

SVEUČILIŠTE U RIJECI

TEHNIČKI FAKULTET

Preddiplomski sveučilišni studij elektrotehnike

Završni rad

AUTOREGRESIJSKI MODELI U OBRADI SIGNALA

Rijeka, 2. kolovoza 2020.

Denis Mijolović
0069066432

SVEUČILIŠTE U RIJECI

TEHNIČKI FAKULTET

Preddiplomski sveučilišni studij elektrotehnike

Završni rad

AUTOREGRESIJSKI MODELI U OBRADI SIGNALA

Mentor: doc. dr. sc. Ivan Dražić

Komentor: prof. dr. sc. Viktor Sučić

Rijeka, 2. kolovoza 2020.

Denis Mijolović
0069066432

Na mjesto ovo stranice je potrebno umetnunti izvornik zadatka

IZJAVA

Sukladno članku 8. Pravilnika o završnom radu, završnom ispitu i završetku prediplomskih sveučilišnih studija/stručnih studija Tehničkog fakulteta Sveučilišta u Rijeci od 1. veljače 2020., izjavljujem da sam samostalno izradio završni rad prema zadatku preuzetom dana 18. ožujka 2019. godine.

Rijeka, 2. kolovoza 2020.

Denis Mijolović

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Mauris et ante a mauris consequat tempor. Donec vulputate, nulla pharetra finibus finibus, ipsum ante molestie mauris, in tincidunt libero augue sit amet arcu. In hac habitasse platea dictumst. Nulla sagittis ante ac augue volutpat cursus. Ut gravida odio malesuada, vestibulum mi at, dignissim leo. Morbi semper finibus est, id maximus mauris auctor eget.

Sadržaj

1. Uvod	3
1.1. Važnost prognoziranja u inženjerstvu	5
2. Teorijske osnove signala i sustava	8
2.1. Vrste signala	8
2.2. Vrste sustava	9
2.2.1. Svojstvo kauzalnosti sustava	10
3. Opće metodologije prognoziranja	12
3.1. Subjektivne prognoze	12
3.2. Prognoze temeljene na statističkim modelima	12
3.2.1. Procjena maksimalne vrijednosti	13
3.2.2. Metoda najmanjih kvadrata	15
4. Osnovni statistički pokazatelji signala i vremenski nizovi	18
4.1. Vremenski nizovi	18
4.1.1. Primjeri analize vremenskih nizova u elektrotehnici	19
4.2. Osnovni statistički pokazatelji	21
4.3. Stohastički procesi	22
4.4. Bijeli šum	23
5. Autoregresivni modeli	24
5.1. Autoregresijski model (AR)	24
5.1.1. Dodatne pretpostavke	27
5.1.2. Metode određivanja vrijednosti nepoznatih koeficijenata	27
5.1.3. Autoregresijski model 1. reda	28
5.1.4. Autoregresijski model 2. reda	28
5.2. Pomična srednja vrijednost (MA)	28
5.3. ARMA (p,q) proces	28
5.4. ARIMA proces	28
6. Pokus	29
7. Zaključak	30

Bibliografija

31

1. Uvod

(Definicija podataka i informacija na zanimljiv način...ako je takvo što moguće)

Vratimo li se u prošlost, važnost podataka i njihova interpretacija u svrhu stvaranja informacija bila je neminovna za civilizacijske strukture koje danas poznajemo. Od početka razvoja primitivnih tehnologija i njezinoga korištenja za optimizaciju životnih procesa, do razvoja trgovine i financijskih instrumenata; čovjek je sve efikasnije koračao prema evolucijskom vrhu uspješnom utilizacijom svoje okoline. Daljnjim razvojem, čovjek je nadišao barijeru vlastite memorije za obradu podataka, te se počeo približavati točki zasićenja - uslijed koje se javila potreba za stvaranjem boljeg i efikasnijega načina obrade podataka i njihove pretvorbe u sažetiju i jednako kvalitetnu informaciju.

Kao i kod većine velikih znanstvenih otkrića koja su obilježila svoje razdoblje kao pretkretnice stoljeća u ratnim okolnostima, tako je i tijekom Drugog svjetskog rata otkriven novi izum kao odgovor na Enigmu - stroj koji je kodirao njemačke vojne i diplomatske poruke te čija se šifra smatrala nerješivom. Uzevši u obzir širenje nacističke Njemačke i prijetnju koja je prijetila ljudskim slobodama, javila se jasna potreba za stvaranjem rješenja koje će nadići čovjekova fizička i psihička ograničenja te biti u mogućnosti procesuirati veliku količinu podataka u vrlo kratkom vremenu – radeći neprestano i ne osjećajući umor te, konzekventno, ne stvarajući greške. Imajući to na umu, Englezi su tada izumili *Colossus* – uređaj koji je pomogao kriptanalitičarima da dešifriraju njemačke poruke i dao im stratešku prednost pri prognoziranju daljnjih koraka tijekom ratnog razdoblja. Taj uređaj se u današnje vrijeme smatra pretečom prvoga računala kakvim ga danas poznajemo.

Baš kao i tada, te kroz cijelu ljudsku povijest, ispravna interpretacija podataka u informaciju koju krajnji korisnik može iskoristiti u svrhu kvalitetnijeg donošenja važnih odluka, iskazala se kao značajan faktor u očuvanju civilizacijske i sveopće stabilnosti – stanja koje je najpotentnije za budući razvoj, a time i blagostanje. Kako je civilizacija težila ka tehnološkom društvu, tako je eksponencijalno rasla i količina podataka koju je bilo potrebno interpretirati; te se javila potreba za sve preciznijim prognoziranjem budućega ponašanja sustava kako ne bi došlo do destrukcije vitalnih procesa i porasta volatilnosti sustava - a samim time i povećanjem vjerojatnosti za njegovim urušavanjem.

Slijedno tome, važnost prognoziranja proizlazi iz dvije osnovne činjenice:

1. budućnost je nepredvidiva
2. poduzete akcije u vrijeme donošenja odluke vrlo često nemaju trenutne posljedice sve do određenoga trenutka u budućnosti

Imajući na umu prethodno navedene činjenice, slijedno je zaključiti da precizne prognoze budućih događaja pospješuju efikasnost postupka donošenja odluka. Štoviše, većina odluka od posebnog poslovnog ili političkog značaja imaju nužan uvjet pozitivnih prognostičkih indikatora prezentiranih kroz studije izvedivosti ili ispitivanja javnoga mnijenja – čija je zadaća što preciznije prognozirati buduću potražnju za proizvodima i uslugama koje su predmet proizvodnih, industrijskih ili inih procesa. U konkretnijem, mikroekonomskom smislu, dalo bi se konstatirati da je prognoziranje implementirano u gotovo sve odluke koje su dio naše svakodnevice: poduzeće će početi izgradnju nove proizvodne jedinice kako bi osiguralo rastuću potražnju na tržištu; zaposleni radnici početi će štedjeti kako bi si mogli priuštiti godišni odmor ili kupnju vrijednosnica s ciljem ostvarivanja prava na buduću kapitalnu dobit; rektor sveučilišta donesti će odluku o otvaranju novog studijskog programa kako bi povećao broj studenata i uskladio broj obrazovanih stručnjaka s budućim tržišnim potrebama. Ukratko, svaki od navedenih procesa zahtjeva inicijalnu percepciju utjecaja donešenih odluka na povećanje vjerojatnosti ostvarivanja zadanih ciljeva – i to prije donošenja samih odluka.

