa

$$dY_t = \left(\frac{\partial U}{\partial t} + \sum_{k=1}^d \mu_t^k \frac{\partial U}{\partial x_t^k} + \frac{1}{2} \sum_{j=1}^m \sum_{i,k=1}^d \Sigma_t^{i,j} \Sigma_t^{k,j} \frac{\partial^2 U}{\partial x_t^i \partial x_t^k} \right) dt + \sum_{j=1}^m \sum_{i=1}^d \Sigma_t^{i,j} \frac{\partial U}{\partial x_t^i} dW_t^j, \quad (2.94)$$

gde su parcijalni izvodi računati u tački $(t, \mathbf{X}_t)$.

Možemo još više generalizovati Itoovu formulu i dobiti **višedimenzionalnu Itoovu formulu** za dovoljno glatku vektorsku funkciju $\mathbf{U} : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^l$ višedimenzionalne stohastičke difuzije $\mathbf{X}_t$. Definišimo l -dimenzionalni vektorski proces $\mathbf{Y}_t = \mathbf{U}(t, \mathbf{X}_t)$. Da bismo dobili uopšteni oblik, moramo primeniti multikomponentnu Itoovu formulu na svaku koordinatu. Dakle, dobijamo

$$dY_t^p = \left(\frac{\partial U^p}{\partial t} + \sum_{k=1}^d \mu_t^k \frac{\partial U^p}{\partial x_t^k} + \frac{1}{2} \sum_{j=1}^m \sum_{i,k=1}^d \Sigma_t^{i,j} \Sigma_t^{k,j} \frac{\partial^2 U^p}{\partial x_t^i \partial x_t^k} \right) dt + \sum_{j=1}^m \sum_{i=1}^d \Sigma_t^{i,j} \frac{\partial U^p}{\partial x_t^i} dW_t^j, \quad (2.95)$$

za $p = 1, 2, \dots, l$, gde su parcijalni izvodi računati u tački $(t, \mathbf{X}_t)$.

Mogli bismo čak Itoovu formulu generalizovati u **stohastički Tejlorov razvoj**, ali za to bi nam bilo potrebno mnogo vremena i strana. Za početak bismo morali uvesti višedimenzionalne Itoove integrale. Za više informacija o ovoj temi predlažemo knjigu [9]. Međutim, iako nećemo pričati o stohastičkoj verziji Tejlorovog razvoja, ona je ključna za jednu metodu diskretizacije o kojoj će kasnije biti reči. Ovu metodu ćemo dati u krajnjoj formuli bez objašnjenja kako je ona dobijena, jer se za nju koristi ovaj razvoj.

Ostaje nam još da pomenemo **višedimenzionalnu verziju Ornštajn-Ulenbekovog procesa**. Ovaj vektorski proces je definisan kao rešenje sledeće stohastičke diferencijalne jednačine

$$d\mathbf{X}_t = \boldsymbol{\theta}(\boldsymbol{\mu} - \mathbf{X}_t)dt + \boldsymbol{\sigma}d\mathbf{W}_t, \quad (2.96)$$

gde je $\boldsymbol{\theta}$ $d \times d$ invertibilna realna matrica, $\boldsymbol{\mu}$ je d -dimenzionalni realni vektor, a $\boldsymbol{\sigma}$ je $d \times m$ realna matrica takva da je kvadratna matrica $\boldsymbol{\sigma}\boldsymbol{\sigma}^\top$ pozitivno definitna. Vektorski proces $\mathbf{W}$ je m -dimenzionalno Braunovo kretanje. Može se pokazati da d -dimenzionalni Ornštajn-Ulenbekov proces $\mathbf{X}$ ima d -dimenzionalnu normalnu raspodelu prelaska, tj.

$$\mathbf{X}(t+h) \mid \mathbf{X}(t) \sim \mathcal{N}_d(\mathbf{M}_t(h), \boldsymbol{\Sigma}(h)),$$

gde su očekivana vrednost $\mathbf{M}$ i matrica kovarijacije $\boldsymbol{\Sigma}$ date sa

$$\mathbf{M}_t(h) = \boldsymbol{\mu} + e^{-\boldsymbol{\theta}h}(\mathbf{X}(t) - \boldsymbol{\mu}), \quad (2.97)$$

$$\boldsymbol{\Sigma}(h) = \int_t^{t+h} e^{-\boldsymbol{\theta}(t-s)} \boldsymbol{\sigma}\boldsymbol{\sigma}^\top e^{-\boldsymbol{\theta}^\top(t-s)} ds. \quad (2.98)$$

Izraz za kovarijacionu matricu se izvodi koristeći Itoovu izometriju.

Na kraju ovog poglavlja dodatno ćemo se priseliti i osnovne logike iza ocene parametara, principa maksimalne verodostojnosti, testiranja hipoteza i sl.

2.3.10 Ocene parametara

Već znamo da su parametri modela presudni u razlikovanju dva modela. Na primer, linearnom regresijom možemo modelovati broj Nobelovih nagrada na osnovu broja pojedine čokolade po glavi stanovnika, ali možemo modelovati i sreću pojedinca na osnovu godišnje zarade. Oba ova prosta modela koriste linearnu regresiju, ali parametri modela će odrediti kako ta prava izgleda. Zato je ključno da u svakom modelu imamo dobro ispitan način za određivanje parametara i ispitivanje njihove efikasnosti.

U linearnoj regresiji imamo dobro poznat **metod najmanjih kvadrata** koji minimizuje sumu kvadrata reziduala. Bilo da je reč o linearnoj regresiji ili bilo kom drugom modelu mašinskog učenja, uvek postoji neka funkcija koja se minimizuje. Ove funkcije se najčešće nazivaju **funkcijama gubitka** ili **funkcijama troška**. U statistici se najčešće funkcija gubitka zasniva na funkciji verodostojnosti.

Princip maksimalne verodostojnosti zasniva se na tome da maksimizujemo verovatnoću da je upravo naš opisani model generisao podatke koje imamo. Kada maksimizujemo ovu veličinu, tada dobijamo željene parametre. U daljem radu videćemo da nekad nije moguće dovoljno brzo i efikasno optimizovati funkciju verodostojnosti, pogotovo kada je broj parametara velik (reda veličine 5), a funkcija izrazito nelinearna. U tom slučaju ćemo morati da pravimo neke smislene aproksimacije koje će učiniti problem optimizacije mnogostruko lakšim.

Dve su najčešće mere koliko je neka ocena parametra dobra.

Kažemo da je ocena parametra **postojana** ako sa povećanjem obima uzorka, ocena parametra konvergira ka pravoj vrednosti parametra. Preciznije, sa povećanjem uzorka, empirijska raspodela ocenjenog parametra postaje izrazito koncentrisana oko prave vrednosti parametra. Formalno govoreći, neka je θ_n ocena parametra θ na osnovu uzorka obima n . Ako važi

$$p - \lim_{n \rightarrow \infty} \theta_n = \theta, \quad (2.99)$$

onda kažemo da je ocena postojana.

Sa druge strane, ocena je **nepriistrasna** ako u proseku dostiže pravu vrednost parametra, tj. očekivanje empirijske raspodele ocene treba da bude jednako pravoj vrednosti parametra. Formalno zapisano, ocena parametra je nepriistrasna ako je

$$\mathbb{E}[\theta_n] = \theta. \quad (2.100)$$

2.3.11 Testovi saglasnosti

Često želimo da testiramo hipotezu o nekom parametru ili o tome koliko nam je model zapravo dobar, što je često najvažnija stvar koju možemo uraditi. Ako model nije dobro postavljen, šta god radili sa njim neće nam pomoći da dobro predvidimo i modelujemo željenu pojavu. Zato ćemo se osvrnuti na neke načine **testiranja saglasnosti** modela.

Ono što želimo da testiramo jeste da li podaci prate određenu raspodelu, a to možemo pomoću sledeće hipoteze:

$$H_0 : F = F_0 \text{ protiv } H_1 : F \neq F_0. \quad (2.101)$$

Ono što odmah možemo primetiti iz postavke problema jeste da, za razliku od testiranja hipoteze o parametrima, ovde ne želimo da odbacimo H_0 u korist H_1 . Razlog tome je jednostavan: postoji beskonačno mnogo alternativa protiv kojih možemo testirati nultu hipotezu. Iz tog razloga je aparat za testiranje ovakvih testova dosta lakši ako želimo da minimizujemo grešku prve vrste, dok je stvar znatno kompleksnija ako se bavimo greškom druge vrste i moći testa.

Zato ćemo pomenuti neke testove koji se bave testiranjem hipoteze (2.101). Pre toga uvodimo **empirijsku funkciju raspodele** definisanu sa

$$F_n(x) = \frac{1}{n} \sum_{k=1}^n \mathbf{1}(X_k \leq x), \text{ za svako } x \in \mathbb{R}. \quad (2.102)$$

Prethodna formula nije ništa drugo do broj slučajnih veličina iz prostog slučajnog uzorka $\mathbf{X}$ koje su manje od x . Zato empirijsku funkciju raspodele možemo posmatrati kao proporciju

$$F_n(x) = \frac{\#\{k = 1, 2, \dots, n \mid X_k \leq x\}}{n}, \text{ za svako } x \in \mathbb{R}.$$

Na osnovu jakog zakona velikih brojeva imamo da je za svako $x \in \mathbb{R}$,

$$F_n(x) \xrightarrow[n \rightarrow \infty]{s.s.} F(x), \quad (2.103)$$

Međutim, ovakav vid konvergencije zahteva da se prvo fiksira x , pa da se za takvo x odredi n iz definicije limesa. Znamo da definicija kaže: „za svako $\varepsilon > 0$, postoji $n(\varepsilon) \in \mathbb{N}$ takvo da ...” Kada pričamo o konvergenciji tačka-po-tačka, naše n ne zavisi samo od ε , već i od x . To nam govori da ovakva konvergencija nije dovoljno dobra za našu test statistiku jer ne želimo da nam konstrukcija, pa samim tim i osobine, test statistike zavise od podataka. Da bismo dobili konvergenciju za sve x odjednom, gde n neće zavisiti od x , potrebna nam je uniformna konvergencija koja nažalost nije garantovana zakonom velikih brojeva. Zbog toga nam treba **centralna teorema matematičke statistike**.

Teorema 2.3.4 (Glivenko-Kanteli). *Neka je F funkcija raspodele obeležja X i F_n empirijska funkcija raspodele prostog slučajnog uzorka $\mathbf{X}$, obima n , iz populacije sa obeležjem X . Tada važi*

$$\|F - F_n\|_{\infty} = \sup_{x \in \mathbb{R}} |F(x) - F_n(x)| \xrightarrow[n \rightarrow \infty]{s.s.} 0. \quad (2.104)$$

Time smo dobili prvu i najjednostavniju test statistiku za testiranje hipoteze saglasnosti (2.101), datu sa

$$KS_n = \|F_n - F_0\|_{\infty}. \quad (2.105)$$

Prethodna test statistika naziva se **test statistika Kolmogorov-Smirnova**. Nultu hipotezu odbacujemo kada je KS_n veće od neke kritične tačke C_n . Neke od poznatijih test statistika zasnovanih na empirijskoj funkciji raspodele su i **Anderson-Darlingova test statistika**, kao i **Kramer-fon Mizesova test statistika**, ali o njima neće biti reči u ovom radu. Pomenute test statistike su varijacija Kolmogorov-Smirnovljeve i dobijaju se promenom metrike koja generiše normu.

Videli smo kako konstruisati KS test statistiku, kao i to da je takva konstrukcija dobra. Sada želimo da vidimo neke od osobina ovog testa. Pre svega, želimo da vidimo pojedine

teorijske rezultate, kao na primer, koja je granična raspodela test statistike i kojom brzinom test statistika konvergira ka njoj.

Lako se pokaže iz osobine Bernulijevih slučajnih veličina i centralne granične teoreme da, za svako $x \in \mathbb{R}$,

$$\sqrt{n}(F_n(x) - F(x)) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, F(x)(1 - F(x))). \quad (2.106)$$

Ovaj rezultat jeste važan, ali nam ne govori mnogo o test statistici KS_n . Razlog je isti kao i maločas. Naime, prethodnu konvergenciju vidimo tačka-po-tačka, što znači da je gledamo nezavisno za svaku tačku $x \in \mathbb{R}$ ponaosob. Da li možemo pojačati prethodni rezultat kao u slučaju Glivenko-Kantelijeve teoreme? Ne toliko jednostavno. Primitimo da želimo da nađemo graničnu raspodelu za

$$KS_n = \sup_{x \in \mathbb{R}} |F(x) - F_n(x)|.$$

Mi zapravo ne posmatramo supremum slučajne veličine, već slučajnog procesa

$$\{|F(x) - F_n(x)| : x \in \mathbb{R}\}. \quad (2.107)$$

Zbog toga moramo naći graničan slučajni proces, a ne graničnu slučajnu veličinu. Prvo što nam može pasti na pamet jeste Braunovo kretanje jer u njegovoj konstrukciji figuriše normalna raspodela, a možemo ga smatrati i neprekidnim (po x) uopštenjem normalne raspodele. Međutim, znamo da za Braunovo kretanje $\{W_t \mid t > 0\}$, u tački t , važi

$$W_t \sim \mathcal{N}(0, t),$$

što implicira da disperzija ovog procesa teži beskonačnosti. S druge strane, vidimo da za naš proces (2.107) važi

$$\mathbb{D}[F_n(x) - F(x)] = \mathbb{D}[F_n(x)] = \frac{F(x)(1 - F(x))}{n} \xrightarrow{x \rightarrow \infty} 0.$$

Dakle, Braunovo kretanje neće biti graničan proces, ali hoće modifikovano Braunovo kretanje, odnosno Braunov most, što nam daje sledeća teorema.

Teorema 2.3.5 (Donsker). *Neka je F **neprekidna** funkcija raspodele obeležja X i neka je F_n empirijska funkcija raspodele prostog slučajnog uzorka $\mathbf{X}$, obima n , iz populacije sa obeležjem X . Tada važi*

$$\sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow[n \rightarrow \infty]{d} \sup_{0 \leq t \leq 1} |B_t|, \quad (2.108)$$

gde je $\{B_t \mid 0 \leq t \leq 1\}$ Braunov most na $[0, 1]$.

Vidimo da smo nametnuli dodatni uslov o neprekidnosti funkcije raspodele F . Time smo automatski ograničili skup raspodela za koje možemo testirati hipotezu (2.101). Razlog tome je očigledan. Ako želimo da dobijemo neprekidan graničan proces gde neprekidnost gledamo po x , moramo imati slučajan proces koji je takođe neprekidan po x . Kako $F_n(x)$ očigledno nije neprekidno, to $F(x)$ mora biti. Ovo ne znači da ne možemo primeniti KS

test na funkcije koje nisu neprekidne, ali znači da se u tom slučaju ne možemo osloniti na matematički aparat razvijen za neprekidne funkcije i koristiti prethodnu teoremu.

Supremum Braunovog mosta nam ništa ne govori, pa ćemo navesti još jednu teoremu koja nam daje raspodelu tog supremuma [4].

Teorema 2.3.6 (Kolmogorova raspodela). *Neka je $\{B_t \mid 0 \leq t \leq 1\}$ Braunov most na $[0, 1]$. Tada je*

$$\mathbb{P}\left(\sup_{0 \leq t \leq 1} |B_t| \leq x\right) = 1 - 2 \sum_{n=1}^{\infty} (-1)^{n+1} e^{-2n^2 x^2} = L(x), \quad x > 0. \quad (2.109)$$

Vrednost granične raspodele L računa se numerički.

Pre nego što završimo poglavlje, podsetimo se još kako merimo da li je neki test dobar ili nije. Nakon toga ukratko ćemo se podsetiti Monte Karlo i butstrep (eng. *bootstrap*) metoda koje su neizostavne u svakodnevnim statističkim simulacijama.

2.3.12 Mera i moć test

Kada govorimo o tome koliko dobro test radi u odnosu na nultu hipotezu, možemo pomenuti kriterijum **mere testa**. Mera testa je maksimalna verovatnoća pravljenja greške **prve vrste**, odnosno verovatnoća odbacivanja nulte hipoteze kada je ona tačna. Dakle, za dobre testove očekujemo da mera bude bliska 0.05.

Sa druge strane, kada govorimo o tome koliko je neki test dobar protiv određenih alternativa, onda govorimo o **moći testa**. Moć statističkog testa je verovatnoća da test tačno odbaci nultu hipotezu kada je alternativna hipoteza tačna. Odnosno, moć testa je povezana sa greškom **druge vrste**.

2.3.13 Statističke metode simulacije

Dve, možda najvažnije, metode u statističkim simulacijama su Monte Karlov metod i metoda butstrepa.

Suštinski, Monte Karlov metod je ponavljanje iste simulacije više puta, gde u svakoj iteraciji ocenjujemo parametar ili računamo test statistiku ili bilo šta vezano za statističko zaključivanje. Kada to uradimo dovoljan broj puta dobijamo empirijsku raspodelu ocenjenih parametara, odnosno test statistika. Bez ovog metoda ne bismo mogli da donosimo nikakve empirijske statističke zaključke. Kako je nekad skup ili je nemoguće više puta generisati podatke iz istog modela, tada u pomoć pristiže butstrep. U suštini butstrep je samo poduzorkovanje sa ponavljanjem iz polaznog uzorka. Slično kao Monte Karlov metod, prvo generišemo podatke iz određenog modela. Posle toga poduzorkujemo sa ponavljanjem i ocenimo parametar ili test statistiku. Onda ponovo poduzorkujemo novi uzorak i tako dok ne dobijemo niz ocena parametara ili test statistika. Na osnovu tog niza zaključujemo o raspodeli pomenute ocene.

Ovim, najзад, završavamo poglavlje i počinjemo sa prvim problemom pri statističkom zaključivanju u stohastičkim difuzijama - diskretizacijom modela.

Diskretizacija stohastičkih diferencijalnih jednačina

U ovom poglavlju želimo da se fokusiramo na problem diskretizacije stohastičkih diferencijalnih jednačina kao važnom koraku ka statističkom zaključivanju. Primena statistike zahteva prilično dobre aproksimacije stohastičke diferencijalne jednačine. Iz tog razloga uvodimo koncept reda srednje-kvadratne konvergencije koji će nam pomoći da poredimo diskretizacije i vidimo zašto diskretizacija na osnovu Ojlerove metode neće biti dovoljno dobra iako je prosta za implementaciju. Takođe, statističke metode se često oslanjaju na generisanje i simuliranje trajektorija stohastičke diferencijalne jednačine. Prema tome, potrebna je dobra diskretizacija koja će ovo učiniti efikasnim i računarski povoljnim. Pored toga, jedna od glavnih osobina koju mora da poseduje diskretizacija, jeste da sačuva najvažnije osobine stohastičke difuzije kao modela i to za najveću moguću dužinu koraka diskretizacije h . Osobine koje želimo da sačuvamo diskretizacijom su na primer: hipoeleptičnost, geometrijska ergodičnost, ali i oscilatorna dinamika kao što su amplitude, frekvencije i faze oscilacije.

U ovom poglavlju predstavimo neke od najvažnijih rezultata iz ovogodišnjeg rada [2] na pomenutu temu. Rad koristi metode deljenja toka vektorskog polja na stohastičke diferencijalne jednačine određenog tipa i pokazuje superiornost ovih metoda u odnosu na metode bazirane na Ojlerovoj diskretizaciji.

Na početku ovog poglavlja uvešćemo pojmove kao što su red srednje-kvadratne konvergencije i još nekoliko metoda diskretizacije za stohastičke diferencijalne jednačine. Potom ćemo opisati šta je to hipoeleptičnost i zašto je važno očuvati je kroz diskretizaciju, a zatim ćemo uvesti klasu modela sa kojima ćemo raditi. Na kraju ćemo se fokusirati na rezultate iz rada o redu srednje-kvadratne konvergencije metoda deljenja, kao i o tome kako ove metode čuvaju osobine modela.

3.1 Red srednje-kvadratne konvergencije

Razmatrajmo diskretizovani vremenski interval $[0, T]$ sa ekvidistantnim tačkama dužine $h = t_i - t_{i-1}$, takav da je $0 = t_0 < t_1 < \dots < t_n = T$. Označićemo sa $\{\tilde{\mathbf{X}}(t_i)\}_{i=0,1,\dots,n}$

numeričko rešenje generičke d -dimenzionalne stohastičke diferencijalne jednačine čije je stvarno rešenje stohastička difuzija $\{\mathbf{X}(t) \mid t \in [0, T]\}$. Neka je $\tilde{\mathbf{X}}(t_0) = X_0$.

Kažemo da numeričko rešenje $\tilde{\mathbf{X}}$ srednje-kvadratno konvergira **redom konvergencije** $p > 0$ ka pravom rešenju ako postoji konstanta $C > 0$ takva da je

$$\max_{t_i \in [0, T]} \left(\mathbb{E} \left[\left\| \mathbf{X}(t_i) - \tilde{\mathbf{X}}(t_i) \right\|^2 \right] \right)^{1/2} \leq Ch^p. \quad (3.1)$$

Neka $\mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i)$ označava tačno rešenje stohastičke diferencijalne jednačine u trenutku t_i , ako je u trenutku t_{i-1} bilo $\mathbf{X}(t_{i-1}) = \mathbf{x}$. Slično, neka $\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}(t_i)$ označava numeričko rešenje jednog koraka unapred stohastičke diferencijalne jednačine u trenutku t_i , ako je u trenutku t_{i-1} bilo $\tilde{\mathbf{X}}(t_{i-1}) = \mathbf{x}$.

Kažemo da je numerička aproksimacija zasnovana na jednom koraku unapred, $\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}(t_i)$, **srednje-kvadratno postojana reda** $p > 0$, ako za sve t_i , $i = 0, 1, \dots, n$ i sve $\mathbf{x} \in \mathbb{R}^d$ važi

$$\left\| \mathbb{E} \left[\mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}(t_i) \right] \right\| = \mathcal{O}(h^{p+1}), \quad (3.2)$$

$$\left(\mathbb{E} \left[\left\| \mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}(t_i) \right\|^2 \right] \right)^{1/2} = \mathcal{O}(h^{p+\frac{1}{2}}). \quad (3.3)$$

Pored srednje-kvadratne postojanosti, ograničenost drugog momenta numeričkog rešenja mora biti pokazana. U slučaju kada je zadovoljena globalna Lipsčicovost, ovo je zagarantovano linearnim rastom. Međutim, mi želimo da generalizujemo model i da ne pretpostavljamo da vektor drifta zadovoljava globalno Lipsčicov uslov. Zato definišemo srednje-kvadratnu ograničenost.

Numeričko rešenje $\tilde{\mathbf{X}}(t_i)$ je **srednje-kvadratno ograničeno** ako postoji konstanta $\tilde{K}(T) > 0$ takva da

$$\max_{t_i \in [0, T]} \mathbb{E} \left[\left\| \tilde{\mathbf{X}}(t_i) \right\|^2 \right] \leq \tilde{K}(T) (1 + \mathbb{E} [\left\| \mathbf{X}_0 \right\|^2]). \quad (3.4)$$

Sledeća teorema nam daje karakterizaciju kada numeričko rešenje ima red srednje-kvadratne konvergencije p . Primetimo da je teorema vrlo slična onoj koju smo videli u diskretnom slučaju u poglavlju 2, s tim što sada nećemo pretpostaviti globalnu Lipsčicovost, već u određenoj meri oslabljene uslove uz srednje-kvadratnu ograničenost.

Teorema 3.1.1. *Neka je $\tilde{\mathbf{X}}(t_i)$ numeričko rešenje u trenutku t_i stohastičke diferencijalne jednačine sa početnim uslovom $\mathbf{X}_0$. Neka je ovo rešenje konstruisano metodom jednog koraka $\mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i)$. Pretpostavimo da je numeričko rešenje srednje-kvadratno postojan red $p > 0$ i da je srednje-kvadratno ograničeno. Ako za funkciju drifta stohastičke diferencijalne jednačine $\boldsymbol{\mu}$ važi*

- (1) *funkcija drifta $\boldsymbol{\mu}$ je globalno Lipsčic neprekidna sa jedne strane, tj. ako postoji konstanta $c_1 > 0$ takva da je*

$$\langle \mathbf{x} - \mathbf{y}, \boldsymbol{\mu}(\mathbf{x}) - \boldsymbol{\mu}(\mathbf{y}) \rangle \leq c_1 \|\mathbf{x} - \mathbf{y}\|^2, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d, \quad (3.5)$$

(2) funkcija drifta $\boldsymbol{\mu}$ raste najviše polinomijalno, tj. postoje konstante $c_2 > 0$ i $\xi \geq 1$ takve da

$$\|\boldsymbol{\mu}(\mathbf{x}) - \boldsymbol{\mu}(\mathbf{y})\|^2 \leq c_2 (1 + \|\mathbf{x}\|^{2\xi-2} + \|\mathbf{y}\|^{2\xi-2}) \|\mathbf{x} - \mathbf{y}\|^2, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d, \quad (3.6)$$

onda numeričko rešenje konvergira srednje-kvadratno sa redom konvergencije p .

Prethodna teorema će nam pomoći da pokažemo da je red srednje-kvadratne konvergencije za metode deljenja Li-Trotera i Stranga $p = 1$. U nastavku ćemo prikazati neke poznate metode diskretizacije stohastičkih diferencijalnih jednačina.

3.2 Metode diskretizacije

Pretpostavimo da želimo da diskretizujemo sledeću d -dimenzionalnu stohastičku diferencijalnu jednačinu

$$d\mathbf{X}_t = \boldsymbol{\mu}(\mathbf{X}_t)dt + \boldsymbol{\Sigma}(\mathbf{X}_t)d\mathbf{W}_t, \quad (3.7)$$

gde je $\mathbf{W}$ m -dimenzionalno Braunovo kretanje. Diskretizacija intervala $[0, T]$ i korak h su definisani kao i do sada.

3.2.1 Metode Ojler-Marujame

Kao i u diskretnom slučaju, počinjemo od najjednostavnije **metode Ojler-Marujame** koja je dobijena kao adaptacija Ojlerove metode u stohastičkom slučaju i glasi

$$\tilde{\mathbf{X}}^{\text{OM}}(t_i) = \tilde{\mathbf{X}}^{\text{OM}}(t_{i-1}) + h\boldsymbol{\mu}(\tilde{\mathbf{X}}^{\text{OM}}(t_{i-1})) + \sqrt{h}\boldsymbol{\Sigma}(\tilde{\mathbf{X}}^{\text{OM}}(t_{i-1}))\boldsymbol{\varepsilon}_{i-1}, \quad (3.8)$$

gde je $\boldsymbol{\varepsilon}_i \stackrel{i.i.d.}{\sim} \mathcal{N}_m(\mathbf{0}, \mathbf{I})$, za $i = 1, 2, \dots, n$.

Ako nametnemo globalnu Lipšicovu neprekidnost na funkciju drifta i difuzije stohastičke diferencijalne jednačine, pokazano je da Ojler-Marujamova metoda ima red srednje-kvadratne konvergencije $p = 1$, odnosno $p = 1/2$, kada je šum aditivan, odnosno multiplikativan, redom. Multiplikativan šum u ovom slučaju znači da intenzitet šuma, odnosno funkcija difuzije zavisi od trenutnog stanja procesa. Međutim, ako želimo da bar jedan od koeficijenata stohastičke diferencijalne jednačine raste brže od linearne funkcije, kao u slučaju FicHju-Nagumovog modela u kom imamo polinomijalni rast, onda ova metoda ne konvergira u srednje-kvadratnom smislu.

Kao što vidimo, ovaj metod jeste intuitivan i lak za implementaciju, ali ne radi već u slučaju nekih poznatih modela kao što je FicHju-Nagumov, a još uvek nije bilo ni reči o tome da li čuva geometrijska svojstva modela. Da bi se zaobišao ovaj nedostatak konvergencije, pribegava se modifikaciji ove metode. Ovde ćemo predstaviti neke od primera.

Prva varijanta dizajnirana je za polinomijalni rast i Lipšicovu neprekidnost sa jedne strane za funkciju drifta, odnosno globalnu Lipšicovost za funkciju difuzije, a zasniva se

na „ukroćivanju pertubacija” i izbegava velike vrednosti prouzrokovane polinomijalnim rastom drifta. Metoda je definisana na sledeći način

$$\tilde{\mathbf{X}}^{\text{UOM}}(t_i) = \tilde{\mathbf{X}}^{\text{UOM}}(t_{i-1}) + \frac{h\boldsymbol{\mu}(\tilde{\mathbf{X}}^{\text{UOM}}(t_{i-1}))}{1 + h\|\boldsymbol{\mu}(\tilde{\mathbf{X}}^{\text{UOM}}(t_{i-1}))\|} + \sqrt{h}\boldsymbol{\Sigma}(\tilde{\mathbf{X}}^{\text{UOM}}(t_{i-1}))\boldsymbol{\varepsilon}_{i-1}. \quad (3.9)$$

Nazvaćemo je **ukroćenom Ojler-Marujamovom metodom** (eng. *tamed Euler-Maruyama method*). Za nju se može pokazati da se poklapa sa Ojler-Marujamovom metodom do drugog stepena u Tejlorovom razvoju. Ova činjenica je poslužila da se pokaže da je red srednje-kvadratne konvergencije isti kao i u slučaju Ojler-Marujame, s tim što su uslovi za koeficijente stohastičke diferencijalne jednačine oslabljeni.

Sledeća metoda slabi pretpostavke i za funkciju difuzije, tj. ukroćuje i funkciju drifta i funkciju difuzije i dopušta da obe ove funkcije rastu polinomijalno i da su Lipšic neprekidne sa jedne strane. Metoda je definisana na sledeći način

$$\tilde{\mathbf{X}}^{\text{UDOM}}(t_i) = \tilde{\mathbf{X}}^{\text{UDOM}}(t_{i-1}) + \frac{h\boldsymbol{\mu}(\tilde{\mathbf{X}}^{\text{UDOM}}(t_{i-1})) + \sqrt{h}\boldsymbol{\Sigma}(\tilde{\mathbf{X}}^{\text{UDOM}}(t_{i-1}))\boldsymbol{\varepsilon}_{i-1}}{1 + h\|\boldsymbol{\mu}(\tilde{\mathbf{X}}^{\text{UDOM}}(t_{i-1}))\| + \sqrt{h}\|\boldsymbol{\Sigma}(\tilde{\mathbf{X}}^{\text{UDOM}}(t_{i-1}))\boldsymbol{\varepsilon}_{i-1}\|}. \quad (3.10)$$

Ovu metodu ćemo nazivati **Ojler-Marujamovom metodom ukroćene difuzije** (eng. *diffusion tamed Euler-Maruyama method*). Može se pokazati da ova metoda konvergira u srednje-kvadratnom smislu sa redom konvergencije $p = 1/2$.

Pored navedene dve, poznate su tzv. okrnjene varijante Ojler-Marujamove metode koje nemaju red srednje-kvadratne konvergencije i date su sa

$$\tilde{\mathbf{X}}^{\text{OOM}}(t_i) = \tilde{\mathbf{X}}^{\text{OOM}}(t_{i-1}) + \frac{h\boldsymbol{\mu}(\tilde{\mathbf{X}}^{\text{OOM}}(t_{i-1}))}{\max\left\{1, h\|\boldsymbol{\mu}(\tilde{\mathbf{X}}^{\text{OOM}}(t_{i-1}))\|\right\}} + \sqrt{h}\boldsymbol{\Sigma}(\tilde{\mathbf{X}}^{\text{UOM}}(t_{i-1}))\boldsymbol{\varepsilon}_{i-1}, \quad (3.11)$$

$$\tilde{\mathbf{X}}^{\text{ODOM}}(t_i) = \tilde{\mathbf{X}}^{\text{ODOM}}(t_{i-1}) + \frac{h\boldsymbol{\mu}(\tilde{\mathbf{X}}^{\text{ODOM}}(t_{i-1})) + \sqrt{h}\boldsymbol{\Sigma}(\tilde{\mathbf{X}}^{\text{ODOM}}(t_{i-1}))\boldsymbol{\varepsilon}_{i-1}}{\max\left\{1, h\|\boldsymbol{\mu}(\tilde{\mathbf{X}}^{\text{ODOM}}(t_{i-1}))\| + \sqrt{h}\|\boldsymbol{\Sigma}(\tilde{\mathbf{X}}^{\text{ODOM}}(t_{i-1}))\boldsymbol{\varepsilon}_{i-1}\|\right\}}. \quad (3.12)$$

Prethodne dve metode ćemo nazivati **okrnjenom Ojler-Marujamovom metodom** (eng. *truncated Euler-Maruyama method*) i **okrnjenom difuzijom Ojler-Marujamove metode** (eng. *diffusion truncated Euler-Maruyama method*), redom.

3.2.2 Metode višeg reda

Da bi se povećao red konvergencije, pribegava se daljem razvoju stohastičke verzije Tejlorove formule i dodavanju više članova u metodu. Nakon Ojler-Marujamove najjednostavnije metode sledi **metoda Majlštajna** koja dodaje član drugog reda iz Itoove formule. Međutim, formule metoda višeg reda su jako kompleksne i zahtevaju dosta članova i razvoj stohastičke verzije Tejlorove formule. Kako smo preskočili razvoj ove verzije Tejlorove teoreme, to ćemo preskočiti i višedimenzionalne metode višeg reda i

pokazati formule samo u slučaju da imamo jednodimenzionalnu stohastičku diferencijalnu jednačinu. Majlštajnova metoda u jednoj dimenziji je data sledećom formulom

$$\tilde{X}^M(t_i) = \tilde{X}_{i-1}^M + h\mu(\tilde{X}_{i-1}^M) + \sqrt{h}\sigma(\tilde{X}_{i-1}^M)\varepsilon_{i-1} + \frac{h}{2}\sigma(\tilde{X}_{i-1}^M)\sigma'(\tilde{X}_{i-1}^M)(\varepsilon_{i-1}^2 - 1). \quad (3.13)$$

Ova metoda ima red srednje-kvadratne konvergencije $p = 1$, bez pretpostavke o tipu šuma, ali u slučaju kad su koeficijenti stohastičke diferencijalne jednačine globalno Lipšic neprekidni. U suprotnom, mora se primeniti ukroćivanje kao i u slučaju Ojler-Marujamove metode. Primetimo još da kada je šum aditivan, tj. σ nije funkcija od x , Majlštajnova metoda se podudara sa metodom Ojler-Marujame.

Korak dalje, stižemo do **metode 1.5** predložene u [9]. Opet, ova metoda se zasniva na razvoju više elemenata u stohastičkoj verziji Tejlorovog razvoja. Ova metoda ima kompleksan zapis čak i u jednodimenzionalnom slučaju, a on glasi

$$\begin{aligned} \tilde{X}^{1.5}(t_i) = & \tilde{X}_{i-1}^{1.5} + h\mu(\tilde{X}_{i-1}^{1.5}) + \sqrt{h}\sigma(\tilde{X}_{i-1}^{1.5})\varepsilon_{i-1} + \frac{h}{2}\sigma(\tilde{X}_{i-1}^{1.5})\sigma'(\tilde{X}_{i-1}^{1.5})(\varepsilon_{i-1}^2 - 1) \\ & + \frac{h^{3/2}}{\sqrt{3}}\sigma(\tilde{X}_{i-1}^{1.5})\mu'(\tilde{X}_{i-1}^{1.5})\eta_{i-1} + \frac{h^2}{2}\left(\mu(\tilde{X}_{i-1}^{1.5})\mu'(\tilde{X}_{i-1}^{1.5}) + \frac{1}{2}\sigma^2(\tilde{X}_{i-1}^{1.5})\mu''(\tilde{X}_{i-1}^{1.5})\right) \\ & + h^{3/2}\left(\mu(\tilde{X}_{i-1}^{1.5})\sigma'(\tilde{X}_{i-1}^{1.5}) + \frac{1}{2}\sigma(\tilde{X}_{i-1}^{1.5})\sigma''(\tilde{X}_{i-1}^{1.5})\right)\left(\varepsilon_{i-1} - \frac{1}{\sqrt{3}}\eta_{i-1}\right) \\ & + \frac{h^{3/2}}{2}\sigma(\tilde{X}_{i-1}^{1.5})\left(\tilde{X}_{i-1}^{1.5}(\tilde{X}_{i-1}^{1.5})\sigma'(\tilde{X}_{i-1}^{1.5})\right)'\left(\frac{1}{3}\varepsilon_{i-1}^2 - 1\right)\varepsilon_{i-1}, \end{aligned} \quad (3.14)$$

gde su $\varepsilon_i, \eta_i \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$ međusobno zavisni nizovi normalnih slučajnih veličina, takvi da je

$$\mathbb{E}[\varepsilon_i \eta_i] = \frac{\sqrt{3}}{2}, \quad i = 1, 2, \dots, n. \quad (3.15)$$

Vidimo da metode višeg reda zasnovane na dodavanju ostalih članova u stohastičkoj verziji Tejlorovog razvoja nisu održive, tj. sa svakim novim članom formula postaje sve kompleksnija, a implementacija iste postaje isuviše zahatevna. Dok ove formule povećavaju red konvergencije pod jakom pretpostavkom globalne Lipšicovosti, dodatno se mora prilagoditi svaka od prethodne dve formule, kao u slučaju Ojler-Marujamove metode.

Pokazano je da sve pomenute metode zasnovane na Ojler-Marujamovoj, kao i Majlštajnova metoda, ne čuvaju geometrijska svojstva neprekidnog modela (kao što su hipoeliptičnost, simetričnost, amplitude i frekvencije kod oscilatornih modela, itd). U slučaju metode 1.5 u radu [3] pokazano je da ova metoda čuva hipoeliptičnost. Na primeru ćemo empirijski pokazati da ova metoda čuva i geometrijska svojstva kao što su amplituda i frekvencija kod oscilatornih modela. Sa druge strane, sama implementacija metode je čini nepraktičnom za upotrebu u opštem slučaju.

U nastavku poglavlja ćemo videti zašto su metode deljenja superiornije od prethodno pomenutih iako imaju red srednje-kvadratne konvergencije $p = 1$. Pokazaćemo da čuvaju hipoeliptičnost i da pogodno izabrana Li-Troterova metoda dovodi do Gausove verovatnoće prelaska, što će omogućiti ocenu parametara i testiranje hipoteza.

Pre nego što uvedemo metode deljenja sa kojima ćemo raditi do kraja ovog rada, prvo ćemo prodiskutovati još jednu osobinu modela slučajnih difuzija - hipoeliptičnost.

3.3 Struktura šuma: eliptičnost i hipoeliptičnost

Pretpostavimo da u modelu imamo aditivan šum, tj. matrica difuzije ne zavisi od promenljivih u modelu. U zavisnosti od strukture šuma, razlikujemo dva tipa modela.

U prvi tip spadaju **eliptične** stohastičke diferencijalne jednačine koje karakteriše nesingularna matrica difuzije, odnosno matrica $\Sigma\Sigma^\top$ ima potpun rang, tj. $\det(\Sigma\Sigma^\top) \neq 0$. Dakle, modeli u kojima je $m = d$ sa dijagonalnom matricom difuzije $\Sigma = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_d]$, gde je $\sigma_i > 0$, za $i = 1, 2, \dots, d$, jesu eliptični modeli.

Drugi tip modela čine modeli koji imaju singularnu matricu difuzije. Razmotrimo slučaj kada je $m = d - 1$ i kada je Σ data sa

$$\Sigma = \begin{bmatrix} \mathbf{0}_{d-1}^\top \\ \mathbf{\Gamma} \end{bmatrix}, \quad (3.16)$$

gde je $\mathbf{\Gamma} = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_{d-1}] \in \mathbb{R}^{d-1} \times \mathbb{R}^{d-1}$, takva da je $\sigma_i > 0$, za $i = 1, 2, \dots, d - 1$. Tada se prva kolona rešenja $\mathbf{X}(t)$ naziva **glatkom** (eng. *smooth*) jer šum ne utiče direktno na nju. Za ostale kolone kažemo da su **hrapave** (eng. *rough*) jer šum utiče direktno na njih. U opisanom slučaju dati model je često **hipoeliptičan**. To znači da verovatnoća prelaska ima glatku funkciju gustine čak i kada $\Sigma\Sigma^\top$ nema potpun rang.

Potreban i dovoljan uslov da proces $\mathbf{X}(t)$ bude hipoeliptičan, jeste da se bar jedna od hrapavih koordinata pojavljuje u glatkoj komponenti funkcije drifta. Konkretno, u prethodno opisanom primeru mora da važi sledeći uslov

$$(\forall \mathbf{x} \in \mathbb{R}^d)(\exists j = 1, 2, \dots, d - 1) \langle \nabla_r \mu_1(\mathbf{x}), \Sigma^j \rangle \neq 0, \quad (3.17)$$

gde je Σ^j j -ta kolona matrice $\mathbf{\Gamma}$ u opštem slučaju, a

$$\nabla_r \mu_1(\mathbf{x}) = \begin{bmatrix} \frac{\partial \mu_1(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial \mu_1(\mathbf{x})}{\partial x_d} \end{bmatrix} \quad (3.18)$$

gradijent sa parcijalnim izvodima prve koordinate vektora difuzije μ_1 , po svim hrapavim koordinatama.

Prethodno razmatranje se može uopštiti na više glatkih koordinata tako što za svaku glatku koordinatu mora da postoji bar jedna hrapava koja se propagira kroz funkciju drifta glatke koordinate.

Često u realnom svetu ne želimo da dodajemo šum na sve koordinate modela, već želimo da šum bude propagiran na ceo sistem kroz određene koordinate koje će biti zadužene, između ostalog, za modelovanje šuma. Zbog toga nam je važno da numerička

diskretizacija može da sačuva ovu osobinu. Primer eliptične stohastičke diferencijalne jednačine bi bio

$$\begin{cases} dX_t &= f(X_t, Y_t)dt + \sigma_1 dW_t, \\ dY_t &= g(X_t, Y_t)dt + \sigma_2 dW_t, \end{cases} \quad (3.19)$$

gde su $\sigma_1, \sigma_2 > 0$. Primer hipoeliptične stohastičke diferencijalne jednačine, bio bi za npr $\sigma_1 = 0$. Tada bi $w \times 2$ matrica difuzije imala rang 1, ali bi se šum propagirao kroz funkciju drifta f preko koordinate Y . Primer stohastičke diferencijalne jednačine koja nije ni eliptična ni hipoeliptična je

$$\begin{cases} dX_t &= f(X_t, Y_t)dt \\ dY_t &= g(X_t, Y_t, Z_t)dt, \\ dZ_t &= h(X_t, Y_t, Z_t)dt + \sigma dW_t. \end{cases} \quad (3.20)$$

Tada se šum propagira kroz Z do Y koje je u ovom slučaju glatka koordinata, ali se ne propagira do X . Tako rang nije potpun, ali ni šum ne dolazi do svih koordinata.

U nastavku poglavlja fokusiraćemo se na model, pretpostavke modela i diskretizaciju metodom deljenja, kao i osobine te diskretizacije.

3.4 Model i osobine

Posmatrajmo sledeću d -dimenzionalnu stohastičku diferencijalnu jednačinu

$$\begin{aligned} d\mathbf{X}_t &= \boldsymbol{\mu}(\mathbf{X}_t)dt + \boldsymbol{\Sigma}d\mathbf{W}_t \\ &= \mathbf{A}\mathbf{X}(t) + \mathbf{N}(\mathbf{X}(t)) + \boldsymbol{\Sigma}d\mathbf{W}_t, \quad \mathbf{X}(0) = \mathbf{X}_0, \quad t \in [0, T], \end{aligned} \quad (3.21)$$

gde je $\mathbf{W}$ m -dimenzionalno Braunovo kretanje, $T > 0$, $\mathbf{A} \in \mathbb{R}^{d \times d}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times m}$, a vektorske funkcije $\boldsymbol{\mu}, \mathbf{N} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ lokalno Lipšic-neprekidne. Inicijalna vrednost $\mathbf{X}_0$ je $\mathcal{F}_0$ merljiva $\mathbb{R}^d$ slučajna veličina koja je nezavisna od Braunovog kretanja $\mathbf{W}$ i takva da je $\mathbb{E}[\|\mathbf{X}_0\|^2] < \infty$. Pretpostavićemo da prethodna stohastička diferencijalna jednačina ima jedinstveno jako rešenje pod uslovima datim u knjizi [8], tj. da je rešenje definisano na celom intervalu $[0, T]$ i da njegove trajektorije ne izdivergiraju u beskonačnost u konačnom vremenu. Formalno, zahtevamo postojanje slučajne difuzije $\mathbf{X}(t)$ adaptirane na $\mathcal{F}(t)$ koja ima neprekidne trajektorije koje zadovoljavaju

$$\mathbf{X}(t) = \mathbf{X}_0 + \int_0^t \boldsymbol{\mu}(\mathbf{X}(s))ds + \int_0^t \boldsymbol{\Sigma}d\mathbf{W}(s), \quad t \in [0, T]. \quad (3.22)$$

Sada ćemo uvesti neke pretpostavke modela koje će biti neophodne da bismo dokazali različite osobine.

Pretpostavka 1 (P1). *Funkcija $\mathbf{N}$ je globalno Lipšic neprekidna sa jedne strane, tj. postoji konstanta $c_1 > 0$ takva da je*

$$\langle \mathbf{x} - \mathbf{y}, \mathbf{N}(\mathbf{x}) - \mathbf{N}(\mathbf{y}) \rangle \leq c_1 \|\mathbf{x} - \mathbf{y}\|^2, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d. \quad (3.23)$$

Pretpostavka 2 (P2). Funkcija $\mathbf{N}$ raste najviše polinomijalno, tj. postoje konstante $c_2 > 0$ i $\xi \geq 1$ takve da

$$\|\mathbf{N}(\mathbf{x}) - \mathbf{N}(\mathbf{y})\|^2 \leq c_2 (1 + \|\mathbf{x}\|^{2\xi-2} + \|\mathbf{y}\|^{2\xi-2}) \|\mathbf{x} - \mathbf{y}\|^2, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d. \quad (3.24)$$

Pretpostavka (P1) služi da oslabi uslov o globalnoj Lipšicovosti, dok pretpostavka (P2) služi da osigura ograničenost momenata stohastičke difuzije $\mathbf{X}(t)$. Konkretno, postoji konstanta $K(T) > 0$ takva da je

$$\max_{t \in [0, T]} \mathbb{E} [\|\mathbf{X}(t)\|^2] \leq K(T) (1 + \mathbb{E} [\|\mathbf{X}_0\|^2]). \quad (3.25)$$

Takođe, slučajna difuzija $\{\mathbf{X} \mid t \in [0, T]\}$ je Markovljev proces. Označimo sa $\mathcal{B}(\mathbb{R}^d)$ Borelovu σ -algebru na $\mathbb{R}^d$. Tada definišemo verovatnoću prelaska difuzije $\mathbf{X}$ sa

$$P_t(\mathcal{A}, \mathbf{x}) = \mathbb{P}(\mathbf{X}(t) \in \mathcal{A} \mid \mathbf{X}(0) = \mathbf{x}), \quad (3.26)$$

gde je $\mathcal{A} \in \mathcal{B}(\mathbb{R}^d)$. Prethodna definicija opisuje verovatnoću da je proces dostigao Borelov skup $\mathcal{A} \subset \mathbb{R}^d$ u trenutku $t > 0$ ako je krenuo iz pozicije $\mathbf{x} \in \mathbb{R}^d$ u trenutku 0.