Ostanemo li na mikroekonomskoj razini, možemo reći da je poslovanje poduzeća zahvaćeno trima faktorima: makroekonomskim prilikama, industrijskim pokazateljima i samim uređenjem poduzeća. Obzirom da direktor takvoga poduzeća najčešće ima mogućnost direktnoga djelovanja jedino na poslijednja dva faktora, od iznimne je važnosti da bude što bolje informiran o vanjskim utjecajima i trendovima na koje ne može izravno djelovati – poput donošenja novih zakonskih regulativa ili razvoja globalne virusne pandemije. Važnost ispravnog prognoziranja čimbenika na koje se nema direktan utjecaj od iznimne je važnosti za većinu projekta iznimno velike vrijednosti. Tako, na primjer, projekti s velikim početnim ulaganjima poput izgradnje novih jedinica za opskrbu električnom energijom, mogu imati vrijeme izrade studije izvedivosti u trajanju od 10 ili više godina prije nego što li se započnu radovi. Budući da opskrba električnom energijom ovisi o stupnju razvijenosti regije ili područja, potrebno je izraditi valjane prognoze koje će pomoći u procjeni trenutne, kratkoročne i dugoročne isplativosti novih opskrbnih jedinica za opskrbljivača, ostale projektne partnere, te za lokalni i regijalni razvoj. U tom slučaju, prisutnost raznovrsnih rizika utječe na sve dionike projekta – od nacionalne, regionalne ili lokalne uprave ili samouprave; do svih industrijskih poduzeća koja se nalaze na tom području. Ti rizici mogu biti smanjenje likvidnosti tržišta, kreditni status¹, kamatni rizik², valutni rizik³, strateški rizik⁴, stanje tržišta rada i još mnogi drugi. Pritom je važno napomenuti važnost osjetljivosti modela prognoziranja na fluktuacije u svakoj stavci procjene ukupnoga rizika, tj. isti prognostički model se ne bi trebao koristiti za prognoziranje u različitim zakonskim i regulativnim okvirima – što pridonosi kompleksnosti izrade samoga modela i porastu troškova izrade kvalitetne studije.

¹Usporedba potraživanja i dugovanja nekoga poduzeća koje se dalje koriste za utvrđivanje kreditnog rejtinga.

²Moguće smanjenje vrijednosti financijskih instrumenata javno dostupnoga poduzeća uslijed promjene razine kamatnih stopa na tržištu.

³Osjetljivost financijskih instrumenata na fluktuacije tečajeva stranih valuta. Značajno za multinacionalna poduzeća s viševalutnim potraživanjima i dugovanjima.

⁴Rizik za prihode ili kapital koji nastaje kao posljedica neadekvatnih poslovnih odluka ili nepravilnog vođenja poduzeća.

1.1. Važnost prognoziranja u inženjerstvu

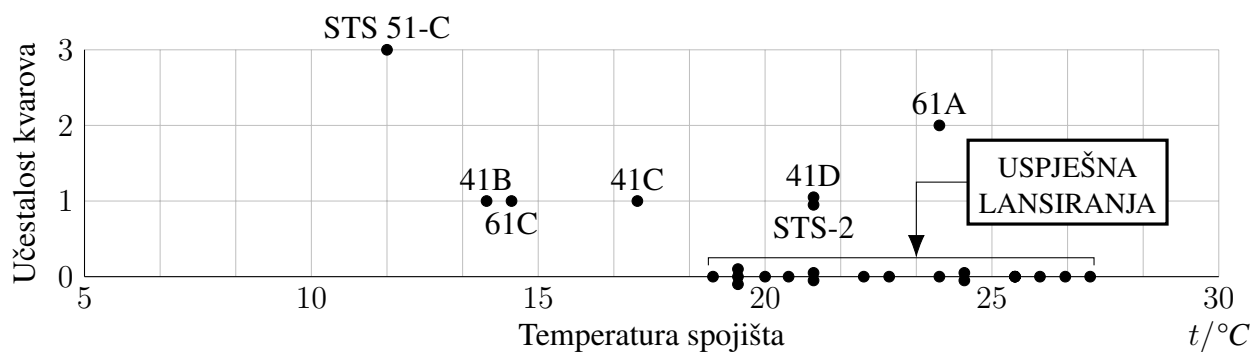
U navedenim primjerima nastojali smo nešto konkretnije opisati važnost prognoziranja kroz opće primjere iz prakse. Kako bismo zaokružili sadržaj ovoga poglavlja, osvrnuti ćemo se na stvarni primjer jedne od najutjecajnijih nesreća koje su obilježile prošlo stoljeće – no ovoga puta fokusirajući se na ispravnoj interpretaciji prognostičkih procesa i važnosti najnasumičnijeg čimbenika u cijelome procesu odlučivanja, tj. čovjeka.

28. siječnja 1986. godine, svemirska letjelica *Challenger* eksplodirala je otprilike 70 sekundi nakon polijetanja iz svemirskog centra John F. Kennedy u Floridi. Sedmero astronauta je poginulo, a svemirska letjelica u potpunosti je uništena. Uzrok nesreće bila je eksplozija spremnika za raketno gorivo, prouzrokovana zapaljenjem plina u pomoćnim raketama za uzlijetanje.

Pomoćne rakete za uzlijetanje bile su predmet zabrinutosti još pri samim počecima razvoja aeronautike. Proizvodnja istih realizira se spajanjem više manjih segmenata uz veliki broj spojišta koji moraju zadovoljavati niz projektom predodređenih uvjeta – među kojima je postizanje nepropusnosti spojišta brtvljenjem. U ovom konkretnom slučaju, korištene su meke profilne brtve s elastičnim deformacijama (također poznate kao o-prsteni) uz sloj ljepila. Prilikom zapaljenja raketnog motora, u sustavu se stvaraju visoke temperature i visoki tlak – što utječe na otapanje ljepila i posljedično habanje brtve, te propuštanja eksplozivnog medija.

Nakon eksplozije, istražni stožer objavio je izvješće u kojemu su navedene okolnosti koje su dovele do uzroka nesreće. Utvrđeno je kako je prethodne večeri, 27. siječnja, donesena odluka o lansiranju letjelice sljedećega dana usprkos nepovoljnoj prognozi temperaturnih uvjeta određenih radnom temperaturom deklariranom od strane proizvođača pomoćnih raketa. Protivno apelu inženjera za obustavom lansiranja, odluka uprave bila je da će se lansiranje ipak provesti, što je rezultiralo kobnim posljedicama.

Kako bi izradili model vjerojatnosti zakazivanja brtvi, koristiti ćemo podatke o učestalosti kvarova brtvi prilikom prethodnih lansiranja – budući da je lansiranje *Challenger*-a bilo 24. lansiranje aktualnog svemirskog programa. Budući da su za svaku pomoćnu raketu korištena tri o-prstena te da su za dotadašnja lansiranja korištene dvije pomoćne rakete po lansiranju, razmatrati ćemo korištenje 6 o-prstena po lansiranju. Uslijed zavisnosti karakteristika brtve o temperaturi, razmotriti ćemo ovisnost zakazivanja brtve o temperaturi koline prilikom lansiranja. Slika 1.1 prikazuje podatke o učestalosti kvarova brtvi pri određenim temperaturama prilikom testnih lansiranja. Iz navedene uočavamo veću učestalost kvarova pri nižim temperaturama – što odgovara svojstvu amorfnih materijala, koji na temperaturama nižim od temperature prelaska u staklo imaju veću tvrdoću i lomljivost.