3.4.1 Metode deljenja za stohastičku diferencijalnu jednačinu

Ovde ćemo predložiti strategiju za deljenje modela sa prethodnim karakteristikama. Podelimo stohastičku diferencijalnu jednačinu (3.21) na sledeće dve

$$d\mathbf{X}^{[1]}(t) = \mathbf{A}\mathbf{X}^{[1]}(t)dt + \Sigma d\mathbf{W}(t), \quad \mathbf{X}^{[1]}(0) = \mathbf{X}_0^{[1]}, \quad t \in [0, T], \quad (3.27)$$

$$d\mathbf{X}^{[2]}(t) = \mathbf{N}(\mathbf{X}^{[2]}(t))dt, \quad \mathbf{X}^{[2]}(0) = \mathbf{X}_0^{[2]}, \quad t \in [0, T]. \quad (3.28)$$

Možemo primetiti da je prva od dve stohastičke diferencijalne jednačine rešiva uvek i da je njeno rešenje Ornštajn-Ulenbekov proces dat sa

$$\mathbf{X}^{[1]}(t) = e^{\mathbf{A}t}\mathbf{X}_0^{[1]} + \int_0^t e^{\mathbf{A}(t-s)}\Sigma d\mathbf{W}(s). \quad (3.29)$$

Već smo videli da prethodna slučajna veličina ima normalnu raspodelu jer Itoov integral ima normalnu raspodelu sa očekivanjem $\mathbf{0}$ i matricom kovarijacije

$$\mathbf{C}(t) = \int_0^t e^{\mathbf{A}(t-s)}\Sigma\Sigma^\top e^{\mathbf{A}^\top(t-s)}ds. \quad (3.30)$$

Dakle, rešenje stohastičke diferencijalne jednačine (3.27) može biti diskretizovano na sledeći način

$$\Phi_h^{[1]}(\mathbf{X}^{[1]}(t_{i-1})) = \mathbf{X}^{[1]}(t_i) = e^{\mathbf{A}h}\mathbf{X}^{[1]}(t_{i-1}) + \boldsymbol{\varepsilon}_{i-1}, \quad i = 1, 2, \dots, n, \quad (3.31)$$

gde je $\boldsymbol{\varepsilon}_i \stackrel{i.i.d}{\sim} \mathcal{N}_d(\mathbf{0}, \mathbf{C}(h))$.

Što se tiče obične diferencijalne jednačine (3.28), moramo napraviti još jednu jaku pretpostavku.

Pretpostavka 3 (P3). Funkcija $\mathbf{N} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ je takva da diferencijalna jednačina (3.28) ima analitičko rešenje.

Pod prethodnom pretpostavkom možemo diskretizovati i običnu diferencijalnu jednačinu i dobiti

$$\Phi_h^{[2]}(\mathbf{X}^{[2]}(t_{i-1})) = \mathbf{X}^{[2]}(t_i) = \mathbf{f}_h(\mathbf{X}^{[2]}(t_{i-1})), \quad i = 1, 2, \dots, n, \quad (3.32)$$

gde je $\mathbf{f}_h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ eksplicitno rešenje obične diferencijalne jednačine (3.28). Pretpostavka (3) se može oslabiti tako da umesto eksplicitnog rešenja koristimo neku numeričku metodu za dobijanje aproksimacije ovog rešenja, ali to bi zakomplikovalo dokaze koji slede. Stoga, radi jednostavnosti pretpostavljamo da je model dat sa pretpostavkom (P3).

Da bismo konačno dobili numeričko rešenje polazne stohastičke diferencijalne jednačine, koristimo metode Li-Trotera i Stranga na osnovu čega dobijamo

$$\tilde{\mathbf{X}}^{[LT]}(t_i) = \left(\Phi_h^{[1]} \circ \Phi_h^{[2]} \right) \left(\tilde{\mathbf{X}}^{[LT]}(t_{i-1}) \right) = e^{\mathbf{A}h} \mathbf{f}_h \left(\tilde{\mathbf{X}}^{[LT]}(t_{i-1}) \right) + \boldsymbol{\varepsilon}_{i-1}, \quad (3.33)$$

$$\tilde{\mathbf{X}}^{[S]}(t_i) = \left(\Phi_{h/2}^{[2]} \circ \Phi_h^{[1]} \circ \Phi_{h/2}^{[2]} \right) \left(\tilde{\mathbf{X}}^{[S]}(t_{i-1}) \right) = \mathbf{f}_{h/2} \left(e^{\mathbf{A}h} \mathbf{f}_{h/2} \left(\tilde{\mathbf{X}}^{[S]}(t_{i-1}) \right) + \boldsymbol{\varepsilon}_{i-1} \right). \quad (3.34)$$

Možemo još naglasiti da su prethodna dva rešenja vrlo jednostavna za implementaciju jer se eksponent matrice $e^{\mathbf{A}h}$ i matrica kovarijacije $\mathbf{C}(h)$ moraju samo jednom izračunati.

3.4.2 Srednje-kvadratna konvergencija za metode deljenja

U narednom delu ćemo iskoristiti teoremu 3.1.1 i pokazati da Li-Troterova i Strangova metoda konvergiraju u srednje-kvadratnom smislu sa redom konvergencije $p = 1$, čak i u slučaju polinomijalnog rasta i globalne Lipšicovosti sa jedne strane. Da bismo iskoristili tu teoremu, potrebno je da dokažemo postojanost ovih metoda reda $p = 1$ i da pokažemo da su one srednje-kvadratno ograničene. Navedeno dokazujemo kroz dve leme. Date leme i dokazi su opisani u radu [2].

Lema 3.4.1 (Srednje-kvadratna postojanost). *Neka važe pretpostavke (P1), (P2) i (P3) i neka su $\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[LT]}(t_i)$ i $\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[S]}(t_i)$ Li-Troterova i Strangova metoda jednog koraka unapred definisana kao u (3.33) i (3.34), redom. Tada su $\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[LT]}(t_i)$ i $\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[S]}(t_i)$ srednje-kvadratno postojane sa redom postojanosti $p = 1$.*

Dokaz. Označimo sa $\|\cdot\|_{L^2} = \sqrt{\mathbb{E}[\|\cdot\|^2]}$ L^2 -normu. Prvo ćemo pokazati postojanost za Li-Troterovu metodu. Koristimo na početku nejednakost trougla da bismo dobili

$$\left\| \mathbb{E} \left[\mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[LT]}(t_i) \right] \right\| \leq \left\| \mathbb{E} \left[\mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[OM]}(t_i) \right] \right\| + \left\| \mathbb{E} \left[\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[OM]}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[LT]}(t_i) \right] \right\|, \quad (3.35)$$

i

$$\left\| \mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[LT]}(t_i) \right\|_{L^2} \leq \left\| \mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[OM]}(t_i) \right\|_{L^2} + \left\| \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[OM]}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[LT]}(t_i) \right\|_{L^2}, \quad (3.36)$$

gde je $\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{OM}]}(t_i)$ Ojler-Marujamovo numeričko rešenje jednog koraka unapred. S obzirom na to da smo pretpostavili aditivnost slučajne greške, iskoristićemo činjenicu da je Ojler-Marujamova metoda srednje-kvadratno konzistentna sa redom $p = 1$, tj.

$$\left\| \mathbb{E} \left[\mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{OM}]}(t_i) \right] \right\| = \mathcal{O}(h^2), \quad (3.37)$$

$$\left\| \mathbf{X}_{t_{i-1}, \mathbf{x}}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{OM}]}(t_i) \right\|_{L^2} = \mathcal{O}(h^{\frac{3}{2}}). \quad (3.38)$$

Ostaje još da se ograniči razlika između Ojler-Marujamove metode i metoda deljenja. Koristeći stohastičku verziju Tejlorovog razvoja na stohastičku diferencijalnu jednačinu (3.27) dobijamo

$$\Phi_h^{[1]}(\mathbf{x}) = \mathbf{x} + \mathbf{A}\mathbf{x}h + \Sigma\sqrt{h}\boldsymbol{\psi}_{i-1} + \mathbf{r}, \quad (3.39)$$

gde $\boldsymbol{\psi}_{i-1} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{I})$ i $\mathbf{r}$ vektor ostatka za čiju svaku komponentu važi $\mathbb{E}[r_j] = \mathcal{O}(h^2)$ i $\mathbb{E}[r_j^2] = \mathcal{O}(h^3)$. Slično, koristeći determinističku verziju Tejlorovog razvoja na (3.32) dobijamo

$$\Phi_h^{[2]}(\mathbf{x}) = \mathbf{x} + \mathbf{N}(\mathbf{x})h + \mathbf{q}, \quad (3.40)$$

gde za svaku komponentu vektora $\mathbf{q}$ važi $q_j = \mathcal{O}(h^2)$. Tada, Li-Troterova aproksimacija jednog koraka ima sledeći razvoj

$$\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{LT}]}(t_i) = \left(\Phi_h^{[1]} \circ \Phi_h^{[2]} \right) (\mathbf{x}) = \Phi_h^{[1]} (\mathbf{x} + \mathbf{N}(\mathbf{x})h + \mathbf{q}) \quad (3.41)$$

$$= \mathbf{x} + \mathbf{N}(\mathbf{x})h + \mathbf{q} + \mathbf{A}\mathbf{x}h + \mathbf{A}\mathbf{N}(\mathbf{x})h^2 + \mathbf{A}\mathbf{q}h + \Sigma\sqrt{h}\boldsymbol{\psi}_{i-1} + \mathbf{r}. \quad (3.42)$$

Kako je Ojler-Marujamov metod dat sledećom jednačinom

$$\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{OM}]}(t_i) = \mathbf{x} + \mathbf{A}\mathbf{x}h + \mathbf{N}(\mathbf{x})h + \Sigma\sqrt{h}\boldsymbol{\psi}_{i-1}, \quad (3.43)$$

to je razlika između ova dva metoda

$$\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{LT}]}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{OM}]}(t_i) = \mathbf{q} + \mathbf{A}\mathbf{N}(\mathbf{x})h^2 + \mathbf{A}\mathbf{q}h + \mathbf{r}, \quad (3.44)$$

odakle vidimo da je

$$\left\| \mathbb{E} \left[\mathbf{X}_{t_{i-1}, \mathbf{x}}^{[\text{LT}]}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{OM}]}(t_i) \right] \right\| = \mathcal{O}(h^2), \quad (3.45)$$

$$\left\| \mathbf{X}_{t_{i-1}, \mathbf{x}}^{[\text{LT}]}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{OM}]}(t_i) \right\|_{L^2} = \mathcal{O}(h^{\frac{3}{2}}), \quad (3.46)$$

čime je dokaz gotov za Li-Troterov metod deljenja. Dokaz za Strangov metod je sličan, ali zahteva nešto obimniji račun. Počinjemo sa definicijom Strangove aproksimacije jednog koraka

$$\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[\text{S}]}(t_i) = \left(\Phi_{h/2}^{[2]} \circ \Phi_h^{[1]} \circ \Phi_{h/2}^{[2]} \right) (\mathbf{x}) \quad (3.47)$$

$$= \left(\Phi_{h/2}^{[2]} \circ \Phi_h^{[1]} \right) \left(\mathbf{x} + \mathbf{N}(\mathbf{x})\frac{h}{2} + \mathbf{q} \right) \quad (3.48)$$

$$= \Phi_{h/2}^{[2]} \left(\mathbf{x} + \mathbf{N}(\mathbf{x})\frac{h}{2} + \mathbf{q} + \mathbf{A}\mathbf{x}h + \mathbf{A}\mathbf{N}(\mathbf{x})h^2 + \mathbf{A}\mathbf{q}h + \Sigma\sqrt{h}\boldsymbol{\psi}_{i-1} + \mathbf{r} \right). \quad (3.49)$$

Radi jednostavnosti zapisa, uvedimo sledeću oznaku

$$\boldsymbol{\omega} = \mathbf{N}(\mathbf{x})\frac{h}{2} + \mathbf{q} + \mathbf{A}\mathbf{x}h + \mathbf{A}\mathbf{N}(\mathbf{x})h^2 + \mathbf{A}\mathbf{q}h + \boldsymbol{\Sigma}\sqrt{h}\boldsymbol{\psi}_{i-1} + \mathbf{r}. \quad (3.50)$$

Tada je $\mathbb{E}[\omega_j] = \mathcal{O}(h)$, kao i $\mathbb{E}[\omega_j^2] = \mathcal{O}(h)$, za $j = 1, 2, \dots, d$. Onda je

$$\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[S]}(t_i) = \Phi_{h/2}^{[2]}(\mathbf{x} + \boldsymbol{\omega}) \quad (3.51)$$

$$= \mathbf{x} + \boldsymbol{\omega} + \mathbf{N}(\mathbf{x} + \boldsymbol{\omega})\frac{h}{2} + \mathbf{q} \quad (3.52)$$

$$= \mathbf{x} + \boldsymbol{\omega} + \mathbf{N}(\mathbf{x})\frac{h}{2} + \boldsymbol{\omega}^*\frac{h}{2} + \mathbf{q}, \quad (3.53)$$

gde je $\boldsymbol{\omega}^*$ dobijena Tejlorovim razvojem oko $\mathbf{x}$ i ima isti red kao i $\boldsymbol{\omega}$, tj. $\mathbb{E}[\omega_j^*] = \mathcal{O}(h)$, kao i $\mathbb{E}[(\omega_j^*)^2] = \mathcal{O}(h)$, za $j = 1, 2, \dots, d$. Tada je

$$\tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[S]}(t_i) - \tilde{\mathbf{X}}_{t_{i-1}, \mathbf{x}}^{[OM]}(t_i) = 2\mathbf{q} + \mathbf{A}\mathbf{N}(\mathbf{x})\frac{h^2}{2} + \mathbf{A}\mathbf{q}h + \mathbf{r} + \boldsymbol{\omega}^*\frac{h}{2}, \quad (3.54)$$

odakle sledi dokaz tvrđenja u slučaju Strangovog metoda. ■

Što se tiče srednje-kvadratnog ograničenja, ovde neće biti izložen dokaz za metode deljenja. Ceo dokaz i neke dodatne pretpostavke mogu se pronaći u radu [2]. Umesto toga, pokušaćemo intuitivno da argumentujemo zašto očekujemo da ovako nešto važi. Prvo, imamo lokalno Lipšicovu običnu diferencijalnu jednačinu 3.27 koja ima globalno eksplicitno rešenje na intervalu $[0, T]$. To znači da se u konačnom vremenu ne dešava nikakva nepredviđena divergencija putanje. Prema tome, kada komponujemo ovo rešenje sa Ornštajn-Ulenbekovim procesom i metodama Li-Trotera i Stranga, ne dodajemo nikakve eksplozije ni divergencije u momentima u konačnom vremenu. Dokaz ovog tvrđenja nije trivijalan i zahteva dodatne pretpostavke o vektorskom polju. Ovde ćemo, svakako, izložiti dodatnu pretpostavku i lemu kako bismo zaokružili priču o srednje-kvadratnoj konvergenciji metoda deljenja.

Pretpostavka 4 (P4). *Za funkciju $\mathbf{f}_h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ definisanu sa (3.32) postoji konstanta $c_4 \geq 0$ takva da za sve $\mathbf{x} \in \mathbb{R}^d$ važi*

$$\|\mathbf{f}_h(\mathbf{x})\|^2 \leq \|\mathbf{x}\|^2 + c_4 h. \quad (3.55)$$

Lema 3.4.2 (Srednje-kvadratno ograničenje). *Ako važe pretpostavke (P3) i (P4) onda su Li-Troterova aproksimacija $\tilde{\mathbf{X}}^{[LT]}(t_i)$ i Strangova aproksimacija $\tilde{\mathbf{X}}^{[S]}(t_i)$ srednje-kvadratno ograničene.*

Na osnovu svega prethodno navedenog imamo sledeću teoremu.

Teorema 3.4.3 (Srednje-kvadratna konvergencija). *Neka su $\tilde{\mathbf{X}}^{[LT]}(t_i)$ i $\tilde{\mathbf{X}}^{[S]}(t_i)$ Li-Troterova i Strangova aproksimacija rešenja stohastičke diferencijalne jednačine (3.21), redom. Pod uslovima teoreme 3.1.1 i lema 3.4.1 i 3.4.2, ove dve aproksimacije konvergiraju u srednjem ka pravom rešenju sa redom konvergencije $p = 1$.*

3.4.3 Očuvanje hipoeliptičnosti

Da bismo dobili analogiju verovatnoće prelaska u diskretnom vremenu, definišemo verovatnoću prelaska u k -koraka numeričkog rešenja $\tilde{\mathbf{X}}(t_i)$ stohastičke diferencijalne jednačine (3.21) kao

$$\tilde{P}_k(\mathcal{A}, \mathbf{x}) = \mathbb{P}(\tilde{\mathbf{X}}(t_k) \in \mathcal{A} \mid \mathbf{X}(0) = \mathbf{x}), \quad (3.56)$$

gde $\mathcal{A} \in \mathcal{B}(\mathbb{R}^d)$ i $\mathbf{x} \in \mathbb{R}^d$. Pretpostavimo sada da je stohastička diferencijalna jednačina (3.21) hipoeliptična. Tada je njena verovatnoća prelaska (3.26) glatka bez obzira na to što rang matrice $\Sigma \Sigma^\top$ nije potpun. Sledeća definicija nam daje analogiju hipoeliptičnosti za numerička rešenja.

Definicija 3.4.1. *Neka je $\tilde{\mathbf{X}}(t_i)$ numeričko rešenje stohastičke diferencijalne jednačine (3.21) i neka je $k \in \mathbb{N}$ najmanji prirodan broj za koji verovatnoća prelaska (3.56) ima glatku gustinu. Tada za numeričko rešenje $\tilde{\mathbf{X}}(t_i)$ kažemo da je hipoeliptično u k -koraka.*

Prethodna definicija nam govori o tome kako numeričke aproksimacije prenose šum na glatke komponente posle k koraka iteracije. Kada je reč o simuliranju trajektorija stohastičke diferencijalne jednačine, ova osobina nije previše važna dokle god je vremenski okvir dovoljno velik i uvek će postojati neko k za koje su standardne metode diskretizacije hipoeliptične. Na primer, Ojler-Marujamove metode su hipoeliptične u 2-koraka. Međutim, slučaj za $k = 1$ je ključan za statističke zaključke. U ovom slučaju često uvodimo novu oznaku za verovatnoću prelaska i pišemo

$$\tilde{P}_h(\mathcal{A}, \mathbf{x}) = \mathbb{P}(\tilde{\mathbf{X}}(t_i) \in \mathcal{A} \mid \tilde{\mathbf{X}}(t_{i-1}) = \mathbf{x}). \quad (3.57)$$

Takođe, kada želimo da ocenimo parametre metodom maksimalne verodostojnosti u slučaju višedimenzionalne raspodele, uočavamo sledeći problem. Uslovna raspodela između svaka dva koraka numeričke aproksimacije rešenja normalna je sa određenim vektorom sredine i kovarijacionom matricom, tj.

$$\tilde{\mathbf{X}}(t_i) \mid \tilde{\mathbf{X}}(t_{i-1}) \sim \mathcal{N}_d \left(\mathbf{m}_i(\tilde{\mathbf{X}}(t_{i-1}), h; \theta), \mathbf{Q}(h; \theta) \right), \quad (3.58)$$

gde je θ skup parametara modela o kojima želimo da donesemo određene statističke zaključke na osnovu diskretizovanog uzorka difuzije $\mathbf{X}(t)$. Pod pretpostavkom da su podaci $\{\tilde{\mathbf{x}}(t_i) \mid i = 0, 1, \dots, n\}$ dostupni, možemo oceniti parametre θ tako što rešimo sledeći optimizacioni problem

$$\operatorname{argmin}_{\theta} \left\{ n \log \det \mathbf{Q}(h; \theta) + \sum_{i=1}^n (\tilde{\mathbf{x}}(t_i) - \mathbf{m}_i(\tilde{\mathbf{x}}(t_{i-1}), h; \theta))^\top \mathbf{Q}(h; \theta)^{-1} (\tilde{\mathbf{x}}(t_i) - \mathbf{m}_i(\tilde{\mathbf{x}}(t_{i-1}), h; \theta)) \right\}. \quad (3.59)$$

Ovaj način ocenjivanja parametara validan je samo ako je inverz $\mathbf{Q}(h; \theta)^{-1}$ dobro definisan, odnosno ako je numerički metod $\tilde{\mathbf{X}}(t_i)$ hipoeliptičan u 1-koraku. Za Ojler-Marujamove metode ovo ne važi, te u njihovom slučaju dobijamo singularne matrice kovarijacije date sa

$$\operatorname{cov}(\tilde{\mathbf{X}}^{[\text{OM}]}(t_i) \mid \tilde{\mathbf{X}}^{[\text{OM}]}(t_{i-1})) = h \Sigma \Sigma^\top. \quad (3.60)$$

Dakle, Ojler-Marujamov metod i sve njegove adaptacije nemaju gustinu prelaska u slučaju hipoeliptičnih difuzija, pa su samim tim nepraktične za korišćenje prilikom statističkog

zaključivanja. Ovaj problem se može rešiti metodama diskretizacije višeg reda, počevši od metode 1.5, ali kao što smo već videli, ovakav pristup nije uopšte lak za implementaciju, a još teže ga je generalizovati u više dimenzija.

Rešenje oba prethodna problema pružaju metodi zasnovani na principima deljenja vektorskog polja, kao što su Li-Troterov ili Strangov metod. Već smo se uverili da su pomenuti jednostavni za implementaciju, a sada ćemo prodiskutovati zašto čuvaju hipoeliptičnost. Oba pomenuta metoda imaju normalan slučajni vektor ξ_i koji ima matricu kovarijacije $\mathbf{C}(h)$ definisanu u jednačini (3.30). S obzirom na to da i matrica drifta $\mathbf{A}$ i matrica difuzije Σ ulaze u $\mathbf{C}(h)$, to su metodi deljenja hipoeliptični u 1-koraku, pod uslovom da je Ornštajn-Ulenbekov proces (3.27) hipoeliptičan.

Dok Strangov metod nema eksplicitnu raspodelu prelaska, Li-Troterov metod nam daje nedegenerisanu normalnu raspodelu.

Teorema 3.4.4. *Neka je model (3.21) hipoeliptičan. Pod pretpostavkom (P3), Li-Troterova aproksimacija $\tilde{\mathbf{X}}(t_i)$ (3.33) je hipoeliptična u 1-koraku i uslovna raspodela između svaka dva koraka je normalna sa sledećim komponentama*

$$\tilde{\mathbf{X}}^{[LT]}(t_i) \mid \tilde{\mathbf{X}}^{[LT]}(t_{i-1}) \sim \mathcal{N}_d \left(e^{\mathbf{A}h} \mathbf{f}_h(\tilde{\mathbf{X}}^{[LT]}(t_{i-1})), \mathbf{C}(h) \right), \quad (3.61)$$

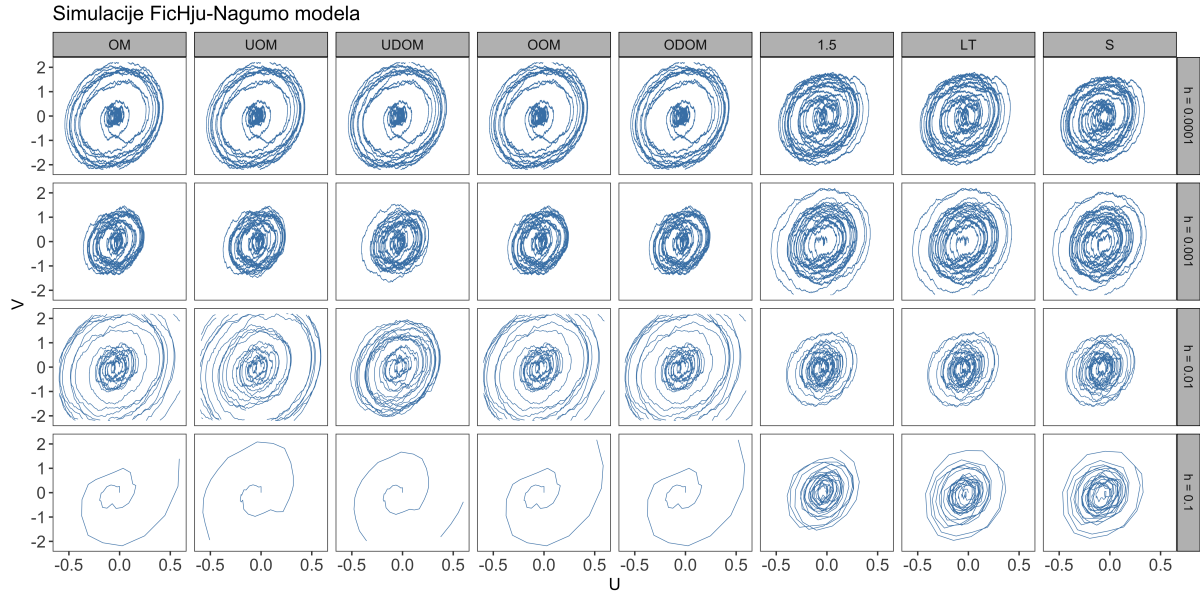
Dokaz. Videli smo da Li-Troterova aproksimacija jeste hipoeliptična u 1-koraku, a normalnost sledi direktno iz svojstva slučajnih vektora ξ_i . ■

3.4.4 Očuvanje strukture modela

U ovom delu ćemo prikazati prednosti metoda deljenja u odnosu na ostale metode kada je reč o geometrijskim svojstvima kao što su amplituda, frekvencija i orbita kod oscilatornih modela. Ovo ćemo pokazati empirijski na primeru FicHju-Nagumovog modela. Prisetimo se formulacije modela

$$\begin{cases} dU_t &= \frac{1}{\varepsilon}(U_t - U_t^3 - V_t + s)dt + \sigma_1 dW_t, \\ dV_t &= (\gamma U_t - V_t + \alpha)dt + \sigma_2 dW_t. \end{cases} \quad (3.62)$$

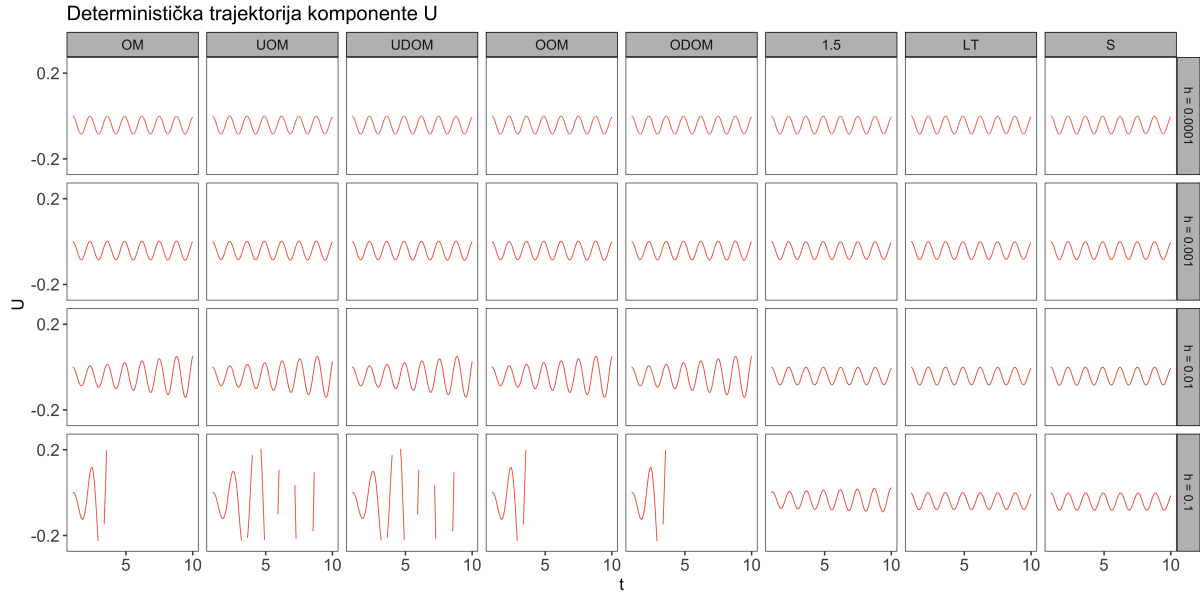
Razmatraćmo najpre vremenski interval $T = 30$. Izabraćemo da je $\sigma_1 = 0$ kako bi model bio hipoeliptičan. Takođe, stavićemo i da je $s = 0$ jer se ovaj parametar smatra poznatim, a za samu simulaciju nije od ključne važnosti. Biramo ostale parametre tako da je $\varepsilon = 1$, $\gamma = 20$, $\alpha = 0.8$, $\sigma_2 = 0.4$. Na slici 3.1 vidimo trajektoriju FicHju-Nagumovog modela simuliranu različitim metodama diskretizacije i za različite veličine koraka diskretizacije h . Metode koje smo koristili su Ojler-Marujamove metode, metoda 1.5 i dve metode deljenja. Primetimo da nismo koristili Majlštajnovu metodu. Razlog za to je što se Majlštajnova metoda i Ojler-Marujamova metoda podudaraju kada je šum aditivan.



Slika 3.1: Simulacija trajektorija FicHju-Nagumovog modela na intervalu dužine $T = 30$, za parametre $\varepsilon = 1$, $s = \sigma_1 = 0$, $\alpha = 0.8$ i $\sigma_2 = 0.4$. U kolonama su prikazane različite metode simulacija i to: Ojler-Marujamova metoda (OM), ukroćena Ojler-Marujamova metoda (UOM), ukroćena difuzija Ojler-Marujamove metode (UDOM), okrnjena Ojler-Marujamova metoda (OOM), okrnjena difuzija Ojler-Marujamove metode (ODOM), metoda 1.5 (1.5), Li-Troterova metoda (LT) i Strangova metoda (S). Kroz redove se menja veličina koraka diskretizacije i to $h = 0.0001$, $h = 0.001$, $h = 0.01$ i $h = 0.1$, odozgo na dole.

Na slici 3.1 se jasno vidi da za veće korake diskretizacije jedino možemo koristiti metodu 1.5 i metode deljenja. Međutim, s obzirom na to da je metoda 1.5 vrlo teška za implementaciju i potrebno je ručno računati višestruke Itoove integrale, vidimo da su metode deljenja superiornije od ostalih. Što se tiče Ojler-Marujamove metode i njenih varijacija, vidimo da one ne uspevaju da sačuvaju orbitu oscilatornog modela i da se sa povećanjem koraka diskretizacije, povećava i greška i to toliko da za $h = 0.1$ trajektorija eksplodira, što znači da drugi momenti nisu sačuvani.

Videli smo da metode deljenja čuvaju orbitu oscilatornih modela, a sada ćemo prikazati njihovu prednost pri čuvanju amplitude i frekvencije. Da bismo ovo bolje uočili, posmatraćemo na trenutak deterministički model, odnosno uzećemo da važi $\sigma_2 = 0$. Takođe, zbog lakšeg upoređivanja, skrat ćemo vremenski interval na $T = 10$. Na slici 3.2 upoređujemo različite metode diskretizacije za simulaciju trajektorije komponente U , determinističke verzije FicHju-Nagumovog modela. Možemo videti da metode zasnovane na Ojler-Marujamovom principu ne čuvaju ni amplitudu ni frekvenciju za velike korake diskretizacije. Ovaj put je takođe uočljivo da ni metoda 1.5 ne čuva amplitudu za jako veliki korak. To znači da su se metode deljenja još jednom pokazale kao najbolje.



Slika 3.2: Simulacija trajektorije U komponente determinističkog FicHju-Nagumovog modela na intervalu dužine $T = 10$. Parametri su zadati sa $\varepsilon = 1$, $s = \sigma_1 = \sigma_2 = 0$ i $\alpha = 0.8$. U kolonama su prikazane različite metode simulacija i to: Ojler-Marujamova metoda (OM), ukroćena Ojler-Marujamova metoda (UOM), ukroćena difuzija Ojler-Marujamove metode (UDOM), okrnjena Ojler-Marujamova metoda (OOM), okrnjena difuzija Ojler-Marujamove metode (ODOM), metoda 1.5 (1.5), Li-Troterova metoda (LT) i Strangova metoda (S). Kroz redove se menja veličina koraka diskretizacije i to $h = 0.0001$, $h = 0.001$, $h = 0.01$ i $h = 0.1$, odozgo na dole.

U radu [2] pokazano je da metode deljenja čuvaju i ergodičnost i graničnu raspodelu, što izlazi van okvira ovog master rada. Za više informacija o navedenom konsultovati pomenuti rad.

Ovim smo završili treće poglavlje. Videli smo neke poznate metode diskretizacije i njihove performanse na realnom primeru. Pokazali smo da metode deljenja imaju red srednje-kvadratne konvergencije $p = 1$, kao i da, za razliku od metoda Ojler-Marujamovog tipa, čuvaju hipoeliptičnost i geometrijska svojstva oscilatornih modela. Videli smo da metoda 1.5 koja je dobila naziv po svom redu srednje-kvadratne konvergencije, dominira nad metodama deljenja po pitanju brzine konvergencije, ali ipak ima veliki problem jer je nepraktična za implementaciju, a to je ono što je ključno pri rešavanju problema iz stvarnog sveta. Sve prethodno nam daje argumente zašto su metode Li-Trotera i Stranga najbolje za korišćenje kada je reč o simulaciji trajektorija stohastičke difuzije u opštem slučaju.

Na kraju, prisetimo se teoreme 3.4.4 koja kaže da raspodela prelaska diskretizovanog modela ima normalnu raspodelu ako koristimo Li-Troterovu metodu deljenja sa odgovarajućim vektorskim poljima. Ovo će biti od ključne važnosti za dalji rad jer će nam omogućiti statističko zaključivanje kao što su ocene parametara i testiranje hipoteza.

Ocenjivanje parametara

U prethodnom poglavlju smo videli da nije nimalo trivijalno stići od postavke modela sa neprekidnim vremenom do dovoljno dobre aproksimacije koja će nam pomoći da izvučemo statističke zaključke. Videli smo da postoje razni načini i predlozi kako se model diskretizuje, ali i da su neki modeli bolji od drugih.

U ovom poglavlju ćemo pomenuti neke novije načine za ocenjivanje parametara, zasnovane na metodi 1.5. Prikazaćemo ih i ukazati na njihove prednosti i mane, kao i na njihove statističke osobine. Takođe, predložićemo ocenjivanje parametara posle diskretizacije Li-Troterovom metodom deljenja. Izabrali smo ovu diskretizaciju iz dva razloga. Prvo, to je jednostavnija od dve metode deljenja za koje smo videli da čuvaju hipoeliptičnost i ostala geometrijska svojstva kod oscilatornih modela. Drugo, ova diskretizacija nam uvek daje normalnu raspodelu prelaska, odakle odmah možemo dobiti funkciju verodostojnosti.

Na kraju ćemo empirijski porediti sve tri metode na primeru FicHju-Nagumovog modela.

4.1 Funkcija verodostojnosti

Cilj ovog poglavlja je oceniti parametar θ koristeći metod maksimalne verodostojnosti. S obzirom na to da ne znamo pravu raspodelu našeg modela, koristićemo raspodelu diskretizovanog modela $\tilde{p}_h$. S obzirom na to da je diskretizovan proces Markovljev, funkciju verodostojnosti možemo zapisati kao

$$\tilde{p}_h(\mathbf{X}_{0:n} \mid \theta) = p(\mathbf{X}_0 \mid \theta) \prod_{k=1}^n \tilde{p}_h(\mathbf{X}_k \mid \mathbf{X}_{k-1}; \theta), \quad (4.1)$$

gde je $p(\mathbf{X}_0 \mid \theta)$ raspodela inicijalne vrednosti procesa. Doprinos ove tačke modela je zanemarljiv za dovoljno veliki uzorak, tako da ćemo je zanemariti pri oceni parametara i pretpostaviti da je inicijalna vrednost deterministička. Prethodna funkcija verodostojnosti predstavlja pseudo funkciju verodostojnosti za tačno rešenje stohastičke diferencijalne jednačine. Optimizacija pseudo funkcije verodostojnosti je česta pojava u statistici i dešava se kada nam je prava raspodela nepoznata ili nedostupna. Tada menjamo tačnu funkciju verodostojnosti aproksimiranom.

U slučaju da je aproksimirana raspodela prelaska normalna sa vektorom sredine $\mathbf{A}_h(\mathbf{X}; \theta)$ i kovarijacionom matricom $\Sigma_h(\theta)$, onda parametre nalazimo tako što rešimo sledeći optimizacioni problem

$$\operatorname{argmin}_{\theta} \left\{ n \log \det \Sigma_h(\theta) + \sum_{k=1}^n (\mathbf{X}_k - \mathbf{A}_h(\mathbf{X}_{k-1}; \theta))^{\top} \Sigma_h(\theta)^{-1} (\mathbf{X}_k - \mathbf{A}_h(\mathbf{X}_{k-1}; \theta)) \right\}. \quad (4.2)$$

Bilo da koristimo metodu 1.5 ili Li-Troterovu metodu, dobićemo isti oblik funkcije koju želimo da optimizujemo. Međutim, razlika će biti u vektoru sredine i kovarijacionoj matrici. Pre nego što nastavimo sa prezentovanjem nekih načina za ocenu parametara, razmotrićemo razliku između ocena u eliptičnom i u hipoeliptičnom slučaju.

4.2 Problemi hipoeliptičnog modela

Prva razlika između eliptičnog i hipoeliptičnog slučaja jeste što se parametri unutar funkcije drifta i funkcije difuzije mogu razdvojiti i posebno oceniti u eliptičnom slučaju. Na primer, u jednodimenzionalnom slučaju gde je stohastička diferencijalna jednačina diskretizovana Ojler-Marujamovom metodom, imamo sledeću ocenu parametara

$$(\hat{\theta}_n, \hat{\sigma}_n^2) = \operatorname{argmin}_{\theta, \sigma^2} \left\{ n \log \Sigma^2(\sigma) + \sum_{k=1}^n \frac{(X_{k+1} - X_k - h\mu(X_{k-1}; \theta))^2}{h\Sigma^2(\sigma)} \right\}, \quad (4.3)$$

gde je μ funkcija drifta, a Σ funkcija difuzije. U ovom slučaju skup parametara drifta θ , nezavisan je od parametra difuzije σ jer optimizacioni problem možemo razbiti na dva dela i minimizovati svaki posebno, bez uticaja jednog na drugi. Kesler [7] je pokazao da ocena $\hat{\theta}_n$ konvergira ka pravo vrednosti θ_0 sa brzinom konvergencije $\sqrt{nh}$. Takođe, kako je

$$\mathbb{E} [(X_{k+1} - X_k - h\mu(X_{k-1}; \theta))^2] = \mathcal{O}(h), \quad (4.4)$$

za sve θ , odatle imamo konvergenciju parametra σ^2 . Ovaj parametar konvergira brzinom $\sqrt{n}$.

Kako prethodni primer nije naročito zanimljiv, pozabavićemo se slučajem kada je stohastička diferencijalna jednačina hipoeliptična. Tada se stvari dodatno komplikuju. Prvo, parametar drifta θ je sada unutar kovarijacione matrice Σ_h . Drugo, disperzija glatke koordinate je reda $\mathcal{O}(h^3)$, dok je disperzija hrapave koordinate reda $\mathcal{O}(h)$. Zbog toga pokazivanje konvergencije parametara postaje kompleksan problem. Da bismo ovo bolje razumeli, uvešćemo jedan generički hipoeliptičan model koji ćemo dalje koristiti u ovom poglavlju. Posmatrajmo model

$$\begin{cases} dX_t &= b(X_t, \mathbf{Y}_t; \psi)dt, \\ dY_t &= \mathbf{B}(X_t, \mathbf{Y}_t; \varphi)dt + \mathbf{\Gamma}(X_t, \mathbf{Y}_t; \sigma)d\mathbf{W}_t, \end{cases} \quad (4.5)$$

gde je X_t slučajni proces koji uzima vrednosti u $\mathbb{R}$, $\mathbf{Y}_t$ slučajni proces koji uzima vrednosti u $\mathbb{R}^p$, $p \geq 1$, $\mathbf{W}_t$ p -dimenzionalno Braunovo kretanje. Funkcije $a : \mathbb{R}^{p+1} \rightarrow \mathbb{R}$ i $\mathbf{A} : \mathbb{R}^{p+1} \rightarrow \mathbb{R}^p$ zavise od nepoznatih parametara ψ i φ , redom. Označimo sa $\boldsymbol{\mu} = (b, \mathbf{B}^{\top})^{\top}$ i $\theta = (\psi, \varphi)$.

Dalje, $\mathbf{\Gamma} : \mathbb{R}^{p+1} \rightarrow \mathbb{R}^{p \times p}$ je parcijalna matrica difuzije koja zavisi od nepoznatog vektora $\boldsymbol{\sigma}$. Potpuna matrica difuzije je data sa

$$\mathbf{C}(x, \mathbf{y}; \boldsymbol{\sigma}) = \begin{bmatrix} \mathbf{0}_p \\ \mathbf{\Gamma}(x, \mathbf{y}; \boldsymbol{\sigma}) \end{bmatrix} \quad (4.6)$$

gde je $\mathbf{0}_p$ p -dimenzionalni vektor nula. Pretpostavićemo dijagonalni šum, odnosno

$$\mathbf{\Gamma} = \begin{bmatrix} \sigma_1(x, \mathbf{y}; \sigma) & 0 & \dots & 0 \\ 0 & \sigma_2(x, \mathbf{y}; \sigma) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_p(x, \mathbf{y}; \sigma) \end{bmatrix}. \quad (4.7)$$

Označimo sa $\bar{\mathbf{\Gamma}} = (\sigma_1(x, \mathbf{y}; \sigma), \sigma_2(x, \mathbf{y}; \sigma), \dots, \sigma_p(x, \mathbf{y}; \sigma))$ vektor elemenata sa dijagonale parcijalne matrice difuzije. Za kraj, uvešćemo oznaku težinskog Laplasovog operatora

$$\nabla_{\bar{\mathbf{\Gamma}}}^2(\cdot) = \sum_{i=1}^p \sigma_i^2(x, \mathbf{y}; \sigma) \frac{\partial^2}{\partial y_j^2}(\cdot). \quad (4.8)$$

Radi kraćeg zapisa, uvedimo oznaku $\mathbf{Z}_t = (X_t, \mathbf{Y}_t)$. U nastavku poglavlja ćemo videti dva načina kako možemo oceniti parametre pomoću gore pomenute funkcije i pokazati da su ocene postojane i asimptotski normalne. Oba načina koriste metod 1.5 za diskretizaciju, te ćemo videti nešto bliže kako ova diskretizacija izgleda. Za posmatrani model (4.5) diskretizacija 1.5 ima oblik

$$\begin{bmatrix} \tilde{X}_{k+1} \\ \tilde{\mathbf{Y}}_{k+1} \end{bmatrix} = \begin{bmatrix} \tilde{X}_k \\ \tilde{\mathbf{Y}}_k \end{bmatrix} + \mathbf{M}_h \left(\begin{bmatrix} \tilde{X}_{k+1} \\ \tilde{\mathbf{Y}}_{k+1} \end{bmatrix}; \theta \right) + \boldsymbol{\varepsilon}_k, \quad (4.9)$$

gde je

$$\mathbf{M}_h^{(i)}(\mathbf{z}; \theta) = h \boldsymbol{\mu}^{(i)}(\mathbf{z}; \theta) + \frac{h^2}{2} \langle \boldsymbol{\mu}(\mathbf{z}; \theta), \nabla \boldsymbol{\mu}^{(i)}(\mathbf{z}; \theta) \rangle + \frac{h^2}{4} \nabla_{\bar{\mathbf{\Gamma}}}^2 \boldsymbol{\mu}^{(i)}(\mathbf{z}; \theta), \quad (4.10)$$

za $i = 1, 2, \dots, p$. Takođe je, $\boldsymbol{\varepsilon}_k \sim \mathcal{N}_{p+1}(\mathbf{0}, \boldsymbol{\Sigma}_h(\tilde{\mathbf{Z}}_k; \psi, \sigma))$ niz nezavisnih slučajnih vektora za koje važi

$$\boldsymbol{\Sigma}_h(\mathbf{z}; \psi, \sigma) = \begin{bmatrix} (\nabla_{\mathbf{y}} b(\mathbf{z}; \psi)) \mathbf{\Gamma} \mathbf{\Gamma}^\top(\mathbf{z}; \sigma) (\nabla_{\mathbf{y}} b(\mathbf{z}; \psi))^\top \frac{h^3}{3} & \nabla_{\mathbf{y}} b(\mathbf{z}; \psi) \mathbf{\Gamma} \mathbf{\Gamma}^\top(\mathbf{z}; \sigma) \frac{h^2}{2} \\ \mathbf{\Gamma} \mathbf{\Gamma}^\top(\mathbf{z}; \sigma) (\nabla_{\mathbf{y}} b(\mathbf{z}; \psi))^\top \frac{h^2}{2} & \mathbf{\Gamma} \mathbf{\Gamma}^\top(\mathbf{z}; \sigma) h \end{bmatrix}. \quad (4.11)$$

Prethodna matrica je aproksimirana po svakoj koordinati tako da je najmanji red od h prisutan.

4.3 Iterativni pristup: Ditlevsen-Samson

U radu [3], autorke S. Ditlevsen i A. Samson nisu uspele da pokažu postojanost i asimptotsku normalnost minimizujući (4.2) odjednom, već su podelili optimizacioni problem na

dva podproblema: jedan za glatke koordinate stohastičke diferencijalne jednačine (4.5), a drugi za hrapave. Preciznije, posmatrale su optimizacioni problem za parametre ψ koordinate X odvojeno od optimizacionog problema za parametre (φ, σ^2) koordinate $\mathbf{Y}$. Na kraju, ocena parametara je dobijena iterativnom metodom.