Slika 1.1. Podaci o učestalosti kvarova brtvi prilikom testnih lansiranja koja su prethodila Challenger-u [5]

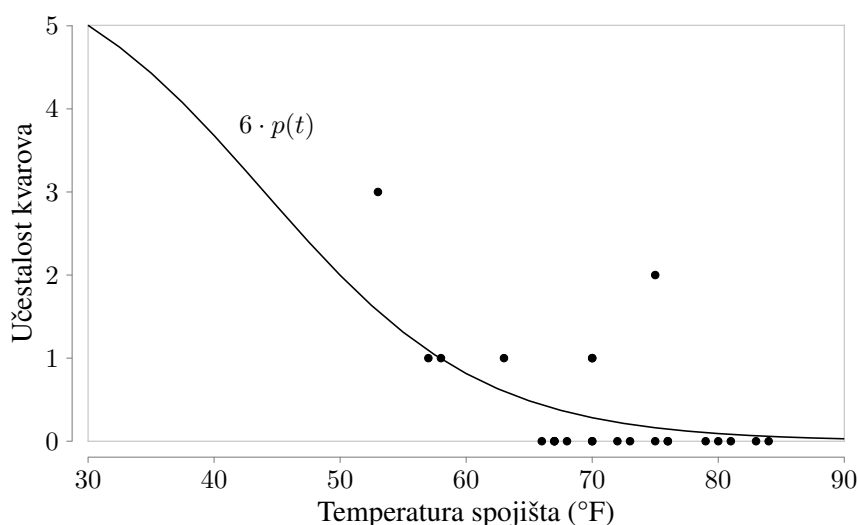
Model prikazan jednačbom (1.1) reprezentira vjerojatnost kvara $p(t)$ pojedine brtve u ovisnosti o temperaturi spojišta t prilikom lansiranja.

$$p(t) = \frac{e^{a+b \cdot t}}{1 + e^{a+b \cdot t}} \quad (1.1)$$

gdje su:

- t temperatura brtve izražena u $^{\circ}F$
- a, b konstante dobivene statističkim modeliranjem metodom procjene maksimalne vrijednosti

Vrijednost konstanti a i b određene su podacima, tj. odabirom vrijednosti a i b tako da dobijemo što točniju aproksimaciju u okolini podatkovnih točaka prikazanih na slici 1.1. Budući da sinteza navedenoga statističkoga modela ne obuhvaća temu ovoga rada, nećemo analizirati statistički izračun konstanti a i b za slučaj Celzijeve temperature, već ćemo koristiti postejeće konstante iz literature ($a = 5.085$ i $b = -0.1156$)[3] u svrhu izrade grafičkog prikaza vjerojatnosti kvara brtvi pri određenim temperaturama – kao što je prikazano na slici 1.2.



Slika 1.2. Grafički prikaz modela (1.1) u odnosu na diskretne vrijednosti mjerenih podataka [3]

Valjano je, temeljem prethodnih primjera, zaključiti da uspješnost ispunjavanja zadanih ciljeva uveliko ovisi o mogućnostima glavnih aktera da što brže i točnije predvide posljedice kako bi se što bolje pripremili za daljnje korake. Pouzdane prognoze upravo to i omogućuju – da se donesu pravovremene odluke koje su temeljene na valjanim planovima. U poglavljima koja slijede, detaljnije ćemo definirati određene pojmove, pomoću kojih ćemo biti u mogućnosti lakše prevesti prognoziranje iz lingvističke apstrakcije u jezik matematike – najprecizniji i najrašireniji jezik koji čovjek poznaje.

2. Teorijske osnove signala i sustava

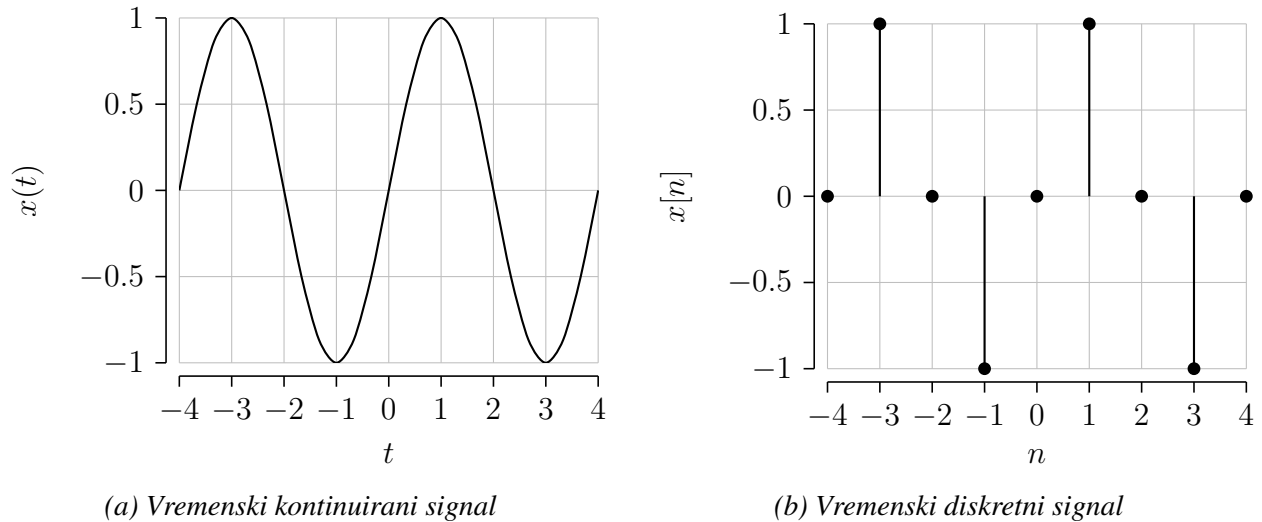
Prije matematičkoga opisa prognostičkih modela, potrebno je pobliže opisati pojam *signala* kao realnu ili kompleksnu funkciju vremena. U elektrotehnici je takav opis signala od iznimne važnosti, pritom posebnu pažnju pridodajemo sinusoidnim signalima zbog karakterističnih svojstva takvih signala. Ta karakteristična svojstva pokazala su se od iznimne koristi za modeliranje fizikalnih pojava pomoću aproksimacije praktičih rezultata dobivenih pripadajućim mjernim metodama. Osim u elektrotehnici, signali su sveprisutni u gotovo svim znanstvenim disciplinama i granama. U širem smislu, valjano je reći da je signal skup podataka ili informacija, pa tako možemo konstatirati da su neki od primjera signala struja i napon, sila i brzina, tlak i protok, govor, video ili burzovni indeksi. Usprkos svojoj sveprisutnosti, signali su u suštini zadržali svoju nominalnu jednostavnost – a to je činjenica da je signal u suštini funkcija jedne ili više nezavisnih varijabli koje opisuju ponašanje promatranih fenomena.

2.1. Vrste signala

Prema prirodi vremenske varijable, dvije najosnovnije vrste signala su:

1. *vremenski kontinuirani signali*: $x(t), t \in \mathcal{R}$,
2. *vremenski diskretni signali*: $x[n], n \in \mathcal{Z}$.

U *vremenski kontinuiranim signalima* vremenska konstanta t je kontinuirana i njezina domena, što rezultira i kontinuiranim vrijednostima slike promatrane funkcije. Navedeno se može grafički predočiti pomoću glatke neprekidne linije kao što je prikazano na Slici 2.1a koja prikazuje vremenski kontinuirani signal $x(t) = \sin(\frac{n\pi}{2})$. Za razliku od vremenski kontinuiranih signala, u *vremenski diskretnim signalima* vremenska konstanta je definirana u točno određenim – diskretnim – vremenima. Posljedično, vrijednosti slike promatrane funkcije biti će poznate samo u prethodno definiranim diskretnim vremenima. Primjer vremenski diskretnog signala $x[n] = \sin(\frac{n\pi}{2})$ grafički je prikazan na Slici 2.1b.