Dva podproblema se zasnivaju na pseudo marginalnim funkcijama verodostojnosti za glatke, odnosno hrapave koordinate i data su sa

$$\begin{aligned} \hat{\psi}_n = \operatorname{argmin}_{\psi} & \left\{ \frac{3}{h^3} \sum_{k=0}^{n-1} \frac{\left(X_{k+1} - X_k - h\mathbf{M}_h^{(1)}(\mathbf{Z}_k; \psi) \right)^2}{\frac{\partial b}{\partial x}(\mathbf{Z}_k; \psi) \mathbf{\Gamma} \mathbf{\Gamma}^\top(\mathbf{Z}_k; \sigma) \left(\frac{\partial b}{\partial x} \right)^\top(\mathbf{Z}_k; \psi)} \right. \\ & \left. + \sum_{k=0}^{n-1} \log \frac{\partial b}{\partial x}(\mathbf{Z}_k; \psi) \mathbf{\Gamma} \mathbf{\Gamma}^\top(\mathbf{Z}_k; \sigma) \left(\frac{\partial b}{\partial x} \right)^\top(\mathbf{Z}_k; \psi) \right\} \end{aligned} \quad (4.12)$$

Na osnovu prvog optimizacionog podproblema (4.12) najpre su ocenjeni parametri ψ glatke koordinate. Pokazano je da ocene ovih parametara konvergiraju brzinom $\sqrt{\frac{n}{h}}$. Kako su parametri hrapave koordinate prisutni samo uz članove višeg reda, pokazano je da oni nemaju uticaja na konvergenciju parametara glatke koordinate. Kada je dobijena ocena parametara $\hat{\psi}_n$ glatkih koordinata, ona je uvrštena u drugi optimizacioni podproblem

$$\begin{aligned} (\hat{\varphi}_n, \hat{\sigma}_n^2) = \operatorname{argmin}_{\varphi, \sigma^2} & \left\{ \frac{3}{h^3} \sum_{k=0}^{n-1} \log \mathbf{\Gamma} \mathbf{\Gamma}^\top(\mathbf{Z}_k; \sigma) \right. \\ & \left. + \sum_{k=0}^{n-1} (\mathbf{Y}_{k+1} - \mathbf{Y}_k - h\mathbf{M}_h(\mathbf{Z}_k; \theta))^\top (h\mathbf{\Gamma} \mathbf{\Gamma}^\top(\mathbf{Z}_k; \sigma))^{-1} (\mathbf{Y}_{k+1} - \mathbf{Y}_k - h\mathbf{M}_h(\mathbf{Z}_k; \theta)) \right\}. \end{aligned} \quad (4.13)$$

Tako su ocenjeni parametri hrapavih koordinata $\hat{\varphi}_n$, kao i parametar difuzije $\hat{\sigma}_n^2$. Pokazano je da $\hat{\varphi}_n$ konvergira ka pravo vrednosti brzinom $\sqrt{nh}$, dok $\hat{\sigma}_n^2$ konvergira brzinom $\sqrt{n}$, što je identično eliptičnom slučaju.

Slaba tačka ovog pristupa, pored toga što je potrebno iterativno ocenjivati parametre, jeste što je da bi se pokazala konvergencija parametara hrapave koordinate i parametra difuzije, potrebno da se parametri glatke koordinate fiksiraju na prave vrednosti. U nastavku ćemo kratko predstaviti glavne teoreme ovog rada.

Stav 4.3.1. *Pretpostavimo da funkcija drifta glatke koordinate ima oblik*

$$b(x, \mathbf{y}; \psi) = b_x(x; \psi) + b_{\mathbf{y}}(\mathbf{y}) \text{ ili } b(x, \mathbf{y}; \psi) = b_x(x) + \psi b_{\mathbf{y}}(x, \mathbf{y}). \quad (4.14)$$

Neka je ψ_0 prava vrednost parametra ψ . Ako su (φ, σ^2) poznati i ako je $h \rightarrow 0$, $nh \rightarrow \infty$, onda je

$$\hat{\psi}_n \xrightarrow{p} \psi_0. \quad (4.15)$$

Stav 4.3.2. *Neka su (φ_0, σ_0^2) prave vrednosti parametara (φ, σ^2) . Ako je ψ poznat i ako je $h \rightarrow 0$, $nh \rightarrow \infty$, onda je*

$$(\hat{\varphi}_n, \hat{\sigma}_n^2) \xrightarrow{p} (\varphi_0, \sigma_0^2). \quad (4.16)$$

Teorema 4.3.3. *Neka je ν stacionarna funkcija gustine za model (4.5). Ako $h \rightarrow 0$, $nh \rightarrow \infty$ i $nh^2 \rightarrow \infty$, onda je*

$$\sqrt{nh}(\hat{\varphi}_n - \varphi_0) \xrightarrow{d} \mathcal{N}\left(0, \left(\nu\left((\nabla_{\varphi}\mathbf{B}(\cdot, \varphi_0))^{\top} (\mathbf{\Gamma}\mathbf{\Gamma}^{\top}(\cdot, \sigma_0))^{-1} (\nabla_{\varphi}\mathbf{B}(\cdot, \varphi_0))\right)\right)^{-1}\right) \quad (4.17)$$

$$\sqrt{n}(\hat{\sigma}_n^2 - \sigma_0^2) \xrightarrow{d} \mathcal{N}\left(0, 2\left(\nu\left((\nabla_{\sigma}\mathbf{\Gamma}\mathbf{\Gamma}^{\top}(\cdot, \varphi_0))^{\top} (\mathbf{\Gamma}\mathbf{\Gamma}^{\top}(\cdot, \sigma_0))^{-1} (\nabla_{\sigma}\mathbf{\Gamma}\mathbf{\Gamma}^{\top}(\cdot, \varphi_0))\right)\right)^2\right), \quad (4.18)$$

gde je $\nu(f(\cdot)) = \int f(x)d\nu(x)$.

Teorema 4.3.4. *Pretpostavimo da funkcija drifta glatke koordinate ima oblik*

$$b(x, \mathbf{y}; \psi) = b_x(x; \psi) + b_{\mathbf{y}}(\mathbf{y}) \text{ ili } b(x, \mathbf{y}; \psi) = b_x(x) + \psi b_{\mathbf{y}}(x, \mathbf{y}). \quad (4.19)$$

Neka je ν stacionarna funkcija gustine za model (4.5). Ako $h \rightarrow 0$, $nh \rightarrow \infty$ i $nh^2 \rightarrow \infty$, onda je

$$\sqrt{\frac{n}{h}}(\hat{\psi}_n - \psi_0) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{3}\left(\nu\left((\nabla_{\psi}b(\cdot, \psi_0))^{\top} ((\nabla_x b(\cdot, \psi_0))\mathbf{\Gamma}\mathbf{\Gamma}^{\top}(\cdot, \sigma_0)(\nabla_x b(\cdot, \psi_0))^{\top})^{-1} \nabla_{\psi}b(\cdot, \varphi_0))\right)^{-1}\right). \quad (4.20)$$

Da bismo dokazali ova tvrđenja, potrebno je pratiti upute iz rada [7]. Dokazi prethodnih teorema kao i svih pomoćnih tvrđenja mogu se naći u radu [3]. Glavni deo za pokazivanje prethodnog je sledeći stav koji pokazuje ograničenja momenata diskretizacije 1.5.

Stav 4.3.5 (Ograničenja momenata). *Neka su X_t i $\mathbf{Y}_t$ tačna rešenja stohastičke diferencijalne jednačine (4.5) koju smo diskretizovali metodom 1.5 kao što je dato u (4.9). Tada je*

$$\mathbb{E}\left[X_{k+1} - X_k - h\mathbf{M}_h^{(1)}(\mathbf{Z}_k; \theta_0) \mid \mathbf{Z}_k = \mathbf{z}\right] = \mathcal{O}(h^3), \quad (4.21)$$

$$\mathbb{E}\left[\mathbf{Y}_{k+1} - \mathbf{Y}_k - h\mathbf{M}_h^{(2)}(\mathbf{Z}_k; \theta_0) \mid \mathbf{Z}_k = \mathbf{z}\right] = \mathcal{O}(h^3), \quad (4.22)$$

$$\mathbb{E}\left[\left(X_{k+1} - X_k - h\mathbf{M}_h^{(1)}(\mathbf{Z}_k; \theta_0)\right)^2 \mid \mathbf{Z}_k = \mathbf{z}\right] = \frac{h^3}{3} (\nabla_{\mathbf{y}}b)^{\top} \mathbf{\Gamma}\mathbf{\Gamma}^{\top} (\nabla_{\mathbf{y}}b) + \mathcal{O}(h^4), \quad (4.23)$$

$$\mathbb{E}\left[\left(X_{k+1} - X_k - h\mathbf{M}_h^{(1)}(\mathbf{Z}_k; \theta_0)\right) \left(\mathbf{Y}_{k+1} - \mathbf{Y}_k - h\mathbf{M}_h^{(2)}(\mathbf{Z}_k; \theta_0)\right)^{\top} \mid \mathbf{Z}_k = \mathbf{z}\right] = h\mathbf{\Gamma}\mathbf{\Gamma}^{\top} + \mathcal{O}(h^2), \quad (4.24)$$

$$\mathbb{E}\left[(X_{k+1} - X_k - h\mathbf{M}_h^{(1)}(\mathbf{Z}_k; \theta_0))^4 \mid \mathbf{Z}_k = \mathbf{z}\right] = \mathcal{O}(h^4), \quad (4.25)$$

$$\mathbb{E}\left[\left(\left(X_{k+1} - X_k - h\mathbf{M}_h^{(1)}(\mathbf{Z}_k; \theta_0)\right) \left(\mathbf{Y}_{k+1} - \mathbf{Y}_k - h\mathbf{M}_h^{(2)}(\mathbf{Z}_k; \theta_0)\right)^{\top}\right)^2 \mid \mathbf{Z}_k = \mathbf{z}\right] = \mathcal{O}(h^2). \quad (4.26)$$

Videli smo neke osobine diskretizacije metodom 1.5 i kako je ona iskorišćena za ocenu parametara u hipoeliptičnom scenariju. Prethodni pristup jeste dosta inovativan i rešava problem hipoeliptičnosti, ali i dalje postoji potreba za iterativnim rešenjem, što nekad nije praktično. Takođe, u dokazima postojanosti koriste se prave vrednosti parametara, što nije realno. Zato ćemo pomenuti sledeći pristup koji rešava ova dva problema.

4.4 Ocenjivanje bez iteracija: Melnjikova

U radu [11], autorka predlaže drugačiji pristup ocenjivanja parametara u dvodimenzionalnom modelu. Za razliku od prethodnog pristupa, optimizacioni problem (4.2) se ne deli na dva podproblema, ali se i dalje vodi računa o različitim brzinama konvergencije i eventualnim zavisnostima između parametara. To znači da u dokazima postojanosti i asimptotske normalnosti i dalje imamo podelu parametara po koordinatama. U radu se prvo pokaže postojanost i konvergencija za skup parametara glatke koordinate ψ . Kao i u prethodnom pristupu, dokaz prati standardne upute iz rada [7]. Ovaj put se vrednost parametra hrapave koordinate φ ne fiksira na tačnu, već može biti proizvoljna. Zatim se ocenjuju parametri φ i σ^2 . Ovi parametri neće konvergirati za proizvoljnu vrednost ψ . Međutim, u radu je pokazana postojanost i asimptotska normalnost ocena ova dva parametra kada se koristi ocena $\hat{\psi}_n$. Sve to uz dodatnu pretpostavku o linearnosti parametara koju imamo i u prethodnom pristupu.

Takođe, razlika ovog rada u odnosu na rad [3] jeste u diskretizaciji. Melnjikova koristi diskretizaciju zasnovanu na lokalnoj linearizaciji. Ovaj metod nismo opisali pre, ali u suštini se zasniva na Tejlorovom razvoju funkcije drifta i potom rešavanju linearne stohastičke diferencijalne jednačine. Ako se drift razvije zaključno sa drugim stepenom koraka h , dobijena diskretizacija se podudara sa metodom 1.5. Ovaj metod je dosta intuitivniji ali i računski praktičniji. Takođe, kako se ova dva metoda poklapaju, to znači da stav 4.3.5 o ograničenjima momenata za metod 1.5 važi i za metod zasnovan na lokalnoj linearizaciji.

Sada ćemo nešto detaljnije izdiskutovati doprinos ovog rada.

Pretpostavka o funkciji drifta glatke koordinate jeste da su parametri te koordinate linearni, tj. da se ona može zapisati kao

$$b(x, \mathbf{y}; \psi) = f(x, y) + \langle \psi, g(x) \rangle, \quad (4.27)$$

gde su funkcije f i g odgovarajućih dimenzija i takve da važe sve prethodne pretpostavke o modelu.

Uz prethodnu pretpostavku se može pokazati postojanost i asimptotska normalnost ocene $\hat{\psi}_n$ parametra ψ . Asimptotska disperzija se u određenoj meri razlikuje od one koju smo videli kod Ditlevsen-Samsoninog principa. Razlog tome krije se u činjenici da sada imamo kovarijaciju između koordinata koja nije važna kada se ocenjivanje podeli na 2 optimizaciona problema za glatku i hrapavu promenljivu. Brzina konvergencije, međutim, ostaje ista. Prethodna pretpostavka o linearnosti parametara glatke koordinate potrebna je da bi se dokazala asimptotska normalnost. Međutim, nije potrebno fiksirati parametre φ i σ^2 na prave vrednosti, kao što su to uradile Ditlevsen i Samson.

Dokaz postojanosti i asimptotske normalnosti za parametre hrapave koordinate dosta je kompleksniji i zahteva poznavanje gustoće niza slučajnih veličina (eng. *tightness*), o čemu neće biti reč u ovom radu. U suštini, može se pokazati da uvek postoji niz ocena $\hat{\psi}_n$ takav da je niz $(\hat{\psi}_n - \psi_0)$ gust. Tada se uz gustoću i postojanost ocene $\hat{\psi}_n$ pokazuje da su ocene $\hat{\varphi}_n$ i $\hat{\sigma}_n$ postojane i asimptotski normalne. Ovde ponovo moramo iskoristiti dodatnu pretpostavku o linearnosti parametra ψ . Postoji mogućnost slabljenja ovog uslova uz

Lipšicovu neprekidnost i dodatne tehnikaliije o kojima ovde neće biti više reči. Dobijene brzine konvergencije se poklapaju sa onima u Ditlevsen-Samsoninom pristupu, s razlikom što parametri nisu fiksirani na prave vrednosti.

Pored opisanog algoritma ocenjivanja, u istom radu [11] je ponuđen još jedan način ocenjivanja koji se bazira na uslovnoj metodi najmanjih kvadrata.

4.5 Metoda najmanjih kvadrata: Melnjikova

Nekada je, iz više razloga, dosta efikasnije razdvojiti parametre iz funkcije drifta od parametara iz funkcije difuzije. Prvo, time smanjujemo dimenzionalnost optimizacionog problema i tako dobijamo na efikasnosti. Drugo, u slučaju višedimenzionalnih hipoeliptičnih stohastičkih diferencijalnih jednačina lakše je generalizovati metodu najmanjih kvadrata zasnovanu na parametrizaciji drifta.

Ideja ovog pristupa jeste da se optimizuju razlike kvadrata između diskretizovanog uzorka $\mathbf{Z}_k$ i odgovarajuće očekivane vrednosti. Diskretizacija se može obaviti bilo kojom metodom koju smo videli do sada. U ovom radu autorka se bazira na diskretizaciji zasnovanoj na lokalnoj linearizaciji. Dakle, ocena parametara je data sa

$$\hat{\psi}_n = \operatorname{argmin}_{\psi} \sum_{k=0}^{n-1} \left(X_{k+1} - \mathbf{M}_h^{(1)}(X_k, \mathbf{Y}_k; \psi, \varphi) \right)^2, \quad (4.28)$$

$$\hat{\varphi}_n = \operatorname{argmin}_{\varphi} \sum_{k=0}^{n-1} \left\| \mathbf{Y}_{k+1} - \mathbf{M}_h^{(2)}(X_k, \mathbf{Y}_k; \psi, \varphi) \right\|^2. \quad (4.29)$$

Ovaj optimizacioni problem će takođe dati postojeće i asimptotski normalne ocene. Prednost ovog postupka u odnosu na prethodni jeste što su problemi nezavisni i nije potrebno fiksirati nepoznati parametar. Takođe, nema potrebe za pretpostavkom o linearnom parametru glatke koordinate. Međutim, asimptotska disperzija se razlikuje od prethodna dva pristupa. Može se desiti da ovaj pristup da mnogo gore ocene od prethodnog jer je manje robustan u prisustvu autlajera. Takođe, performanse ovog pristupa su loše kada je koeficijent difuzije velik.

Primetimo da nismo ocenili parametar σ^2 . Možemo primetiti i da ostali parametri konvergiraju nevezano za parametar σ^2 . Jedan od načina je da ubacimo ocenjene parametre drifta u neki od prethodna dva pristupa i da tako ocenimo parametar difuzije. Takođe, u radu [11] je predloženo da u slučaju dvodimenzionalne hipoeliptične stohastičke diferencijalne jednačine koja ima aditivan šum, parametar σ^2 ocenimo na osnovu uzorka hrapave koordinate i to na sledeći način

$$\hat{\sigma}_n^2 = \frac{1}{nh} \sum_{k=0}^{n-1} (Y_{k+1} - Y_k)^2. \quad (4.30)$$

Pokazano je da je ova ocena postojana i asimptotski normalna.

Videli smo dva načina ocenjivanja parametara koja su izučavana u prethodne dve godine. Sada ćemo predložiti jedan na osnovu Li-Troterove diskretizacije. Ovaj način nije suštinski drugačiji, ali želimo da istražimo kako se ponaša u poređenju sa prethodna dva.

4.6 Nov predlog ocenjivanja parametara

Počinjemo sa diskretizacijom modela (4.5) kao što smo to radili u poglavlju 3. Koristeći jednačine (3.27) i (3.28)

$$\tilde{\mathbf{Z}}^{[\text{LT}]}(t_i) = e^{\mathbf{A}(\theta)h} \mathbf{f}_h \left(\tilde{\mathbf{Z}}^{[\text{LT}]}(t_{i-1}) \right) + \boldsymbol{\varepsilon}_{i-1}(\theta), \quad (4.31)$$

gde je $\mathbf{f}_h$ rešenje nelinearne diferencijalne jednačine, a $\mathbf{A}$ matrica iz linearne stohastičke diferencijalne jednačine. Slučajni vektori $\boldsymbol{\varepsilon}_i$ imaju normalnu raspodelu $\boldsymbol{\varepsilon}_i \stackrel{i.i.d}{\sim} \mathcal{N}_d(\mathbf{0}, \mathbf{C}_h(\theta))$, gde je matrica kovarijacije data sa

$$\mathbf{C}_h(\theta) = \int_0^t e^{\mathbf{A}(\theta)(t-s)} \boldsymbol{\Gamma}(\theta) \boldsymbol{\Gamma}^\top(\theta) e^{\mathbf{A}^\top(\theta)(t-s)} ds. \quad (4.32)$$

Radi lakšeg zapisa, uvedimo oznaku za vektor sredine

$$\mathbf{M}_h(\mathbf{z}; \theta) = e^{\mathbf{A}(\theta)h} \mathbf{f}_h(\mathbf{z}; \theta). \quad (4.33)$$

Ova diskretizacija, nažalost, nema ista ograničenja momenata kao diskretizacije 1.5 i lokalna linearizacija. Preciznije, prvi momenti za glatke, odnosno hrapave koordinate, poklapaju se do na h^2 . Navedeno nam daje sledeći stav.

Stav 4.6.1 (Ograničenja momenata). *Neka je $\mathbf{Z}_t = (X_t, \mathbf{Y}_t)$ tačno rešenje stohastičke diferencijalne jednačine (4.5) koju smo diskretizovali metodom deljenja Li-Trotera. Tada je*

$$\mathbb{E} \left[X_{k+1} - \mathbf{M}_h^{(1)}(\mathbf{Z}_k; \theta_0) \mid \mathbf{Z}_k = \mathbf{z} \right] = \mathcal{O}(h^2), \quad (4.34)$$

$$\mathbb{E} \left[\mathbf{Y}_{k+1} - \mathbf{M}_h^{(2)}(\mathbf{Z}_k; \theta_0) \mid \mathbf{Z}_k = \mathbf{z} \right] = \mathcal{O}(h^2), \quad (4.35)$$

$$\mathbb{E} \left[\left(X_{k+1} - \mathbf{M}_h^{(1)}(\mathbf{Z}_k; \theta_0) \right)^2 \mid \mathbf{Z}_k = \mathbf{z} \right] = \frac{h^3}{3} (\nabla_{\mathbf{y}} b) \boldsymbol{\Gamma} \boldsymbol{\Gamma}^\top (\nabla_{\mathbf{y}} b)^\top + \mathcal{O}(h^4), \quad (4.36)$$

$$\mathbb{E} \left[\left(X_{k+1} - \mathbf{M}_h^{(1)}(\mathbf{Z}_k; \theta_0) \right) \left(\mathbf{Y}_{k+1} - \mathbf{M}_h^{(2)}(\mathbf{Z}_k; \theta_0) \right)^\top \mid \mathbf{Z}_k = \mathbf{z} \right] = h \boldsymbol{\Gamma} \boldsymbol{\Gamma}^\top + \mathcal{O}(h^2), \quad (4.37)$$

$$\mathbb{E} \left[(X_{k+1} - \mathbf{M}_h(\mathbf{Z}_k; \theta_0))^4 \mid \mathbf{Z}_k = \mathbf{z} \right] = \mathcal{O}(h^4), \quad (4.38)$$

$$\mathbb{E} \left[\left(\left(X_{k+1} - \mathbf{M}_h^{(1)}(\mathbf{Z}_k; \theta_0) \right) \left(\mathbf{Y}_{k+1} - \mathbf{M}_h^{(2)}(\mathbf{Z}_k; \theta_0) \right)^\top \right)^2 \mid \mathbf{Z}_k = \mathbf{z} \right] = \mathcal{O}(h^2). \quad (4.39)$$

Posledica prethodnog stava jeste da ne možemo iskoristiti ideju koju su Ditlevsenova i Samsonova iskoristile za dokazivanje postojanosti jer nam nedostaje preciznost u aproksimaciji. Sa druge strane, kako se u radu Melnjikove nigde ne fiksiraju parametri na pravu vrednost, za razliku od rada Ditlevseonve i Samsonove (stavovi 4.3.1 i 4.3.2), to se ograničenje prvog momenta smanjuje sa $\mathcal{O}(h^3)$ na $\mathcal{O}(h^2)$.

Primetimo da se prethodni stav neće promeniti kada ne fiksiramo parametre na pravu vrednost, tako da i u našem slučaju možemo iskoristiti isti dokaz za postojanost i asimptotsku normalnost parametara koji je Melnjikova predložila. Iz tog razloga ovde više neće

biti reči o statističkim osobinama parametara, već ćemo završiti poglavlje sa numeričkim eksperimentima i uporediti prikazane metode na primeru FicHju-Nagumovog modela.

Dakle, parametre ćemo oceniti na sledeći način

$$\hat{\theta}_n^{[LT]} = \underset{\theta}{\operatorname{argmin}} \left\{ n \log \det \mathbf{C}_h(\theta) + \sum_{k=1}^n (\mathbf{Z}_k - \mathbf{M}_h(\mathbf{Z}_{k-1}; \theta))^{\top} \mathbf{C}_h(\theta)^{-1} (\mathbf{Z}_k - \mathbf{M}_h(\mathbf{Z}_{k-1}; \theta)) \right\}. \quad (4.40)$$

Upoređićemo ovaj metod sa metodom Melnjikove, pre čega ćemo predložiti još dva načina ocenjivanja zasnovana na Tejlorovom razvoju nelinearne funkcije drifta $\mathbf{M}_h$ i kovarijacione matrice $\mathbf{C}_h$ u okolini $h = 0$. Ovo ćemo demonstrirati na primeru FicHju-Nagumovog modela.