Slika 2.1. Prikaz vremenski kontinuiranoga i vremenski diskretnoga signala funkcije $\sin(\frac{n\pi}{2})$

Neki od primjera vremenski kontinuiranih signala su brzina, pozicija vozila, govor ili zvuk iz audio sustava; dok su primjeri vremenski diskretnih signala mjesečne vrijednosti dionica, demografski pokazatelji poput nataliteta i sl..

Kao što smo već prethodno ustanovili, vremenski diskretni signali mogu reprezentirati bilo koje promatrani fenomen čije se vrijednosti mogu dobiti za točno određeni trenutak – što nas vodi do zaključka o postojanju mogućnosti prikazivanja vremenski kontinuiranih signala pomoću vremenski diskretnih signala. To je moguće sukcesivnim *uzorkovanjem* vremenski kontinuiranoga signala (eng. *sampling*), čime dolazimo do diskretnih vrijednosti signala. Rezolucijom uzorkovanog signala moguće je upravljati pomoću podešavanja frekvencije uzorkovanja (eng. *sampling rate*). Tako je prethodno navedeni diskretni signal prikazan na Slici 2.1b načinjen uzorkovanjem kontinuiranoga signala na Slici 2.1a. Navedena klasa signala vrlo je korisna prilikom digitalne obrade signala. Uzmemo li u obzir prethodno spomenutu prednost visoke brzine računalnog procesuiranja podataka, ali i postojeća ograničenja računalnih sustava radnom memorijom, procesorskim performansama, te digitaliziranom reprezentacijom podataka; važnost vremenski diskretnih signala i uzorkovanja je neminovna – što je dokazano širokom i ekonomičnom integracijom funkcionalnosti obrade signala u različita tehnološka i ina rješenja (npr. mogućnost mobilnih uređaja da prepoznaju glasovne naredbe). Naravno, operacija uzorkovanja signala nije toliko jednostavna kao što smo opisali jer kvalitetna izvedba zahtjeva dobro poznavanje spektralne analize – što uključuje znanja iz područja matematičkih transformacija do realnih ograničenja poput preklapanja signala ili prisutnosti šuma. Analogno vrijedi i za rekonstrukciju uzorkovanih signala u vremenski kontinuirane signale.

2.2. Vrste sustava

Uz signale se nerijetko veže i ideja *sustava*, te je stoga potrebno definirati koncept sustava kao spona između ulaznoga i izlaznoga signala. Za razliku od jednostavnosti ideje signala, navedeno će

možda izazvati određene poteškoće – budući da je sustav u općem smislu shvaćen kao cjelokupno i dovršeno projektantsko rješenje (npr. sustav za prijenos električne energije), dok je matematička reprezentacija takvoga sustava predložena kao model sustava. Štoviše, opisujući sustav u užem smislu, i dalje dolazimo do vrlo širokoga polja istraživanja i unutar točno određenoga znanstvenoga područja. Zadržimo li se isključivo na području elektrotehnike, sustavi mogu biti određeni kao sustavi programske podrške, sustavi regulacije, elektronički sustavi, ugradbeni sustavi, mehanički sustavi, elektromehanički sustavi itd.. Valjano je zaključiti da je potrebno odrediti nužne granice idejnoga razmatranja sustava kako bi izbjegli neizbježnu apstrakciju beskonačne kvantizacije opisa sustava. Iz toga proizlazi ideja i potreba za prethodno spomenutim razmatranjem sustava kao poveznice između ulaznoga i izlaznoga signala, što se ispostavilo kao koristan alat inženjerima prilikom analize propagacije ulaznih signala u regulacijskim krugovima ili prilikom prognozirivanja ponašanja sustava na određene pobude tijekom projektiranja samoga sustava.

Matematički, sustavi predstavljaju modele koji reprezentiraju transformaciju ulaznog signala $x(t)$ u izlazni signal $y(t)$. Navedeno se može opisati relacijom $y(t) = Hx(t)$, gdje H predstavlja matematički model transformacije signala – tj. sustav. Analogna relacija vrijedi i za slučaj transformacije vremenski diskretnoga signala $x[n]$, kao što je prikazano na Slici 2.2.

Kao i kod signala, dvije osnovne vrste sustava prema prirodi vremenske varijable su:

1. *Vremenski kontinuirani sustavi*: ulazni signal $x(t)$ i izlazni signal $y(t)$ sustava su vremenski kontinuirani,
2. *Vremenski diskretni sustavi*: ulazni signal $x[n]$ i izlazni signal $y[n]$ su vremenski diskretni.



Slika 2.2. Blokovski prikaz vremenski kontinuiranoga sustava (a) i vremenski diskretnoga sustava (b).

2.2.1. Svojstvo kauzalnosti sustava

Iako signali i sustavi imaju niz matematičkih svojstva koja se koriste u svakodnevnoj primjeni, u svrhu pojašnjenja budićih termina ukratko ćemo pojasniti pojam kauzalnosti sustava.

Sukladno definiciji, sustav je *kauzalan* ukoliko signal koji propagira kroz sustav ovisi isključivo o trenutnim ili prošlim vrijednostima ulaznoga signala. Drugim rječima, izlazni signal ne predviđa vrijednosti odziva na buduće vrijednosti ulaznoga signala. Konzekventno, sustav *nije kauzalan* ukoliko vrijednost izlaznoga signala ovisi o budućoj vrijednosti ulaznoga signala. Primjer

kauzalnog sustava je

$$y[n] = \frac{1}{M} \sum_{k=0}^{M-1} x[n-k] \quad (2.1)$$

dok je

$$y[n] = \sum_{k=-\infty}^n x[n-k] \quad (2.2)$$

primjer sustava koji nije kauzalan.

Iako su kauzalni sustavi vrlo važni zbog prirode realnih rješenja koja aktuiraju temeljem prethodno očitanih vrijednosti, već sada možemo naslutiti važnost nekauzalnih sustava prilikom uočavanja kretanja trendova prilikom prognoziranja ili sinteze sustava regulacije s ciljem odstranjivanja šumova i sl..

3. Opće metodologije prognoziranja

Obzirom na sveprisutnost prognoziranja, baš kao i signala i sustava, klasificirati ćemo dvije osnovne vrste prognoza – subjektivne prognoze i prognoze temeljene na statističkim modelima.[4]

3.1. Subjektivne prognoze

Subjektivne prognoze temelje se na pogađanjima, prethodnim iskustvima i intuiciji. Takve prognoze nemaju predodređena i utvrđena pravila procesuiranja informacija, već se svode na donošenje zaključaka temeljem osobnog mišljenja osobe koja iznosi prognozu. Na primjer, dva će ekonomska analitičara iznesti različite prognoze za ponašanje tržišta i vrijednosti financijskih instrumenata u slučaju kriznih situacija poput pandemije koronavirusa. Dok će jedan analitičar smatrati da će vrijednosti financijskih instrumenata imati negativan trend u nadolazećem kvartalu uslijed restriktivnih mjera koje imaju nepovoljan utjecaj na tržište, drugi će prognozirati rast temeljem povećanog priljeva kapitala od strane povećanog broja ulagača koji prvi puta ulaze na tržište kapitala uslijed straha neiskorištavanja volatilitnosti tržišta za ostvarivanje dugoročnih kapitalnih dobitaka (eng. *fear of missing-out*). Iako ovakve metode prognoziranja uistinu mogu rezultirati donekle točnim prognozama, kao što je to bilo u slučaju tzv. kratkih prodaja (eng. *short-sell*) prilikom kraha američkog tržišta nekretnina ili prilikom razotkrivanja prikrivanja neispravnih rezultata mjerenja dušičnih oksida dizelskih automobila Volkswagen grupe; one su dokazano manje točne od prognoza temeljenih na statističkim modelima.

3.2. Prognoze temeljene na statističkim modelima

Za razliku od subjektivnih prognoza, prognoze koje su temeljene na statističkim modelima prate jasno određeni i sistematizirani kvantitativni obrazac definiran međudnosom utjecajnih varijabli. Navedeni modeli mogu se razvrstati na *kauzalne modele* i *nekauzalne modele*. [4]

Analogno prethodno spomenutim kauzalnim signalima, kauzalni prognostički modeli nastoje detaljno opisati matematičku genezu diskretnih mjernih podataka i pokušati pronaći relaciju koja opisuje determinirani način dobivanja vrijednosti zasebnih varijabli u svakoj točki prošloga vremena. Takvi modeli kreću od premise nužnosti što boljšeg poznavanja prošloga ponašanja kako bi se predvidjela budućnosti. Nadalje, kauzalne prognostičke signale možemo dalje podijeliti na *kvalitativne* – koji daju informaciju o tendenciji rasta ili pada budućih vrijednosti – te *kvantitativne*, koji daju informaciju o brojčanoj vrijednosti rasta ili pada budućih vrijednosti.

U odnosu na kauzalne prognostičke modele, nekauzalni prognostički modeli nastoje predvidjeti buduće iznose vrijednosti podataka bez dubljega razumjevanja geneze prethodno mjerenih podataka; već fokusirajući se na projekciju njihova međusobnog odnosa na budućnost. Takvi mo-

deli imaju prednost ekonomičnosti budući da se ne orijentiraju na detaljna istraživanja ponašanja svakog parametra, što je povoljno za prognoziranje trenda čimbenika niže važnosti. Navedena prednost može se gledati i kao nedostatak, budući da premisa o istovjetnosti ponašanja vrijednosti signala u prošlosti i budućnosti ne predviđa pojavu smetnji u sustavu – poput pojave prenapona mreže, kratkog spoja u visokonaponskim energetske sustavima ili pojave velike gospodarske krize.

Naravno, navedeno vrijedi u slučaju jednostavnih prognostičkih modela, te ćemo u nadolazećim poglavljima proučiti detalje kompleksnih nekauzalnih prognostičkih modela temeljenih na vremenskim nizovima. No, prije toga ćemo definirati dvije opće prognostičke metode koje će nam služiti kao osnova budućih razmatranja.

3.2.1. Procjena maksimalne vrijednosti

Procjena maksimalne vrijednosti jedna je od općih metoda prognoziranja koja služi za analizu kompleksnijih sustava koji nemaju jasno raznoljiv način određivanja koeficijenata procjene. Ova metoda čini temelj za razne ostale metode procjene vrijednosti, a najveća joj je prednost primjenjivost u bilo kojem problemu za koji je moguće odrediti zajedničku distribuciju vjerojatnosti temeljem promatranja, tj. mjerenja.

Pretpostavimo li analizu nasumičnog uzorka podataka s n uzorkovanja x_1, \dots, x_n koji je određen distribucijom vjerojatnosti kontinuirane varijable $f(x, \theta)$ sa jednim nepoznatim parametrom θ , izradom vektorske funkcije koja kao članove ima skup uzorkovanih podataka $X = (x_1, \dots, x_n)$ možemo doći do izraza za zajedničku distribuciju vjerojatnosti [1]

$$f(x, \theta) = f(x_1, \theta)f(x_2, \theta) \dots f(x_n, \theta). \quad (3.1)$$

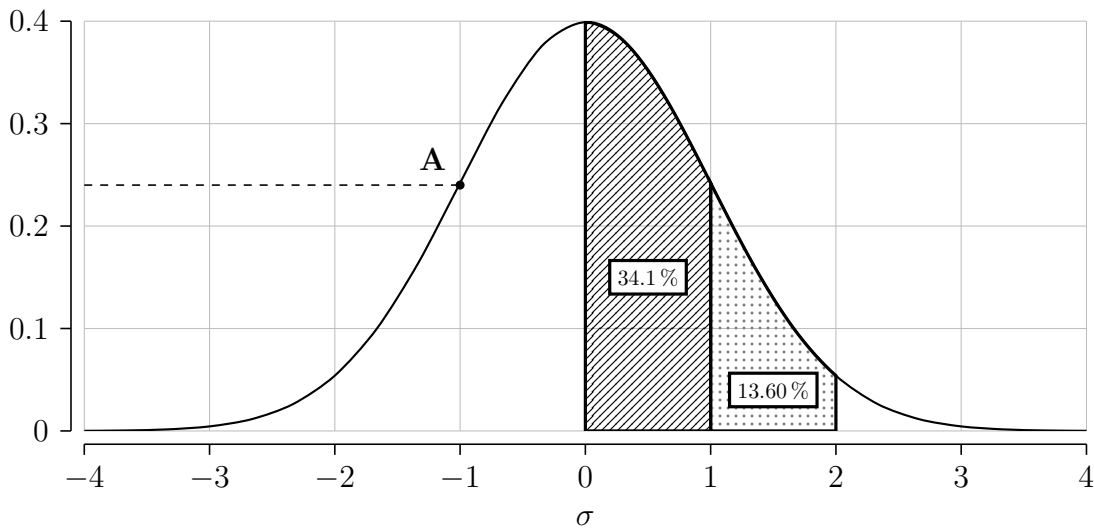
Iz navedenog izraza možemo zaključiti kako je zajednička distribucija vjerojatnosti rezultatni produkt distribucije vjerojatnosti za svakog pojedinog člana x_n . Izraz za zajedničku distribuciju vjerojatnosti $f(x, \theta)$ iz jednadžbe 3.1 koristi se u teoriji vjerojatnosti za procjenu vrijednosti skupa X kada je poznata vrijednost parametra θ . Ukoliko vrijedi

$$f(x_1, \theta) > f(x_2, \theta), \quad (3.2)$$

možemo konvencionalno konstatirati kako je veća vjerojatnost pojavljivanja člana x_1 kao vrijednosti skupa X u nekom određenom trenutku promatranja kada vrijedi ta relacija. Iako navedena konstatacija vrijedi, u prognoziranju je puno učestalije razmatranje vrijednosti θ obzirom na češću primjenu prilikom analize statističke interferencije – koja nastoji opisati ponašanje nepoznatog parametra θ u ovisnosti o različitim ulaznim vrijednostima skupa X koje su nam poznate. Takva instanca funkcije zajedničke distribucije poznata je i kao funkcija vjerodostojnosti.

Kako bi lakše predočili razliku između vjerojatnosti i vjerodostojnosti, pretpostavimo zadani skup podataka s normalnom distribucijom kao što je prikazano na slici 3.1. Vjerojatnost možemo opisati kao površinu ispod krivulje normalne distribucije skupa X . Iz navedenoga, odaberemo li nasumičan član navedenoga skupa, vjerojatnost da je nasumični član skupa unutar granica $\mu \pm 1\sigma$ iznosi 68.27 %.

Za razliku od vjerojatnosti, vjerodostojnost možemo opisati kao vrijednost ordinatnog položaja točke A. Kako bismo znali odrediti prethodno navedeni položaj, potrebno je poznavati srednju vrijednost μ i standardnu devijaciju σ promatranog skupa – temeljem kojih možemo odrediti točku A kao sjecište pravca okomitog na apcisu u točki $\mu \pm \sigma$ i krivulje normalne distribucije skupa X .



Slika 3.1. Prikaz normalne distribucije s naznačenom vjerojatnošću i vjerodostojnošću

Sada kada znamo osnovnu razliku između vjerojatnosti i vjerodostojnosti, promotrimo li funkciju vjerodostojnosti $f(x, \theta)$ uz konstantnu izmjerenu vrijednost x uz relaciju

$$f(x, \theta_1) > f(x, \theta_2), \quad (3.3)$$

dolazimo do zaključka kako se veća vrijednost nepoznatog parametra θ_2 može pripisati kao vjerodostojnija, a samim time je i vjerojatnosti pojavljivanja izmjerenoga člana x kao vrijednosti skupa X veća. Takva razmatranja ne mogu se pripisati *a priori*, već se pretpostavljena vrijednost nepoznatog parametra θ dodatno testira substitucijom u jednadžbu 3.1 – nakon čega se vrši usporedba dobivenih vrijednosti sa izmjerenim vrijednostima iz skupa X [2].