4.7 Numerički eksperimenti

Rešavanje optimizacionog problema (4.40) nije uvek izvodljivo u praksi. Primetimo da su i funkcija drifta $\mathbf{M}_h$ i matrica kovarijacije $\mathbf{C}_h$ jako nelinearne po parametrima θ . Ovo nije slučaj u ostalim metodama diskretizacije, kao što su metoda 1.5 ili lokalna linearizacija. Sve metode, osim metoda zasnovanih na principima deljenja vektorskog polja, koriste neku vrstu Tejlorovog razvoja i imaju polinomijalni oblik po promenljivoj h . Time se najčešće dobija i polinom po parametru θ ako već nismo imali nepolinomijalne nelinearne transformacije ovih parametara u početnom modelu. Sa druge strane, posle kompozicije rešenja metodom Li-Trotera, parametre i korak diskretizacije h transformišemo nelinearno, funkcijama kao što su eksponencijalna ili neka trigonometrijska. Zbog toga ćemo funkciju drifta i kovarijacionu matricu razviti u Tejlorov red drugog stepena. Takva funkcija drifta će se razlikovati od one dobijene u lokalnoj linearizaciji, ali će kovarijaciona matrica biti ista. To će nam dati još jednu ocenu parametara. Dakle, prvo ćemo parametre ocenjivati jednačinom (4.40), a potom ćemo tu jednačinu razviti Tejlorovom formulom, pa ponovo oceniti parametre. Ove dve metode će nam dati različite parametre koje ćemo uporediti. Ako se funkcija drifta razvije do drugog stepena od h , numerički se može pokazati da su svojstva kao što su amplituda i frekvencija i dalje sačuvana. Međutim, ako uzmemo u obzir grublji razvoj do prvog stepena, tada više nemamo očuvanost ovih geometrijskih svojstava. Bez obzira, pogledaćemo i ove ocene jer će zapravo biti ono što bi nam Ojler-Marujamov metod dao u eliptičnom smislu.

Dakle, imamo ocenu dobijenu aproksimiranjem funkcija drifta i kovarijacione matrice

$$\hat{\theta}_n^{[LTa]} = \underset{\theta}{\operatorname{argmin}} \left\{ n \log \det \tilde{\mathbf{C}}_h(\theta) + \sum_{k=1}^n \left(\mathbf{Z}_k - \tilde{\mathbf{M}}_h(\mathbf{Z}_{k-1}; \theta) \right)^{\top} \tilde{\mathbf{C}}_h(\theta)^{-1} \left(\mathbf{Z}_k - \tilde{\mathbf{M}}_h(\mathbf{Z}_{k-1}; \theta) \right) \right\}, \quad (4.41)$$

gde su $\tilde{\mathbf{M}}_h$ i $\tilde{\mathbf{C}}_h$ aproksimacije dobijene Tejlorovim razvojem u okolini $h = 0$ zaključno sa stepenom h^2 .

Možemo iskoristiti ideju o marginalnoj metodi najmanjih kvadrata i primeniti je u Li-Troterovoj diskretizaciji. Uz to, da bismo potpuno ubrzali ocenjivanje parametara, aproksimiraćemo funkciju drifta Tejlorovim razvojem prvog stepena, a kovarijacionu matricu ćemo zadržati tako da na svakoj koordinati učestvuje najmanji stepen od h . Tada

za FicHju-Nagumov model (4.48) imamo sledeće optimizacione probleme

$$(\hat{\alpha}_n, \hat{\gamma}_n) = \operatorname{argmin}_{\alpha, \gamma} \sum_{k=0}^{n-1} \left(U_{k+1} - \widetilde{\mathbf{M}}_h^{(1)}(U_k, V_k; \alpha, \gamma, \varepsilon) \right)^2, \quad (4.42)$$

$$\hat{\varepsilon}_n = \operatorname{argmin}_{\varepsilon} \sum_{k=0}^{n-1} \left(V_{k+1} - \widetilde{\mathbf{M}}_h^{(2)}(U_k, V_k; \alpha, \gamma, \varepsilon) \right)^2. \quad (4.43)$$

Da bismo ocenili parametar σ^2 možemo iskoristiti marginalnu funkciju verodostojnosti za koordinatu V_t , odnosno

$$\hat{\sigma}_n^2 = \operatorname{argmin}_{\sigma^2} \left\{ n \log \widetilde{\mathbf{C}}_h^{(2,2)}(\varepsilon, \sigma) + \frac{1}{\widetilde{\mathbf{C}}_h^{(2,2)}(\varepsilon, \sigma)} \sum_{k=1}^n \left(V_{k+1} - \widetilde{\mathbf{M}}_h^{(2)}(U_k, V_k; \alpha, \gamma, \varepsilon) \right)^2 \right\}. \quad (4.44)$$

Funkcije $\widetilde{\mathbf{M}}_h$ i $\widetilde{\mathbf{C}}_h$ su Tejlorovom formulom razvijene do najmanjeg stepena od h . U slučaju FicHju-Nagumovog modela one su date sa

$$\widetilde{\mathbf{M}}_h(u, v; \alpha, \gamma, \varepsilon) = \begin{bmatrix} u + \frac{h}{\varepsilon}(u - u^3 - v) \\ v + h(\gamma u - v + \alpha) \end{bmatrix}, \quad (4.45)$$

i

$$\widetilde{\mathbf{C}}_h(\varepsilon, \sigma) = \sigma^2 \begin{bmatrix} \frac{h^3}{3\varepsilon^2} & \frac{h^2}{2\varepsilon} \\ \frac{h^2}{2\varepsilon} & h \end{bmatrix}. \quad (4.46)$$

Primetimo da je vektor drifta sada isti kao i u Ojler-Marujamovoj diskretizaciji. Dakle, ove ocene odgovaraju onima u Ojler-Marujamovom slučaju kada bi difuzija bila eliptična. Međutim, pokazano je da su i u eliptičnim difuzijama ocene dobijene Ojler-Marujamovom diskretizacijom dosta pristrasne i ta pristrasnost raste sa povećanjem koraka h . Bez obzira na pomenuto, nastavićemo sa ovim ocenama da bismo ih uporedili sa ostalim. Prethodni optimizacioni problem se svodi na sistem jednačina

$$\begin{cases} \hat{\alpha}_n &= \frac{1}{nh} \sum_{k=1}^n (V_k - V_{k-1} + h(\hat{\gamma}_n U_{k-1} - V_{k-1})), \\ \hat{\gamma}_n &= \frac{\sum_{k=1}^n (V_k - V_{k-1} - h(\hat{\alpha}_n - V_{k-1})) U_{k-1}}{h \sum_{k=1}^n U_{k-1}^2}, \\ \hat{\varepsilon}_n &= h \frac{\sum_{k=1}^n (U_{k-1} - U_{k-1}^3 - V_{k-1})^2}{\sum_{k=1}^n (U_k - U_{k-1})(U_{k-1} - U_{k-1}^3 - V_{k-1})}, \\ \hat{\sigma}_n^2 &= \frac{1}{nh} \sum_{k=1}^n (V_k - V_{k-1} - h(\hat{\gamma}_n U_{k-1} - V_{k-1} + \hat{\alpha}_n))^2. \end{cases} \quad (4.47)$$

Sada smo spremni da poredimo ocene. To ćemo učiniti na primeru FicHju-Nagumovog modela. Poređićemo 4 različita načina dobijanja ocena:

1. Diskretizacija lokalnom linearizacijom i potom ocenjivanje svih parametara odjednom, bez iteracije. Ovu ocenu ćemo označavati sa $\hat{\theta}_n^{[LL]}$.
2. Diskretizacija Li-Troterovom metodom i ocenjivanje svih parametara odjednom bez Tejlоровog razvoja. Ovu ocenu ćemo označavati sa $\hat{\theta}_n^{[LT]}$.
3. Diskretizacija Li-Troterovom metodom i ocenjivanje svih parametara odjednom sa Tejlоровim razvojem drugog stepena. Ovu ocenu ćemo označavati sa $\hat{\theta}_n^{[LTa]}$.
4. Diskretizacija Li-Troterovom metodom i ocenjivanje parametara preko marginalnih funkcija verodostojnosti za svaku koordinatu sa Tejlоровim razvojem prvog stepena. Ovu ocenu ćemo označavati sa $\hat{\theta}_n^{[LTm]}$.

Model je dat sa

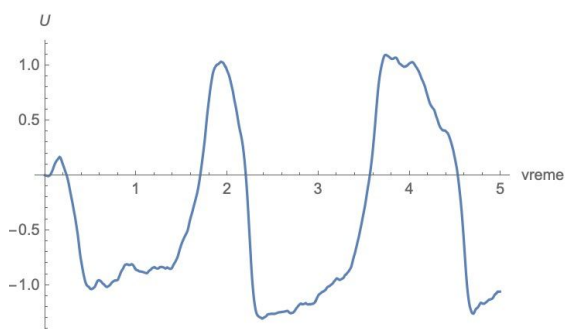
$$\begin{cases} dU_t &= \frac{1}{\varepsilon}(U_t - U_t^3 - V_t)dt, \\ dV_t &= (\gamma U(t) - V(t) + \alpha)dt + \sigma dW_t. \end{cases} \quad (4.48)$$

Prave vrednosti parametara ćemo fiksirati na $(\alpha, \varepsilon, \gamma, \sigma) = (0.3, 0.1, 1.5, 0.6)$. Da bismo imali fer eksperiment, model ćemo simulirati Ojler-Marujamovom metodom za $h = 0.0001$ na vremenskom intervalu $T = 5$. Zatim ćemo izvlačiti svaki stoti element iz uzorka i taj uzorak koristiti za ocenjivanje parametara. Tako izbegavamo sve posledice Ojler-Marujamove diskretizacije jer fiksiramo jako mali korak diskretizacije h . Sa druge strane, poduzorkovanjem izbegavamo posedovanje ogromne količine podataka koja bi prilično usporila simulaciju. Na kraju dobijamo uzorak na intervalu $T = 5$ sa korakom diskretizacije $h = 0.01$ i obimom uzorka $n = 500$.

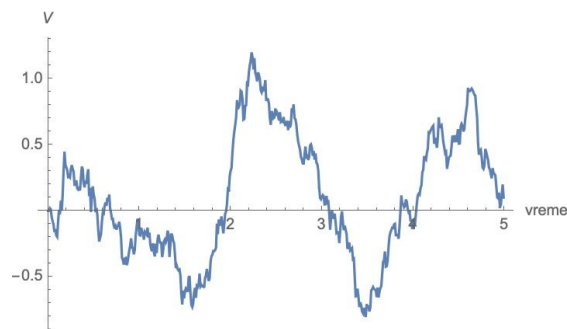
Na slici 4.1 vidimo jedan primer trajektorije FicHju-Nagumovog modela za date parametre i hiperparametre. Jasno se vidi razlika između glatke koordinate i hrapave koordinate, a takođe je primetan stohastički efekat i na glatkoj koordinati. Na slici 4.1c vidimo jednu evoluciju modela kroz vreme, praćenu po obe koordinate.

Da bismo dobili empirijsku raspodelu parametara, koristićemo Monte Karlo metodu. Simuliraćemo nMC puta jednu trajektoriju modela sa pravim vrednostima parametara, a potom ćemo oceniti parametre u svakoj iteraciji. S obzirom na to da za ocenu parametara moramo rešavati optimizacioni problem, broj Monte Karlo iteracija će biti $nMC = 100$. Svesni smo da ovo nije dovoljno da bi se sa preciznošću ocenilo tačno očekivanje za svaki od parametara, ali je dovoljno da vidimo približnu raspodelu parametara na osnovu koje možemo zaključivati o pristrasnosti.

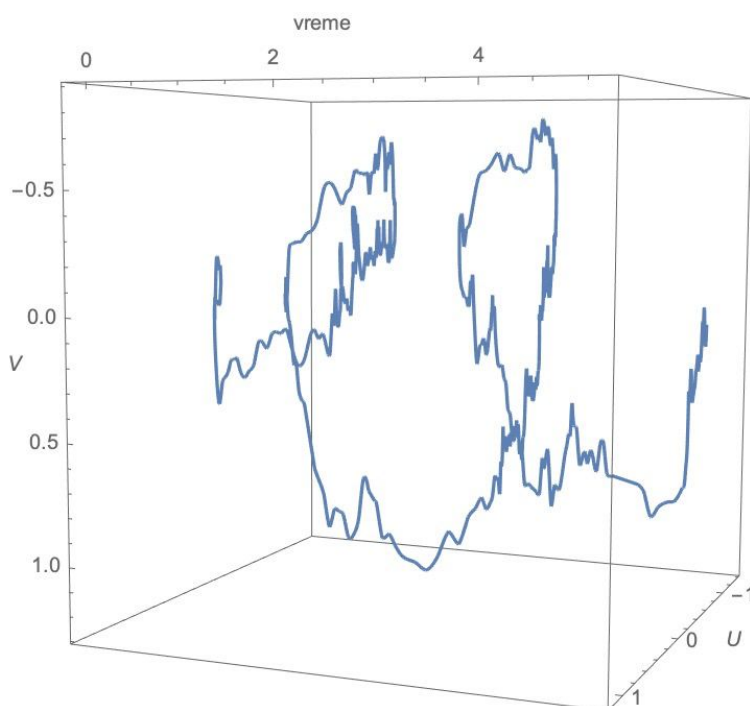
S obzirom na to da je potrebno numerički rešavati optimizacioni problem, mi ćemo samo ponoviti simulaciju iz rada [11] i koristiti funkciju `optim` u programskom jeziku R. Metod optimizacije će biti `CG` kao i u samom radu. Ovaj metod se naziva konjugovanim gradijentnim metodom (eng. *conjugate gradient method*) i zasniva se na iterativnoj optimizaciji u pravcu gradijenta. Kako ova oblast prevazilazi granice ovog rada, o navedenom metodu dalje neće biti reči.



(a) Glatka koordinata, U_t



(b) Hrapava koordinata, V_t



(c) Trajektorija modela kroz vreme za obe koordinate

Slika 4.1: Primer trajektorije FicHju-Nagumovog modela na vremenskom intervalu $T = 5$, za korak diskretizacije $h = 0.01$ i za parametre $(\alpha, \varepsilon, \gamma, \sigma) = (0.3, 0.1, 1.5, 0.6)$.

U nastavku ćemo prikazati empirijske raspodele svakog od parametara dobijenih različitim metodama ocenjivanja. Poređićemo stvarne vrednosti sa ocenjenim i videti koji metod najbolje radi.

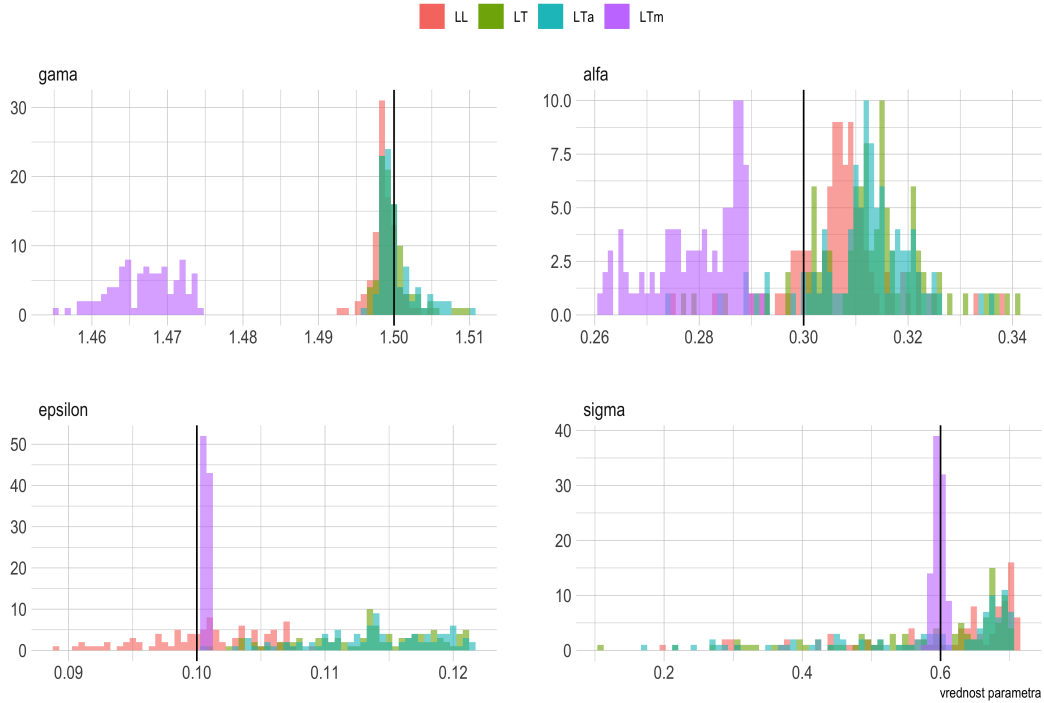
U tabeli 4.1 možemo porediti ocenjene vrednosti parametara i stvarne vrednosti sa kojima smo generisali model. Već smo pomenuli da je obim Monte Karlo simulacije jako mali da bismo donosili sigurne i precizne zaključke, ali možemo pogledati oko kojih vrednosti se ocene kreću.

Primetimo ovde da za rešavanje optimizacionog problema moramo dati početne vrednosti

θ	θ_0	$\hat{\theta}^{[LL]}$	$\hat{\theta}^{[LT]}$	$\hat{\theta}^{[LTa]}$	$\hat{\theta}^{[LTm]}$
γ	1.5	1.498 (0.0002)	1.500 (0.0002)	1.500 (0.0003)	1.467 (0.0034)
α	0.3	0.307 (0.0012)	0.313 (0.0017)	0.311 (0.0015)	0.279 (0.0023)
ε	0.1	0.101 (0.0005)	0.114 (0.0015)	0.114 (0.0015)	0.101 (0.0000)
σ	0.6	0.606 (0.0126)	0.583 (0.0132)	0.575 (0.0143)	0.597 (0.0009)

Tabela 4.1: Srednje vrednosti (standardna odstupanja) ocenjenih parametara za različite metode ocenjivanja.

parametara. S obzirom na to da ne želimo da se bavimo dodatnim numeričkim greškama u ovom poglavlju, svaki put smo pokretali optimizacioni algoritam sa inicijalnim vrednostima postavljenim na prave vrednosti parametara. Takođe, ocene koje odgovaraju koloni $\hat{\theta}^{[LTm]}$, dobijene marginalizacijom funkcije verodostojnosti i dobijene najgrubljom aprosimacijom Tejlorovim razvojem, ne zahtevaju inicijalne vrednosti jer se mogu pronaći kao rešenje linearnog sistema. Ovo nije uopšteni slučaj i zahteva da parametri budu linearni po svim koordinatama. Takođe, ovaj metod u našem slučaju daje najbrže ocene, ali po cenu pristrasnosti. U tabeli vidimo da je u najvećoj meri metod zasnovan na lokalnoj linearizaciji najbolje ocenio parametre. Naravno, opet pominjemo da je broj iteracija premali da bismo sa sigurnošću mogli da zaključimo da je ovaj metod najbolji, ali teorijska podloga nam govori da bi ovaj metod trebalo da nadjača ostale.



Slika 4.2: Raspodele parametara za različite metode ocenjivanja dobijene Monte Karlo simulacijom sa 100 iteracija

Na slici 4.2 možemo videti empirijske raspodele parametara za svaku od metoda ocenjivanja. Ono što odmah vidimo jeste pristrasnost ocena $\hat{\alpha}_n^{[LTm]}$ i $\hat{\gamma}_n^{[LTm]}$. Sa druge strane, ovaj metod daje vrlo dobre rezultate za ostala dva parametra gde disperzija gotovo da ne postoji. Pored toga, ostale metode ocenjivanja daju mnogo tačniju ocenu parametra γ , gde već vidimo da raspodele liče na normalnu. Što se tiče parametara ε i σ , oni su najteži za oceniti ih i vidimo da za 100 iteracija raspodela ovih parametara nije ni počela da se formira. Disperzija je jako velika i sve ocene deluju uniformno raspoređeno. Očekujemo da se ovo ispravi sa povećanjem broja Monte Karlo iteracija. Vidimo da postoji pristrasnost i za ocenu α , kao i da nijedan metod nije dao nepristrasnu ocenu. Ovo se možda popravi sa većim brojem iteracija, međutim, to ne možemo tvrditi sa sigurnošću.

Dakle, u ovom poglavlju smo videli neke od načina kako možemo dobiti ocene parametara u modelima zasnovanim na neprekidnom vremenu. Neke metode su lakše za implementaciju, dok su neke komplikovanije. Neke metode daju nešto bolje ocene u empirijskim studijama, dok za neke imamo sve važne asimptotske rezultate teorijski dokazane. Ovaj problem i dalje ostaje otvoren bez konkretnog rešenja i putokaza kako najbolje oceniti parametre u opštem slučaju. Nadamo se da ćemo se u daljem istraživanju više približiti odgovoru na to pitanje. Do tad, prelazimo na poslednje poglavlje u ovoj priči o statističkom zaključivanju u stohastičkim difuzijama.

Testiranje saglasnosti podataka sa modelom

Ocenjeni parametri modela ne služe ničemu ukoliko model ne odgovara podacima. Upravo zato važan deo svakog statističkog zaključivanja jeste testiranje saglasnosti (eng. *goodness-of-fit*) modela sa podacima. Kao i u slučaju ocenjivanja parametara, ne postoji nijedan opšteprihvaćeni i sveobuhvatni metod koji rešava ovaj problem. Takođe, posmatrano i van okvira stohastičkih difuzija, u nekim prostijim modelima postoji puno testova saglasnosti koji su dobri u određenim situacijama, a loši u nekim drugim. Tako je i sama oblast konstrukcije i biranje testova saglasnosti za proizvoljnu raspodelu velika i kompleksna i nema jedinstvenu metodologiju postupanja u određenim situacijama.

Većina testova koji su konstruisani za stohastičke diferencijalne jednačine bavi se isključivo jednačinama u jednoj dimenziji, kao i onim koje nemaju parametre. S obzirom na to da želimo generalizovani primer, predstavimo test koji je opisan u jednoj dimenziji i predložićemo njegovu adaptaciju za testiranje saglasnosti sa FicHju-Nagumovim modelom. Za opisano ćemo koristiti marginalnu raspodelu hrapave koordinate kao polaznu jednodimenzionalnu stohastičku diferencijalnu jednačinu.

U ovom poglavlju ćemo detaljno predstaviti jedan od mnogih načina testiranja saglasnosti. Ovaj test će se bazirati na testiranju raspodele reziduala modela. Videćemo neke važne teorijske osobine samog testa i numerički ih ispitati.

Pre toga, pomenimo još jedan test koji posmatra model bez parametara.