Konačno, slijedi osnovni postulat metode procjene maksimalne vrijednosti koji nalaže postupak procjenjivanja vrijednosti parametra θ i odabir one vrijednosti koja će dati najvjerodostojniji rezultat. U praktičnoj primjeni, češće se koristi logaritamski oblik funkcije vjerodostojnosti

$$L = L(x, \theta) = \log_e \{f(x, \theta)\}. \quad (3.4)$$

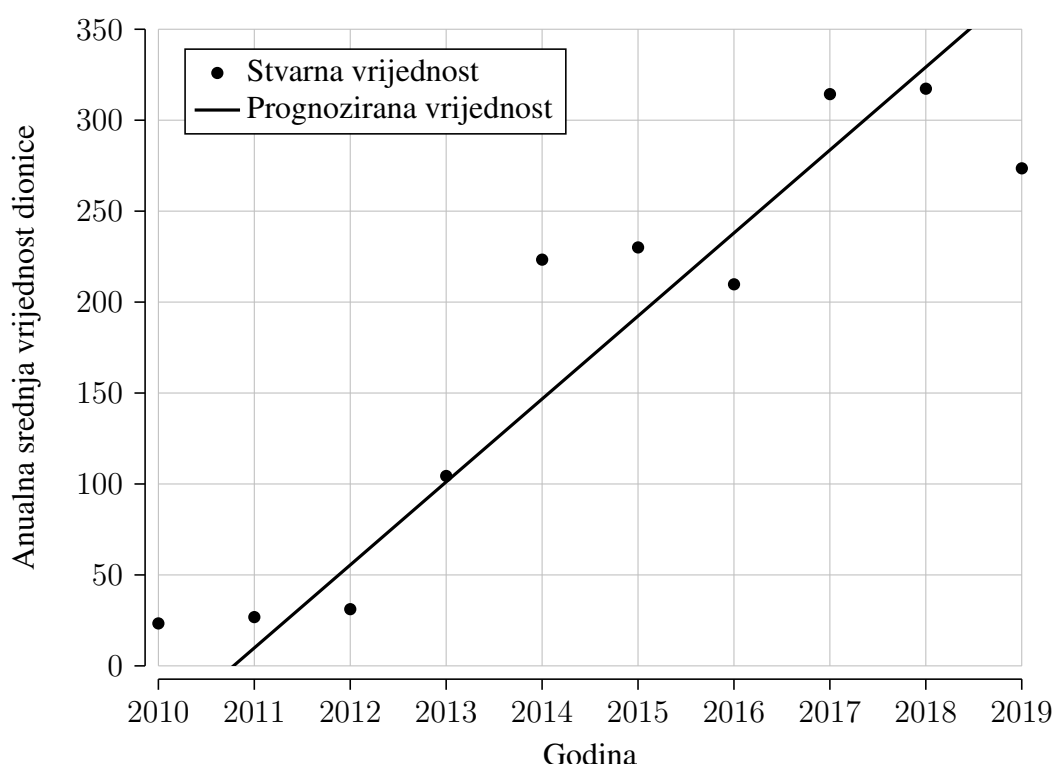
Prethodno stečenim matematičkim znanjem, možemo pretpostaviti proces dobivanja maksimuma funkcije vjerodostojnosti izračunavanjem izraza

$$\frac{\partial L}{\partial \theta} = 0. \quad (3.5)$$

Daljnji proračun nadilazi potrebna znanja za shvaćanje teme ovoga rada te ćemo krenuti s pojašnjenjem sljedeće opće metode prognoziranja – metode najmanjih kvadrata.

3.2.2. Metoda najmanjih kvadrata

Metoda najmanjih kvadrata je jedna od najranijih metoda prognoziranja, no njezina primjena zamjetna je i danas. Cilj korištenja metode najmanjih kvadrata je pronalazak pravca koji minimizira sumu svih kvadratnih grešaka koje nastaju rezultat korelacijskih jednadžbi, kao što je prikazano na slici 3.2. Važno je naglasiti kako prethodno somenuta teza vrijedi isključivo za linearnu metodu najmanjih kvadrata. Uz navedenu, postoji i nelinearna metoda najmanjih kvadrata koju nećemo promatrati.



Slika 3.2. Prikaz stvarne i prognozirane vrijednosti dionica poduzeća "Tesla, Inc" (Izvor: <https://www.macrotrends.net/stocks/charts/TSLA/tesla/stock-price-history>)

Praktičnu primjenu navedene metode je najintuitivnije, poput prethodno spomenute metode procjene maksimalne vrijednosti, predložiti kroz primjenu u procesu investiranja. Uzmemo li primjer investitora koji nastoji donesti odluku o dugoročnom investiranju u poduzeće "Tesla, Inc"

(\$TSLA) s naumom ostvarivanja dobiti kroz idućih 10 godina, korištenjem metode najmanjih kvadrata moguće je izraditi prognostički model koji će pobliže opisati i predočiti trend promjene vrijednosti dionica poduzeća u odnosu na prethodno desetljeće. Rezultat navedene metode jest pravac prikazan na slici 3.2, iz kojega investitor može steći subjektivni dojam o trendu vrijednosti dionice, a koji će mu pomoći pri donošenju odluke o investiranju. Kao i kod otalih metoda, korištenje ove metode može samo pobliže opisati određeni trend u odnosu na referentni vremenski period ili neku drugu nezavisnu varijablu, no u našem konkretnom primjeru, zbog volatilnosti tržišta kapitala nastoji se izbjegavati donošenje konkretnih odluka temeljem metode najmanjih kvadrata uslijed preširokog vremenskog intervala koji može dati krive rezultate prilikom neočekivanih događaja. Kao što je bilo navedeno i kroz prethodni primjer, jedan od neočekivanih događanja može biti i pojava financijske krize – koja se otprilike ponavlja svakih 10 godina. Upravo zbog toga, izrada prognostičkih modela ima više čimbenika i indikatora temeljem kojih se može donesti sud o ispravnosti budućih odluka temeljem dubljih analiza na temelju različitih prognostičkih modela – za koje je nužna logička valjanost korelacije između zavisne i nezavisne varijable. U protivnom, obzirom na proizvoljan odabir zavisne i nezavisne varijable za koju ćemo vršiti promatranja, može doći do logičke neistovjetnosti korelacije i kauzaliteta.

Nakon usustavljivanja osnovne predodžbe metode najmanjih kvadrata, svojstva iste nastojati ćemo opisati kroz analizu nasumičnog mjerenje dostavljenih otpornika, čije rezultate možemo prikazati kao opći izraz

$$E(y) = \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_q x_q, \quad (3.6)$$

gdje su vrijednosti varijabli x_1, x_2, \dots, x_q poznalice (vrijednosti mjerenih otpora), a vrijednosti nepoznatog parametara $\theta_1, \theta_2, \dots, \theta_q$ nepoznanice. Nadalje, izraz 3.6 možemo razvesti kao

$$y = \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_q x_q + \epsilon, \quad (3.7)$$

za koji pretpostavljamo $E(\epsilon) = 0$. Nasumičnu varijablu ϵ nazivamo mjernom pogreškom [1]. Mjerna pogreška ili odstupanje može nastati iz različitih razloga, od kojih su u elektrotehnici najčešći nesavršenosti mjerne opreme, neispravnost mjernog postupka, nevjerodostojno stanje mjernog objekta, vanjski utjecaji, itd.

Izraz 3.6 daje nam linearnu korelaciju vrijednosti y kao sume q članova poznatih vrijednosti x . Želimo li odrediti vrijednost y za točno određeni broj članova, izraz 3.6 možemo zapisati i kao

$$y_i = \theta_1 x_{1i} + \theta_2 x_{2i} + \dots + \theta_q x_{qi} + \epsilon_i, \quad i = 1, \dots, n, \quad (3.8)$$

uz analogno navedenu pretpostavku $E(\epsilon_i) = 0$, te vrijednosti ϵ_i kao mjerne pogreške određenog mjerenja i . Uz sve navedene pretpostavke i poznate vrijednosti podataka, kao jedina nepoz-

nanica nameće se vrijednost nepoznatog parametra θ . "Princip najmanjih kvadrata" nalaže odabir nepoznatih parametara $\theta_1, \dots, \theta_q$ s pripadajućim vrijednostima $\hat{\theta}_1, \dots, \hat{\theta}_q$ koje će minimizirati količinu članova [1]

$$Q(\theta_1, \dots, \theta_q) = \sum_{i=1}^n \{y_i - \theta_1 x_{1i} - \theta_2 x_{2i} - \dots - \theta_q x_{qi}\}^2. \quad (3.9)$$

Korištenjem prethodno stečenih matematičkih znanja, možemo pretpostaviti daljnji pronalazak pripadajućih vrijednosti izračunom minimuma funkcije $Q(\theta_1, \dots, \theta_q)$ parcijalnom derivacijom po svakom θ_i , te izjednačavanjem te parcijalne derivacije s 0 kako slijedi:

$$\frac{\partial Q}{\partial \theta_i} = (-2) \sum_{i=1}^n x_{ji} \{y_i - \theta_1 x_{1i} - \theta_2 x_{2i} - \dots - \theta_q x_{qi}\} = 0, \quad j = 1, \dots, q. \quad (3.10)$$

Jednadžbom 3.10 zaključujemo ovo poglavlje. U sljedećem poglavlju opisati ćemo osnovne statističke pokazatelje signala i vremenskih nizova koji čine osnovu svih statističkih modela, pa tako i onih navedenih u ovom potpoglavlju.

4. Osnovni statistički pokazatelji signala i vremenski nizovi

Kroz prethodna poglavlja ustanovili smo osnovne značajke signala i sustava te prognostike kao discipline koja se bavi predikcijom ponašanja signala temeljem određenih metodologija. Navedene metodologije su kvalitetne koliko i njihova statistička reprezentativnost. Kako bi kvantitativno mogli utvrditi statističku reprezentativnost modela za prognoziranje ponašanja signala, potrebno je izvršiti detaljniju *analizu signala*.

(Dodati Figure 1.9. iz Broersena)

U sklopu studijskog programa svučilišnih studija elektrotehnike, primat kroz više godine studija drži spektralna analiza signala, kojom se nastoji opisati frekvencijske karakteristike promatranog iznalog signala (poput amplitudno-frekvencijske karakteristike, fazno-frekvencijske karakteristike ili pojačanja sustava) Navedeno se najčešće realizira kroz određivanje prijenosnih i inih funkcija sustava koje su od značaja za ponašanje elektroničkih sustava čija su baza elektroničke komponente čije karakteristike nisu nužno konstantne pri određenim frekvencijskim pojavama – što omogućuje korištenje određenih elektroničkih sustava u svrhu niskopropusnoga ili visokopropusnoga filtriranja, te propusnoga filtriranja određenog frekvencijskog opsega.

Drugi pristup analizi signala je analiza ponašanja signala u ovisnosti o vremenu. Štoviše, poznavajući pojam rekonstrukcije vremenski kontinuiranih signala iz vremenski diskretnih signala, dolazimo do zaključka da ponašanje signala u vremenu možemo opisati pomoću beskonačno mnogo diskretnih vrijednosti u infinitezimalno malim vremenskim razmacima, što nas dalje vodi do uvođenja *vremenskih nizova* – koje ćemo detaljnije opisati kroz nadolazeća potpoglavlja.

4.1. Vremenski nizovi

Vremenske nizove povezujemo uz bilo kakavu reprezentaciju vrijednosti u ovisnosti o vremenskoj konstanti – neovisno radi li se o vrijednostima koje prate neko determinističko načelo ili načelo nasumičnosti. Drugim rječima, vremenske nizove možemo opisati kao stohastički signal s kronološki raspoređenim opažanjima uzorkovanim u redovitim vremenskim razmacima. Mogu se pojavljivati kao podaci o fizičkim veličinama, kao ekonomski ili financijski podaci, demografski podaci, itd.. Opažanja ili mjerenja navedenih podataka provedena su u proizvoljnim vremenskim intervalima – svake sekunde, svakoga sata, svakoga dana, svakoga mjeseca, i sl.. U analizi vremenskih nizova, takve skupove podataka uobičajeno reprezentiramo konačnim brojem mjerenja promatranog nasumičnog procesa $X(t)$ za, primjerice, $t = 1, 2, \dots, N$. [1] Vremenske nizove možemo analitički zapisati i na sljedeće načine:

$$X = \{X_1, X_2, \dots\} \quad (4.1)$$

ili

$$Y = \{Y_t : t \in T\}. \quad (4.2)$$

U daljnjim razmatranjima teme ovoga rada, prateći generalni konsenzus u nomenklaturi stručnih izvora, pod analizom vremenskih nizova smatrati ćemo principe i modele koji analiziraju promatrane vrijednosti nasumičnih procesa – koje ćemo detaljnije opisati u nadolazećim potpoglavljima.

4.1.1. Primjeri analize vremenskih nizova u elektrotehnici

Temeljem definicije vremenskog niza, prirodno se nameće korelacija i korištenje istih u ekonomiji i poslovnim sustavima, koji navedene koriste za prognostičko modeliranje u prije spomenute svrhe koje će polučiti konkretni rezultat koji će poslije biti predmet daljnjeg proučavanja.

Korištene metodologije u takvim procesima nisu samostalno svrsihodne u inženjerstvu, tj. nadopunjuju se tehničkom terminologijom i stručnim saznanjima o fenomenima, ograničenjima primjene i fizikalnim pojavama promatranih sustava. Iz toga proizlazi teza o korištenju analize vremenskih nizova kao jednog od alata za daljnje inženjerske analize.

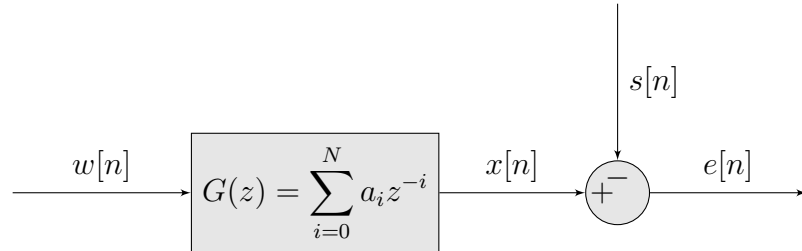
U uvodnom poglavlju naveli smo neke od temeljnih važnosti i značajki prognoiziranja u ostvarivanju tržišnih, strateških i sigurnosnih ciljeva – gdje smo čvrsto ustanovili sveprisutnost potrebe za statističkim modeliranjem, tj. prognoziranjem. Baš kao i u navedenim primjerima ekonomske i upravljačke prirode, prognoziranje i odgovarajuće metodologije od iznimne su koristi današnjim inženjerima pri analizi fizikalnih sustava.