5.1 Negri-Nišijaminov test saglasnosti

Test je predstavljen u radu [12] gde autori I. Negri i J. Nišijama pretpostavljaju model sledećeg oblika

$$X_t dt = S(X_t)dt + \varepsilon \sigma(X_t) dW_t, \quad X_0 = x_0, \quad t \in [0, T]. \quad (5.1)$$

Funkcija drifta S i funkcija difuzije σ zadovoljavaju sve pretpostavke o postojanju i jedinstvenosti rešenja prethodne stohastičke diferencijalne jednačine. Proces W_t je standardno Braunovo kretanje. Autori u radu izučavaju asimptotski slučaj kada $\varepsilon \rightarrow 0$. Nulta

hipoteza koju želimo da testiramo je

$$H_0 : S = S_0, \text{ za dato } S_0, \text{ protiv } H_1 : S \neq S_0. \quad (5.2)$$

U radu je posmatrana sledeća uopštena diskretizacija. Proces $X = \{X_t \mid t \in [0, T]\}$ je posmatran u trenucima

$$0 = t_0^\varepsilon < t_1^\varepsilon < \dots < t_{n(\varepsilon)}^\varepsilon = T, \quad (5.3)$$

tako da je $h_\varepsilon = o(\varepsilon^2)$, kad $\varepsilon \rightarrow 0$, gde je

$$h_\varepsilon = \max_{1 \leq i \leq n(\varepsilon)} |t_i^\varepsilon - t_{i-1}^\varepsilon|. \quad (5.4)$$

Bez umanjenja opštosti, pretpostavlja se $h_\varepsilon \leq 1$ i $\varepsilon \leq 1$. Označimo rešenje obične diferencijalne jednačine

$$\frac{dx_t^S}{dt} = S(x_t^S), \quad x_0^S = x_0, \quad (5.5)$$

sa $x^S = \{x_t^S \mid t \in [0, T]\}$. Tada je alternativna hipoteza zapravo $H_1 : S \in \mathcal{S}$, gde klasu $\mathcal{S}$ funkcija S definišemo tako da

$$\int_0^{u_S} (S(x_t^S) - S_0(x_t^S)) dt \neq 0, \quad \text{za neko } u_S \in [0, T]. \quad (5.6)$$

Naravno, za funkcije $S \in \mathcal{S}$ važe svi regulatorni uslovi za postojanje i jedinstvenost rešenja diferencijalne jednačine. Označimo sada

$$\Sigma_{S,\sigma} = \sqrt{\int_0^T \sigma^2(x_t^S) dt} > 0. \quad (5.7)$$

Pokazano je da je

$$\hat{\Sigma}^\varepsilon = \sqrt{\frac{1}{\varepsilon^2} \sum_{i=1}^{n(\varepsilon)} (X_{t_i^\varepsilon} - X_{t_{i-1}^\varepsilon})^2} \quad (5.8)$$

postojana ocena za $\Sigma_{S,\sigma}$. Test statistika iz rada se bazira na sledećem slučajnom procesu $U^\varepsilon = \{U^\varepsilon(u) \mid u \in [0, T]\}$ datom sa

$$U^\varepsilon(u) = \frac{1}{\varepsilon} \sum_{i=1}^{n(\varepsilon)} \mathbf{1}(t_i^\varepsilon \leq u) \left[X_{t_i^\varepsilon} - X_{t_{i-1}^\varepsilon} - S_0(X_{t_{i-1}^\varepsilon})(t_i^\varepsilon - t_{i-1}^\varepsilon) \right]. \quad (5.9)$$

Na kraju, konstruisana test statistika za koju je pokazano da je postojana protiv alternative $H_1 : S \in \mathcal{S}$ i asimptotski slobodna od raspodele pod H_0 , data je sa

$$\frac{\sup_{u \in [0, T]} |U^\varepsilon(u)|}{\hat{\Sigma}^\varepsilon} \xrightarrow{d} \sup_{t \in [0, T]} |\widetilde{W}_t|, \quad (5.10)$$

gde je $\widetilde{W}_t$ standardno Braunovo kretanje.

Glavni doprinos ovog testa jeste što je asimptotski slobodan od raspodele i što je jedan od prvih testova koji se bazira na diskretizovanom uzorku umesto na uzorku na osnovu

opservacijama iz neprekidnog vremena, kao što su to radili testovi pre ovog. Sa druge strane, mane ovog testa jesu u tome što model ne podrazumeva parametre u funkciji drifta i difuzije, kao i što pretpostavlja parametar ε koji je uslovljen da teži ka nuli. Time se gubi na opštosti modela jer se podrazumeva da je šum u modelu jako mali. Iako bi bilo zanimljivo uporediti ovaj i sledeći test, to neće biti urađeno u ovom master radu i ostaće otvoreno za neka druga istraživanja. Sada prelazimo na sledeći test koji ćemo numerički testirati.

5.2 Lijev test saglasnosti

Pre nego što uvedemo test i model, važno je napomenuti da autor radova [14] i [13] Sangjeol Li (eng. *Sangyeol Lee*) nije ista osoba po kojoj je Li-Troterova diskretizacija dobila naziv (Sofus Li, eng. *Sophus Lie*). U radovima [14] i [13] autori predlažu test saglasnosti na osnovu empirijskog procesa reziduala. U prvom radu se ne pretpostavljaju parametri u funkciji difuzije, dok se drugi rad nadograđuje na prvi, a potom uopštava tako da su parametri dopušteni i u funkciji difuzije. Oba rada se baziraju na modelu u jednoj dimenziji. Pokazano je da asimptotska raspodela empirijskog procesa reziduala ne zavisi od ocenjenih parametara, odnosno da je test statistika zasnovana na ovom procesu slobodna od raspodele. Ovo je pokazano nauštrb obima uzorka. Naime, ako posmatramo uzorak obima n , na vremenskom intervalu T , sa korakom diskretizacije h , gde je $n = T/h$ ceo broj, onda empirijski proces reziduala posmatra samo prvih T umesto svih n reziduala. Time smo izgubili ogromnu većinu podataka, ali smo dobili asimptotski slobodan test, tako da je cena koju smo platili diskutabilno zadovoljavajuća. Sada ćemo malo detaljnije prikazati model i test statistiku.

Posmatrajmo sledeću stohastičku diferencijalnu jednačinu

$$dX_t = a(X_t; \theta)dt + b(X_t; \sigma)dW_t, \quad X_t = x_0, \quad t \in [0, T], \quad (5.11)$$

gde je (θ, σ) $p + q$ -dimenzionalni nepoznat parametar. Funkcije a i b su realne i takve da važe svi regulatorni uslovi za stohastičku diferencijalnu jednačinu. Standardno, proces W_t je standardno Braunovo kretanje. Hipoteza koju želimo da testiramo je

$$H_0 : W_t \text{ je standardno Braunovo kretanje} \quad \text{protiv} \quad H_1 : \neg H_0. \quad (5.12)$$

Dok se prethodni test zasnivao na testiranju funkcionalne hipoteze, odnosno oblika funkcije drifta S , ovaj test se bazira na testiranju raspodele reziduala modela. U ovim radovima se pretpostavlja sledeća diskretizacija procesa X . Neka je $\{h_n\}_{n \in \mathbb{N}}$ niz pozitivnih brojeva takav da $h_n \rightarrow 0$ kad $n \rightarrow \infty$. Neka važe i sledeći uslovi

$$nh_n \rightarrow \infty, \quad nh_n^2 \rightarrow 0, \quad \text{i} \quad \frac{\sqrt{nh_n}}{\ln n} \rightarrow \infty, \quad \text{kad} \quad n \rightarrow \infty. \quad (5.13)$$

Posmatrajmo onda diskretizaciju X_{t_i} , gde je $t_i = ih_n$, za $i = 1, 2, \dots, n$. U ovim radovima se ne bavi ocenjivanjem parametara i podrazumeva se da su parametri ocenjeni tako da budu postojani i asimptotski normalni sa sledećim redovima konvergencije

$$\sqrt{nh_n}(\hat{\theta}_n - \theta) = \mathcal{O}_{\mathbb{P}}(1) \quad \text{i} \quad \sqrt{n}(\hat{\sigma}_n - \sigma) = \mathcal{O}_{\mathbb{P}}(1). \quad (5.14)$$

Primetimo da bar za dve vrste ocenjivanja znamo da ispunjavaju ove osobine, kao što smo videli u prethodnom poglavlju. S obzirom na to da je jedino za ocene dobijene na osnovu lokalne linearizacije formalno pokazano da su parametri postojani i asimptotski normalni sa istim redovima konvergencije, bez pretpostavke o fiksiranju parametara na prave vrednosti, koristićemo te ocene za numeričko testiranje ovog testa.

U radovima [14] i [13] se koristi Ojler-Marujamov metod diskretizacije da bi se dobili reziduali. Ovaj metod, kao što znamo, neće raditi u hipoeliptičnom slučaju. Međutim, u ovim radovima nije naglašeno kako su dobijene ocene, pa ćemo probati da ispitamo sledećim postupkom. Generišemo model Ojler-Marujamovom metodom kao i do sada, sa finim korakom diskretizacije, pa izvlačimo poduzorak da bismo smanjili obim uzorka. Zatim, ocenimo parametre hipoeliptičnog modela nekom od metoda iz prethodnog poglavlja. Nakon toga možemo posmatrati marginalnu raspodelu hrapavih koordinata i nju diskretizovati Ojler-Marujamovom metodom. Tada ćemo posmatrati rezidualne samo hrapavih koordinata i testirati hipotezu o njihovoj normalnosti. Ono što želimo da vidimo jeste da li će ovaj način ocenjivanja parametara i testiranja saglasnosti uticati na efikasnost testa. Svi podaci su poznati, imamo uzorak i sa glatke i sa hrapave koordinate, kao i ocenjene parametre koji ispunjavaju uslove iz radova [14] i [13].

Dakle, posmatrajmo diskretizovani model

$$X_{t_i} - X_{t_{i-1}} \approx h_n a(X_{t_{i-1}}; \theta) + b(X_{t_{i-1}}; \sigma)(W_{t_i} - W_{t_{i-1}}), \quad (5.15)$$

i za njega definišimo rezidualne i ocenjene rezidualne, redom

$$\eta_{ni} = \frac{W_{t_i} - W_{t_{i-1}}}{\sqrt{h_n}}, \quad (5.16)$$

$$\hat{\eta}_{ni} = \frac{X_{t_i} - X_{t_{i-1}} - h_n a(X_{t_{i-1}}; \hat{\theta}_n)}{\sqrt{h_n} b(X_{t_{i-1}}; \hat{\sigma}_n)}. \quad (5.17)$$

Sada definišemo empirijski proces reziduala kao

$$\hat{\xi}_n(x) = \frac{1}{\sqrt{[nh_n]}} \sum_{i=1}^{[nh_n]} (\mathbf{1}(\hat{\eta}_{ni} \leq x) - \Phi(x)), \quad x \in \mathbb{R}, \quad (5.18)$$

gde je Φ funkcija standardne normalne raspodele, a $[\cdot]$ funkcija ceo-deo. Primetimo ovde da je za empirijski proces reziduala iskorišćeno samo prvih $[nh_n]$ elemenata uzorka umesto celog uzorka obima n . Postavlja se pitanje da li je ova restrikcija neophodna. Prema radovima [14] i [13] ovo je jedini način da bi se pokazala asimptotska sloboda od raspodele. Pogledajmo teoremu bez dokaza koja je data u radu [13].

Teorema 5.2.1. *Pod regulatornim pretpostavkama modela, pretpostavkama o koraku diskretizacije h_n i pretpostavkama o asimptotskoj normalnosti ocena parametara važi*

$$\hat{\xi}_n(x) = \frac{1}{\sqrt{[nh_n]}} \sum_{i=1}^{[nh_n]} (\mathbf{1}(\eta_{ni} \leq x) - \Phi(x)) + R_n(x), \quad (5.19)$$

gde je

$$\sup_{x \in \mathbb{R}} |R_n(x)| = \mathcal{O}_{\mathbb{P}}(1). \quad (5.20)$$

To znači, da $\hat{\xi}_n(\Phi^{-1}(t))$ slabo konvergira ka Braunovom mostu B_t .

Dokaz prethodne teoreme je vrlo tehnički i zahteva par strana da bi se sve kalkulacije izvele do kraja, zbog čega ga nećemo ovde predstaviti, ali ćemo iskomentarisati ideju dokaza. Ceo dokaz se može pronaći u pomenutom radu. Za početak, uvedimo sledeće nizove slučajnih veličina

$$\Delta_{ni} = \int_{t_{i-1}}^{t_i} (a(X_s; \theta) - a(X_{t_{i-1}}; \theta)) ds + \int_{t_{i-1}}^{t_i} (b(X_s; \sigma) - b(X_{t_{i-1}}; \sigma)) dW_s, \quad (5.21)$$

$$d_{ni} = a(X_{t_{i-1}}; \hat{\theta}_n) - a(X_{t_{i-1}}; \theta). \quad (5.22)$$

Prvi niz se može razumeti kao pristrasnost modela u svakoj tački diskretizacije. Sada ćemo koristiti osnovne algebarske operacije kako bismo empirijski proces reziduala predstavili preko teorijskih reziduala. Počinjemo tako što model (5.11) zapisujemo u integralnom obliku

$$X_{t_{i-1}} - X_{t_i} = \int_{t_{i-1}}^{t_i} a(X_s; \theta) ds + \int_{t_{i-1}}^{t_i} b(X_s; \sigma) dW_s \quad (5.23)$$

$$= \int_{t_{i-1}}^{t_i} (a(X_s; \theta) - a(X_{t_{i-1}}; \theta)) ds + \int_{t_{i-1}}^{t_i} (b(X_s; \sigma) - b(X_{t_{i-1}}; \sigma)) dW_s \quad (5.24)$$

$$+ h_n a(X_{t_{i-1}}; \theta) + \sqrt{h_n} b(X_{t_{i-1}}; \sigma) \eta_{ni} \quad (5.25)$$

$$= \Delta_{ni} + h_n a(X_{t_{i-1}}; \theta) + \sqrt{h_n} b(X_{t_{i-1}}; \sigma) \eta_{ni}. \quad (5.26)$$

Sa druge strane, iz definicije ocenjenih reziduala (5.17) imamo da je

$$X_{t_{i-1}} - X_{t_i} = \sqrt{h_n} b(X_{t_{i-1}}; \hat{\sigma}_n) \hat{\eta}_{ni} + h_n a(X_{t_{i-1}}; \hat{\theta}_n). \quad (5.27)$$

Iz prethodnog vidimo da je veza između ocenjenih i teorijskih reziduala data sa

$$\hat{\eta}_{ni} = \frac{\sqrt{h_n} b(X_{t_{i-1}}; \sigma) \eta_{ni} + \Delta_{ni} - h_n d_{ni}}{\sqrt{h_n} b(X_{t_{i-1}}; \hat{\sigma}_n)}. \quad (5.28)$$

Konačno, uslov $\hat{\eta}_{ni} \leq x$ je ekvivalentan uslovu

$$\eta_{ni} \leq x \frac{b(X_{t_{i-1}}; \hat{\sigma}_n)}{b(X_{t_{i-1}}; \sigma)} - \frac{\Delta_{ni}}{\sqrt{h_n} b(X_{t_{i-1}}; \sigma)} + \sqrt{h_n} \frac{d_{ni}}{b(X_{t_{i-1}}; \sigma)}. \quad (5.29)$$

U narednim koracima dokaza eksploatisaćemo činjenicu da je raspodela za η_{ni} pod nultom hipotezom standardna normalna. Sada empirijski proces reziduala $\hat{\xi}_n(x)$ možemo zapisati kao

$$\begin{aligned} \hat{\xi}_n(x) &= \frac{1}{\sqrt{[nh_n]}} \sum_{i=1}^{[nh_n]} \left(\mathbf{1} \left(\eta_{ni} \leq x \frac{b(X_{t_{i-1}}; \hat{\sigma}_n)}{b(X_{t_{i-1}}; \sigma)} - \frac{\Delta_{ni}}{\sqrt{h_n} b(X_{t_{i-1}}; \sigma)} + \sqrt{h_n} \frac{d_{ni}}{b(X_{t_{i-1}}; \sigma)} \right) - \Phi(x) \right) \\ &= \frac{1}{\sqrt{[nh_n]}} \sum_{i=1}^{[nh_n]} (\mathbf{1}(\eta_{ni} \leq x) - \Phi(x)) + I_n(x) + II_n(x), \end{aligned} \quad (5.30)$$

gde su

$$\begin{aligned}
I_n(x) &= \frac{1}{\sqrt{[nh_n]}} \sum_{i=1}^{[nh_n]} \left(\Phi \left(x \frac{b(X_{t_{i-1}}; \hat{\sigma}_n)}{b(X_{t_{i-1}}; \sigma)} - \frac{\Delta_{ni}}{\sqrt{h_n} b(X_{t_{i-1}}; \sigma)} + \sqrt{h_n} \frac{d_{ni}}{b(X_{t_{i-1}}; \sigma)} \right) - \Phi(x) \right), \\
II_n(x) &= \frac{1}{\sqrt{[nh_n]}} \sum_{i=1}^{[nh_n]} \left[\mathbf{1} \left(\eta_{ni} \leq x \frac{b(X_{t_{i-1}}; \hat{\sigma}_n)}{b(X_{t_{i-1}}; \sigma)} - \frac{\Delta_{ni}}{\sqrt{h_n} b(X_{t_{i-1}}; \sigma)} + \sqrt{h_n} \frac{d_{ni}}{b(X_{t_{i-1}}; \sigma)} \right) \right. \\
&\quad \left. - \Phi \left(x \frac{b(X_{t_{i-1}}; \hat{\sigma}_n)}{b(X_{t_{i-1}}; \sigma)} - \frac{\Delta_{ni}}{\sqrt{h_n} b(X_{t_{i-1}}; \sigma)} + \sqrt{h_n} \frac{d_{ni}}{b(X_{t_{i-1}}; \sigma)} \right) \right. \\
&\quad \left. + \Phi(x) - \mathbf{1}(\eta_{ni} \leq x) \right]. \tag{5.31}
\end{aligned}$$

Ostatak dokaza je tehnički i na nekoliko strana pokazuje da važi

$$\sup_{x \in \mathbb{R}} |I_n| = \mathcal{O}_{\mathbb{P}}(1) \text{ i } \sup_{x \in \mathbb{R}} |II_n| = \mathcal{O}_{\mathbb{P}}(1), \text{ kad } n \rightarrow \infty. \tag{5.32}$$

Postavlja se pitanje da li je moguće koristiti ceo obim uzorka n umesto samo nekolicine $[nh_n]$ u empirijskom procesu reziduala. Prema radovima [14] i [13] ispostavlja se da ovo nije moguće. Niz slučajnih veličina Δ_{ni} dat u (5.21) ne možemo smatrati uniformno zanemarljivim ako koristimo ceo uzorak i neće biti moguće staviti ga u ostatak R_n . Prisetimo se da h_n ne sme biti previše mali broj jer nh_n mora težiti ka beskonačnosti, što znači da postoji granica koliko možemo smanjiti pristrasnost modela u smislu jednačine (5.21). U praksi korišćenje samo $[nh_n]$ podataka ne mora naneti veliku štetu ako je obim uzorka dovoljno veliki. U suprotnom, sva statistička zaključivanja se dovode u pitanje jer pre svega ocene parametara nisu pouzdane.

Takođe, najčešće ocena parametra σ utiče na graničnu raspodelu. Međutim, to ovde nije slučaj. Opet, razlog tome je što koristimo samo $[nh_n]$ opservacija, dok je ocena $\hat{\sigma}_n^2$ postojana sa redom konvergencije $\sqrt{n}$. To je čini vrlo efikasnom u odnosu na obim uzorka $[nh_n]$.

Prethodni rezultat se može iskoristiti da se testira saglasnost pomoću testova normalnosti kao što su Kolmogorov-Smirnov test ili test Kramer-fon Mizesa, jer zbog teoreme o neprekidnom preslikavanju važi

$$KS_n = \sup_{u \in [0,1]} |\hat{\xi}_n(\Phi^{-1}(u))| \xrightarrow{d} \sup_{t \in [0,1]} |B_t|, \tag{5.33}$$

odnosno,

$$CVM_n = \int_0^1 |\hat{\xi}_n(\Phi^{-1}(u))|^2 du \xrightarrow{d} \int_0^1 |B_t|^2 dt, \tag{5.34}$$

gde je B_t Braunov most.

Odbacujemo nultu hipotezu H_0 da je W_t Braunovo kretanje, što je ekvivalentno hipotezi da X_t prati model difuzije (5.11) za velike vrednosti test statistika KS_n i CVM_n .

5.3 Numerički eksperimenti

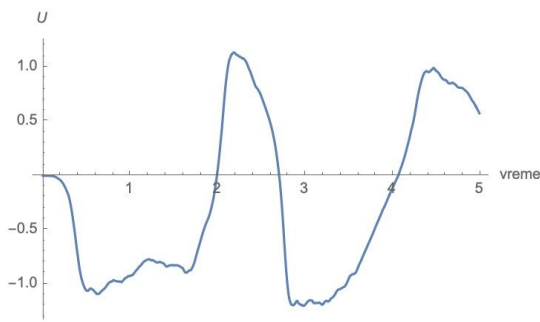
Na kraju poglavlja želimo da testiramo sve što smo do sada videli. Iskoristićemo Lijev test u kombinaciji sa Kolmogorov-Smirnov i Kramer-fon Mizesovom metrikom. Izvršićemo dva eksperimenta. Prvo, pretpostavićemo da je asimptotska raspodela pod nultom hipotezom slobodna od parametara za empirijski proces reziduala kao što je pokazano u prethodnoj teoremi. Oцениćemo meru testa koristeći Monte Karlo simulaciju. Takođe, to ćemo uraditi za različite vrednosti pravih parametara kako bismo se uverili da je mera testa ista bez obzira na njih. Zatim ćemo oceniti moć testa protiv date alternativne hipoteze. Opet, ovo ćemo uraditi za dva seta početnih parametara kako bismo ispitali da li moć testa zavisi od parametara modela.

Kao i do sada, testove ćemo izvoditi na primeru FicHju-Nagumovog modela. Nulta hipoteza će biti da podaci prate ovaj model

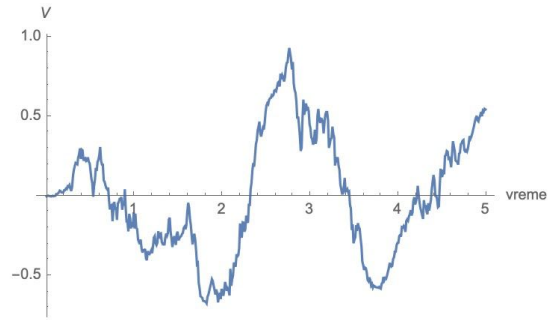
$$\begin{cases} dU_t &= \frac{1}{\varepsilon}(U_t - U_t^3 - V_t)dt, \\ dV_t &= (\gamma U(t) - V(t) + \alpha)dt + \sigma dW_t. \end{cases} \quad (5.35)$$

Za alternativnu hipotezu ćemo pretpostaviti da funkcija difuzije nije aditivna, već zavisi od trenutnog stanja procesa. Ovaj alternativni model ćemo zadati sa

$$\begin{cases} dU_t &= \frac{1}{\varepsilon}(U_t - U_t^3 - V_t)dt, \\ dV_t &= (\gamma U(t) - V(t) + \alpha)dt + \sigma \sqrt{|U_t - V_t|}dW_t. \end{cases} \quad (5.36)$$



(a) Glatka koordinata, U_t

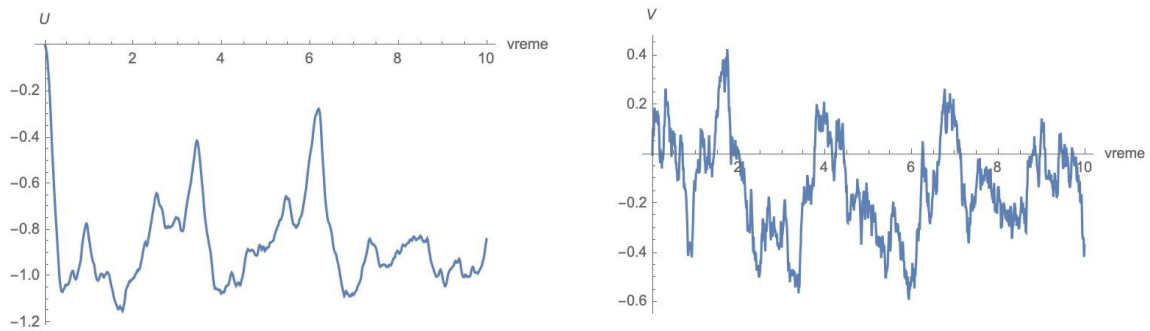


(b) Hrapava koordinata, V_t

Slika 5.1: Primer trajektorija modela pod alternativnom hipotezom H_1 na vremenskom intervalu $T = 5$, za korak diskretizacije $h = 0.01$ i za parametre $(\alpha, \varepsilon, \gamma, \sigma) = (0.3, 0.1, 1.5, 0.6)$.

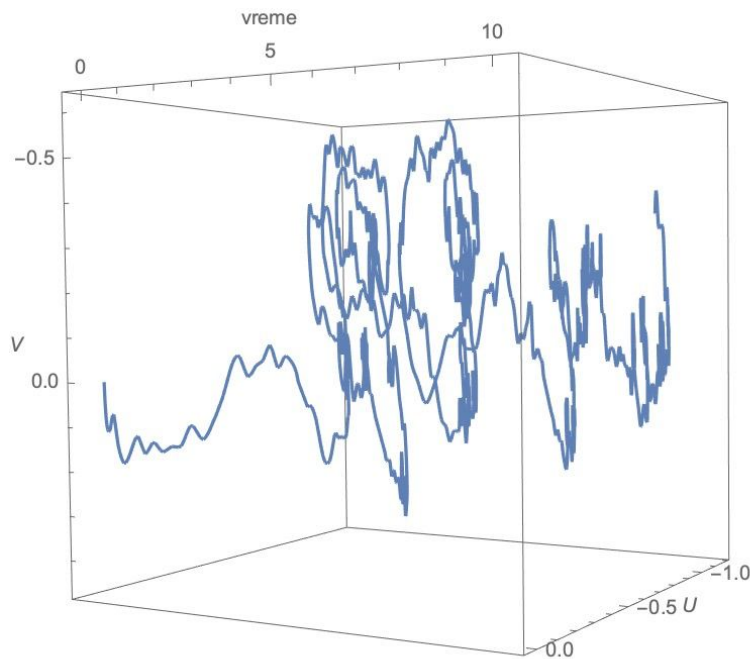
Izbor ovakve alternativne hipoteze je motivisan čestom alternativom protiv Ornštajn-Ulenbekovog procesa, a to je Koks-Ingersol-Rosov model (eng. *Cox-Ingersoll-Ross*) ili skraćeno CIR. S obzirom na to da testiramo rezidualne samo za drugu koordinatu, kada je proces prve koordinate poznat, marginalnu stohastičku diferencijalnu jednačinu za proces V možemo posmatrati kao linearnu. To znači da možemo zadati alternativu sličnu kao

u slučaju CIR modela, s tim što moramo biti oprezni. Ornštajn-Ulenbekov proces, kao i CIR model, definisani su na skupu pozitivnih brojeva, dok FicHju-Nagumov model nije. Iz tog razloga ćemo u alternativnom modelu dodati koren iz apsolutne razlike između koordinata. Razlog razmatranja obe koordinate jeste da bi se smanjio uticaj na raspršenost, kako bi modeli pod nultom i alternativnom hipotezom bili relativno iste disperzije. Na slici 5.1 vidimo primer jedne trajektorije po koordinatama za alternativni model. Ako uporedimo sliku sa 4.1 vidimo da su trajektorije približno iste i ne možemo zaključiti sa slika da dolaze od dva različita modela.



(a) Glatka koordinata, U_t

(b) Hrapava koordinata, V_t



(c) Trajektorija modela kroz vreme za obe koordinate

Slika 5.2: Primer trajektorije FicHju-Nagumovog modela na vremenskom intervalu $T = 10$, za korak diskretizacije $h = 0.01$ i za parametre $(\alpha, \varepsilon, \gamma, \sigma) = (1.2, 0.1, 1.3, 0.4)$.

U oba eksperimenta, posmatraćemo dva seta parametara

$$\theta_1 = (\alpha_1, \varepsilon_1, \gamma_1, \sigma_1) = (0.3, 0.1, 1.5, 0.6), \quad (5.37)$$

$$\theta_2 = (\alpha_2, \varepsilon_2, \gamma_2, \sigma_2) = (1.2, 0.1, 1.3, 0.4). \quad (5.38)$$

Model koji odgovara parametru θ_2 će se dosta razlikovati od modela sa parametrom θ_1 jer izborom parametara menjamo ekvilibrijume modela, kao i njihove stabilnosti. Na slici 5.2 možemo videti jednu trajektoriju FicHju-Nagumovog modela za druge vrednosti početnih parametara. Ovaj put smo proširili vremenski interval $T = 10$, ali ni to nije bio dovoljno veliki interval da vidimo jednu frekvenciju oscilacije koordinata. To ne znači da model nije više oscilatoran, već da su oscilacije dosta ređe zbog jake privlačnosti stabilnog ekvilibrijuma. Dakle, da model nije stohastički, dinamički sistem bi iskonvergirao ka stabilnom ekvilibrijumu, ali zbog slučajnog šuma, model izlazi iz ekvilibrijuma. U terminima akcionog potencijala, kada voltaža pređe određeni prag, potencijal naglo opali i pošalje signal drugom neuronu. U modelu sa drugim setom parametara ovaj događaj je mnogo ređi, što više odgovara modelovanju akcionog potencijala.

Prilikom testiranja ćemo proširiti vremenski okvir na $T = nh = 50$ jer imamo uslov da $nh \rightarrow \infty$, kao i da $n \rightarrow \infty$ i $h \rightarrow 0$. Za $h = 0.01$, imamo obim uzorka $n = 5000$. Opet, podatke simuliramo Ojler-Marujamovom metodom za $h = 0.0001$ i uzimamo svaki stoti iz uzorka.

Prvo želimo da proverimo mere testova KS_n i CVM_n zasnovanih na empirijskom procesu reziduala. To radimo pomoću Monte Karlo simulacije. Postavljamo broj iteracija na $nMC = 5000$. U svakoj iteraciji ćemo generisati uzorak iz nulte hipoteze sa pravim vrednostima parametara za oba seta parametara, kao što je gore opisano. Potom ćemo iskoristiti ocenu parametara dobijenih metodom zasnovanom na lokalnoj linearizaciji. Sada ćemo na osnovu tih parametara oformiti ocenjene rezidualne dobijene Ojler-Marujamovom diskretizacijom samo za marginalnu diskretizovanu stohastičku diferencijalnu jednačinu. S obzirom na to da imamo dve koordinate i da jedna nema šum koji se propagira posle Ojler-Marujamove diskretizacije, posmatraćemo rezidualne dobijene samo za hrapavu koordinatu. Ovde nije očigledno da će ovaj pristup raditi jer menjamo diskretizacije za parametre i za rezidualne. Takođe, koristimo samo deo podataka, tj. samo jednu koordinatu umesto obe. Ovo želimo da istražimo. Kada smo dobili ocenjene rezidualne, računamo test statistiku za njih. Koristimo samo $T = nh$ prvih reziduala. Ovde je takođe logičnije koristiti poslednje jer na početku proces nije stabilan, ali pratićemo uputstva iz rada, a napomenuto ćemo testirati u daljim istraživanjima. U svakoj Monte Karlo iteraciji dobićemo jednu vrednost za svaku od dve test statistike. Na kraju ćemo imati dva niza od 100 vrednosti. Ove nizove ćemo porediti sa odgovarajućim asimptotskim raspodelama za KS_n i CVM_n test statistiku. Brojaćemo koliko smo prosečno puta odbacili nultu hipotezu i time ćemo dobiti ocenu mere testa.

U tabeli 5.1 vidimo ocenjene mere testova za različite prave vrednosti parametara. Primećimo da je željena mera testa 0.05 i da su dobijene mere relativno blizu željene. Razlog zašto nismo dobili očekivanu meru verovatno leži u tome što naš vremenski interval $T = nh = 50$ nije dovoljno velik. Ostaje za buduća istraživanja da se proveriti da li se mera poboljšava za veći vremenski interval. Za sad možemo zaključiti da su, s obzirom na okolnosti, ocenjene mere dobrog reda veličine. Takođe, postoji velika razlika između

θ	KS_n	CVM_n
θ_1	0.0836	0.0774
θ_2	0.0415	0.0353

Tabela 5.1: Ocenjene mere KS_n i CVM_n testova Monte Karlo metodom sa $nMC = 5000$ iteracija. Mere su ocenjene za dva seta pravih parametara θ_1 i θ_2 zadatim u jednačinama (5.37) i (5.38).

mera za različite parametre, što ne bi trebalo da se dešava ako je asimptotska raspodela pod nultom hipotezom slobodna od parametara. Opet, očekujemo da je razlog ovome nedovoljan obim posmatranih reziduala T . Možda bi se mere manje razlikovale da smo posmatrali poslednjih T reziduala, umesto prvih.

U drugom eksperimentu ćemo ocenjivati moć testova protiv date alternative. U svakoj iteraciji Monte Karla simuliraćemo podatke iz alternativne raspodele i testirati test statistikom pod nultom hipotezom. Na kraju će prosek odbačenih nultih hipoteza biti ocena moći.

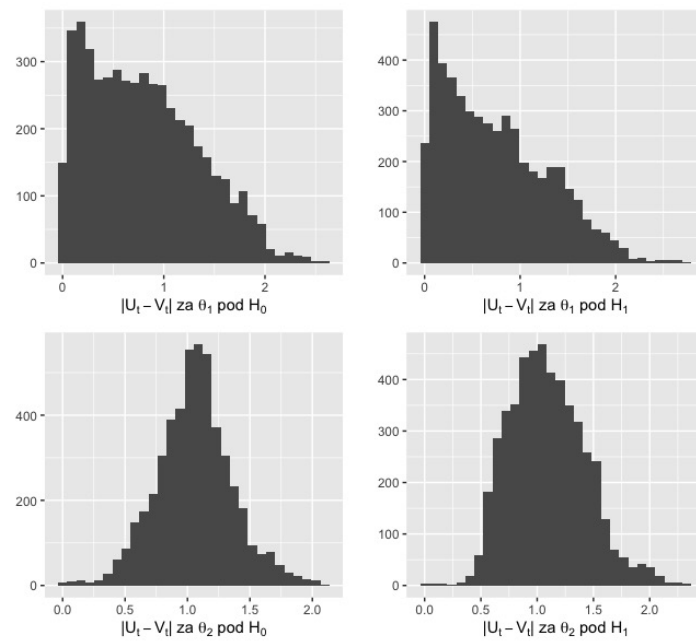
θ	KS_n	CVM_n
θ_1	0.9186	0.9564
θ_2	0.2373	0.2684

Tabela 5.2: Ocenjene moći KS_n i CVM_n testova Monte Karlo metodom sa $nMC = 5000$ iteracija. Moći su ocenjene za dva seta pravih parametara θ_1 (5.37) i θ_2 (5.38) protiv gorepomenutog alternativnog modela H_1 .

U tabeli 5.2 možemo videti da je CVM_n test bolji od KS_n testa u oba slučaja. S druge strane, vidimo da su u prvom slučaju oba testa dovoljno moćna da razlikuju alternativnu hipotezu od nulte, dok u slučaju modela sa stabilnim ekvilibrijumom, ovi testovi nisu pouzdani.

Razlog za to leži u konstrukciji alternativne hipoteze i modela sa stabilnim ekvilibrijumom. Jedina razlika između nulte i alternativne hipoteze jeste što u alternativnoj hipotezi šum množimo sa $\sqrt{|U_t - V_t|}$. U slučaju modela sa stabilnim ekvilibrijumom, ova slučajna veličina u oba modela ima ocenjeno očekivanje oko 1, što znači da se alternativni model ne razlikuje od nultog. Ovo možemo videti na histogramima na slici 5.3. Na slici vidimo apsolutnu razliku između koordinata za jednu trajektoriju procesa simuliranu pod različitim uslovima. U redovima imamo različite početne parametre, dok u kolonama imamo modele iz nulte, odnosno alternativne hipoteze. Vidimo da u slučaju modela sa početnim parametrima θ_2 apsolutna razlika ima prilično simetričnu raspodelu oko vrednosti 1. To znači da je jako teško razlikovati alternativni model od nultog. Dakle, za testiranje moći testa u ovom slučaju moramo koristiti neku bolju alternativu.

Ovime završavamo poglavlje i sam master rad. Ostaje još zaključak gde ćemo ukratko rezimirati sadržaj rada i predložiti dalja istraživanja.



Slika 5.3: Raspodele apsolutne razlike koordinata pod različitim hipotezama i različitim početnim parametrima

Zaključak

U ovom master radu bavili smo se statističkim zaključivanjem u stohastičkim difuzijama. Videli smo da je oblast vrlo kompleksna i zahteva znanje i intuiciju iz raznih grana matematike. Naravno, sve prikazano u radu nije ni približno onome što postoji u ovoj oblasti, ali je zamisao bila da se rad koncipira tako da uvede čitaoca u priču i ukaže na potencijalne probleme i neke od načina da se problemi izbegnu.

Na početku rada smo videli motivaciju iza modelovanja stohastičkim difuzijama, odnosno zašto je važno imati modele sa neprekidnim vremenom koji modeluju određeni šum. Uveli smo primer FicHju-Nagumovog modela koji smo koristili kroz ceo rad za različite tipove simulacija.

Potom smo napravili kratak pregled svih matematičkih pojmova i nekih važnih teorema koje su neophodne da bi se bolje razumelo okruženje u kome radimo. Akcenat je stavljen na numerički deo rešavanja stohastičkih diferencijalnih jednačina, kao i na osnovno razumevanje pojmova iz običnih diferencijalnih jednačina. U delu koji zalazi u koncepte iz verovatnoće i statistike videli smo kako se intuitivno uvode stohastički integrali, a time i stohastičke diferencijalne jednačine, kao i koje sve statističke eksperimente radimo kada se bavimo bilo kakvim modelom.

U trećem poglavlju fokusirali smo se na metode deljenja za stohastičke diferencijalne jednačine. Uverili smo se da su ove metode vrlo uspešne za simulaciju u smislu jednostavnosti, ali i u smislu efikasnosti. Dosta elegantno rešenje koje daje aproksimaciju koja konvergira u srednjem sa redom $p = 1$. Iako ovo nije naročito poseban rezultat, svi ostali rezultati bili su bolji nego u slučaju drugih poznatih metoda diskretizacije. U ovom poglavlju smo se takođe pozabavili i pitanjem hipoeliptičnosti stohastičke diferencijalne jednačine. Videli smo da se ona često bira za modelovanje procesa iz realnog sveta, ali i da nije lako očuvati ovu osobinu posle diskretizacije. Zaključili smo da su diskretizacije na osnovu deljenja najbolje što se tiče simulacije modela.

U nastavku smo najzad započeli statističko zaključivanje. Sama činjenica da ovo nismo mogli na samom početku, dovoljno govori o kompleksnosti problema. Najpre je bilo potrebno uvesti čitaoca u priču i objasniti šta sve postoji i kako od neprekidnog modela doći do diskretizovanog uzorka. Videli smo da i posle celog puta koji smo prešli i dalje nije trivijalno oceniti parametre modela. Ocene parametara u velikoj meri zavise od diskretizacije, zbog čega smo dosta vremena proveli na prethodnom poglavlju. Videli smo

nekoliko načina ocenjivanja. Prvo smo pomenuli način ocenjivanja koji su predložile S. Ditlevsen i A. Samson koje i pored toga što su koristile izrazito kompleksnu diskretizaciju, nisu uspele da ocene parametre odjednom. Razlog tome jeste u hipoeleptičnosti modela i činjenici da se šum propagira sa različitim redovima veličine na glatke, odnosno hrapave koordinate. Posle toga smo se upoznali sa načinom ocenjivanja predloženim od strane A. Melnjikove. Ona je uspela da oceni sve parametre zajedno, bez razdvajanja, i da pokaže njihovu postojanost i asimptotsku normalnost. Takođe je rešila i problem koji se javio kod S. Ditlevsen i A. Samson, a to je diskretizacija. Njen način diskretizacije je dosta intuitivniji i lakše se implementira, a dovodi do identičnog rezultata. Dakle, diskretizacija zasnovana na lokalnoj linearizaciji preovladava u smislu ocene parametara. Na kraju smo predložili dve ocene zasnovane na Li-Troterovoj diskretizaciji. Ove ocene, iako nisu bolje od prethodno predložene, ipak imaju određenih prednosti. Pre svega, ocena zasnovana na Li-Troterovoj aproksimaciji koristi nelinearne funkcije od parametara i koraka diskretizacije h , što znači da se očekuje da ove ocene rade bolje u slučaju kada je korak diskretizacije dovoljno velik da Tejlorska aproksimacija ne daje dobre rezultate. Sa druge strane, ako bismo Tejlorski aproksimirali funkciju drifta i difuzije dobijene iz Li-Troterove aproksimacije samo do prvog stepena, tada bismo dobili ekvivalent Ojler-Marujamove ocene kada bi stohastička diferencijalna jednačina bila eliptična. U slučaju kada su parametri linearni, tada ih možemo oceniti bez numeričkog rešavanja optimizacionog problema. Ovakve ocene neće biti nepristrasne, ali su jako efikasne što se tiče kompjuterske brzine i memorije, tako da u trenucima kada nam je to jako važno, možemo grubo oceniti parametre na ovaj način.

Za kraj, u petom poglavlju smo se bavili problemom testiranja saglasnosti modela sa podacima. Ovaj problem je izrazito važan u statističkom zaključivanju jer ocene parametara nisu važne ako ne možemo verovati modelu. Videli smo, opet, da je to još jedno otvoreno pitanje u nizu pitanja iz ove oblasti. Pomenuli smo dva načina testiranja saglasnosti, oba bazirana na jednodimenzionalnim difuzijama. Prvi test koji smo prikazali jedan je od prvih koji je asimptotski slobodan od raspodele i koji se zasniva na diskretizovanom uzorku. Doduše, mana ovog testa je što postavlja jake pretpostavke o modelu. Iz tog razloga nismo se bavili numeričkim eksperimentima za isti. Drugi test možemo okarakterisati kao prilično interesantan. Objavljen je u radu koji se bavi empirijskim rezidualima i pokazuje da se ukoliko se žrtvuje većina reziduala, može konstruisati asimptotski test slobodan od raspodele pod nultom hipotezom. Ovaj test koristi samo mali procenat svih reziduala i ne zavisi od ocenjenih parametara. Za njega smo izveli određene numeričke simulacije. Za kasnija istraživanja ostaju otvorena sledeća pitanja. Prvo, kako možemo uopštiti ovaj test na više koordinata? Zatim, kako uopštiti test da koristi druge tipove diskretizacije umesto Ojler-Marujamove? Što se numeričkih istraživanja tiče, bilo bi dobro istražiti kako se test ponaša pri većem vremenskom intervalu $T = nh$, kao i da li postoji razlika u korišćenju različitih reziduala za konstrukciju test statistike. Za kraj, može se ispitati i asimptotska sloboda od raspodele pri nultoj hipotezi pomoću butstrep metode.

Ovim zaključujemo ovaj master rad uz nadu i uverenja da ćemo u narednom periodu dobiti odgovore na navedena pitanja.

Literatura

- [1] Sergio Blanes, Fernando Casas, and Ander Murua. Splitting and composition methods in the numerical integration of differential equations. *Bol. Soc. Esp. Mat. Apl.*, 45, 01 2009.
- [2] Evelyn Buckwar, Adeline Samson, Massimiliano Tamborrino, and Irene Tubikanec. Splitting methods for sdes with locally lipschitz drift. an illustration on the fitzhugh-nagumo model, 2021.
- [3] Susanne Ditlevsen and Adeline Samson. Hypoelliptic diffusions: discretization, filtering and inference from complete and partial observations, 2018.
- [4] W. Feller. On the kolmogorov–smirnov limit theorems for empirical distributions. *Annals of Mathematical Statistics*, 19:735–749, 1948.
- [5] Wulfram Gerstner, Werner M. Kistler, Richard Naud, and Liam Paninski. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, USA, 2014.
- [6] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations; 2nd ed.* Springer, Dordrecht, 2006.
- [7] Mathieu Kessler. Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, 24:211–229, 1997.
- [8] Rafail Khasminskii. *Stochastic Stability of Differential Equations*. Mechanics of structural systems. Springer Netherlands, 2012.
- [9] Peter Kloeden and Eckhard Platen. *The Numerical Solution of Stochastic Differential Equations*, volume 23. 01 2011.
- [10] P. Lax and R. D. Richtmyer. Survey of the stability of linear finite difference equations. *Communications on Pure and Applied Mathematics*, 9:267–293, 1956.
- [11] Anna Melnykova. Parametric inference for hypoelliptic ergodic diffusions with full observations. *Statistical Inference for Stochastic Processes*, 23:595–635, October 2020.

- [12] Ilia Negri and Yoichi Nishiyama. Goodness of fit test for small diffusions by discrete observations. *Annals of the Institute of Statistical Mathematics*, 63:211–225, 04 2011.
- [13] Lee Sangyeol. On the goodness of fit test for discretely observed sample form diffusion processes: Divergence measure approach. *Journal of the Korean Mathematical Society*, 47(6):1137–1146, 11 2010.
- [14] Lee Sangyeol and Wee In-Suk. Residual empirical process for diffusion processes. *Journal of the Korean Mathematical Society*, 45(3):683–693, 05 2008.

Biografija

Predrag Pilipović rođen je 4. februara 1997. u Novom Sadu. Do svoje šeste godine živeo je u Bačkoj Palanci, a potom se preselio u Beograd. Osnovnu školu Pavle Savić pohađao je do šestog razreda, a zatim se upisao u Osnovnu školu pri Matematičkoj gimnaziji u kojoj je stekao i diplomu srednje škole. Godine 2016. upisuje Matematički fakultet na modulu statistika, aktuarska i finansijska matematika. Na istom je sa prosečnom ocenom 9,05 diplomirao za tri, umesto za programom predviđene četiri godine. U poslednjoj godini studija obavio je praksu u Rajfajzen banci u okviru tima za modelovanje. Prvu diplomu master studija stekao je od strane Univerziteta u Grenoblu (eng. *University of Grenoble Alpes*) i Politehničkog Instituta u Grenoblu (eng. *Grenoble Institute of Technology*), u Francuskoj, na smeru za primenjenu i industrijsku matematiku. Master rad na temu *Coupled Audio-Visual Multi-Speaker Tracking* napisao je pod mentorstvom profesora Havijera Alamede-Pinede kao deo prakse u Nacionalnom institutu za istraživanja u digitalnim naukama i tehnologijama INRIA (eng. *National Institute for Research in Digital Science and Technology*). Ovaj master program bio je finansiran stipendijom Multidisciplinarnog instituta za veštačku inteligenciju (eng. *Multidisciplinary Institute in Artificial Intelligence*). Godine 2020. upisuje master studije na Matematičkom fakultetu, Univerziteta u Beogradu, dok paralelno dobija posao naučnog saradnika na Matematičkom fakultetu, Univerziteta u Kopenhagenu. U Kopenhagenu je prvi put započeo izučavanje stohastičkih difuzija. U septembru 2021. godine započinje doktorske studije na temu *Multivariate responses, dependence structures and partial observations* na istom univerzitetu pod mentorstvom profesorke Suzane Ditlevsen. Navedeni program doktorskih studija deo je Evropskog projekta *Economic Policy in Complex Environments*¹ (EPOC) pod pokroviteljstvom stipendije Marije Kiri. Ukoliko uspešno odbrani doktorat, Predrag će steći dvojni diplomu Univerziteta u Kopenhagenu i Univerziteta u Bilefeldu.

¹<https://epoc-itn.eu>