Primjer korištenja istih može se uočiti na primjeru ranoga razvoja modernih vojnih rješenja za protuavionsku obranu. Navedena rješenja proizlašla su kao odgovor na postojeći problem brzorasućeg razvoja ofenzivnih oružja za vrijeme 2. svjetskog rata, što je dovelo do prirodne potrebe za odgovarajućim sustavima zaštite u svrhu očuvanja nacionalne sigurnosti. Razvoj protuavionskih obrambenih sustava temeljio se na pronalasku korelacije koja će s visokom uspješnošću predvidjeti poziciju neautoriziranih objekata u zračnom prostoru u odnosu na trenutnu poziciju obrambenog projektila u svrhu najučinkovitijega nahođenja koje će rezultirati terminacijom potencijalne prijetnje.

Iz procesa pronalaska navedenoga rješenja, proizašao je današnji *Wienerov filter*, koji ima široku primjenu u elektrotehnici i drugim inženjerskim granama – ponajviše prilikom obrade signala, digitalne obrade slike, sinteze sustava upravljanja, te u razvoju digitalnih komunikacija. U općem smislu, koristi je za identifikaciju sustava, dekonvoluciju, odšumljivanje i uočavanje signala.

Temeljem prethodno navedenoga povoda za razvojem Wienerova filtra možemo lakše shvatiti njegovu funkciju, a to je statistički proračun ponašanja nepoznatog signala pomoću referentnog

ulaznog signala koji, propagacijom kroz filter, daje izlaznu (prognoziranu) vrijednost. Točnije, za kauzalne diskretne signale s konačnom vrijednošću Wienerov filter pronalazi optimalne koeficijente pomoću statističkog modeliranja pomoću poznatih prethodnih vrijednosti ulaznog i izlaznog signala – populirajući matricu ulaznih signala s prognoziranim vrijednostima autokorelacije ulaznog signala, dok vrši sličnu radnju populacijom matrice izlaznih signala s prognoziranim vrijednostima unakrsne korelacije između izlaznih i ulaznih signala.



Slika 4.1. Blokovski prikaz FIR Wiener filtra za diskretne signale.

Iz blokovskog prikaza FIR Wiener filtra sa slike 4.1, uočavamo propagaciju diskretnog ulaznog signala $w[n]$ kroz Wienerov filter – kojim se modelira učestalost otipkavanja N i odabir koeficijenata a_0, \dots, a_N . Izlazni signal $x[n]$ iz FIR Wienerova filtra zadan je izrazom

$$x[n] = \sum_{i=0}^N a_i w[n-i]. \quad (4.3)$$

Na temelju zadanoga izraza, možemo uočiti prethodno navedeno korištenje vremenskih nizova kao dio većeg "sustava alata" kojima se inženjeri koriste u svakodnevici. U konkretnom slučaju, FIR Wiener filtra prognozira vrijednosti signala $x[n]$ temeljem prethodnih vrijednosti ulaznog signala $w[n-i]$ uz ugođavanje autokorelacijskih koeficijenata a_i – pritom imajući na umu temeljne pretpostavke i uvjete. Jedan od tih uvjeta je minimizacija srednje kvadrane pogreške koja se koristi kao jedan od analitičkih faktora ocjene kvalitete sustava.

Kako bismo pronašli srednju kvadratnu pogrešku sustava $E[e^2[n]]$, potrebno je definirati signal pogreške $e[n]$. U slučaju blokovskog prikaza sa slike 4.1, signal pogreške definiramo kao razliku vrijednosti izlaznog signala iz FIR Wienerova filtra $x[n]$ i vrijednosti referentnog signala $s[n]$, dolazimo do sljedećih izraza:

$$e[n] = x[n] - s[n], \quad (4.4)$$

$$\begin{aligned} E[e^2[n]] &= E[(x[n] - s[n])^2] \\ &= E[x^2[n]] - 2E[x[n]s[n]] + E[s^2[n]] \\ &= E\left[\left(\sum_{i=0}^N a_i w[n-i]\right)^2\right] - 2E\left[\sum_{i=0}^N a_i w[n-i]s[n]\right] + E[s^2[n]]. \end{aligned} \quad (4.5)$$

Parcijalnim deriviranjem jednadžbe 4.5 po autokorelacijskom koeficijentu a_i , možemo doći do konkretnog izraza za izračun minimuma funkcije srednje kvadratne pogreške $\frac{\partial E[e^2[n]]}{\partial a_i}$. Daljnji proračun tog izraza možemo intuitivno povezati s postupkom minimizacije funkcije metode najmanjih kvadrata opisanoj u potpoglavlju 3.2.2., te ga nije potrebno dalje analizirati.

Zaključno, u ovom potpoglavlju smo potvrdili tezu o korisnosti kroštenja statističke analize vremenskih nizova u elektrotehnici. Jedna od konkretnih primjena Wienerova filtra u elektrotehnici je uočavanje pojave neočekivanoga šuma u električnoj mreži i daljnja regulacija sustava temeljem vrijednosti signala pogreške koji nastaje uslijed pojave navedenoga šuma. Korištenje FIR Wienerova filtra i filtra sličnih funkcija od iznimne je koristi za pravovremenost regulacije sustava upravljanja. U slučaju izostanka pravovremene regulacije sustava upravljanja, može doći do oštećenja sustava i njegovih komponenti, pojave nestabilnosti u sustavu, te drugih nepoželjnih svojstva.

(Dodati neki primjer iz elektrotehnike. U Priestleyu ima predetaljno razrađeni primjer za filtriranje (može se koristiti za dobivanje više stranica, više naginje korištenju vremenskih nizova i nekog općeg nasumičnog signala...ovo trenutno napisano je nekako više "stručno", ali se može staviti kao konkretizacija tog općeg primjera), IEEE Xplore kao dodatni primjeri (uz rastavljanje potpoglavnja))

4.2. Osnovni statistički pokazatelji

varijanca i kovarijanca – potrebno poslije za pojašnjenje stacionarnih i nestacionarnih signala **(uzorak (sample space), srednja vrijednost, varijanca, kovarijanca, stacionarni procesi, auto-kovarijanca...)**

Dekking:

Sample Spaces: Page 13,

Montgomery and Runger:

Collecting Engineering Data: Chap 1. Page 5,

Sample Spaces and Events: Chap 2. Page 17,

Mean and Variance of a Discrete Random Variable: Chap 3. Page 66,

Mean and Variance of a Continuous Random Variable: Chap 4. Page 105,

Data Summary and Display: Page 230,

Priestley:

Probability Spaces: 34,

Random Variables: Page 37,

Means, Variances and Moments: Page 47,

Properties of Means and Variances of Random samples: Page 96,

4.3. Stohastički procesi

Montgomery and Rugner:

Random Variables: Chap. 2 Page 53,

Holden:

Stochastic models of time series: Page 45,

Deterministic and Stochastic Trends: Page 80,

Dekking:

Discrete Random Variables: Page 41,

Continuous Random Variables: Page 57,

Variance: Page 96,

Covariance and Correlation: Page 135,

Broersen:

Random Variables: Page 11,

Functions of Random Variables: Page 18,

Stochastic Processes: Page 29,

Autocorrelation Function: Page 31,

Estimation of Mean and Variance: Page 38,

Autocorrelation Estimation: Page 40,

Priestley:

Random Processes: Page 10,

Stationary Random Processes: Page 14,

Random Variables: Page 37,

Stationary Random Processes: Page 100,

Stationary Processes: Page 104,

The Autocovariance and Autocorrelation Functions: Page 106,

Special Discrete Parameter Models (White Noise): Page 114,

(The words “process” and “model” are often used to describe more or less the same system. Process is generally used to describe the actual process of generating the data, which are also called the observations or signal. Each stationary stochastic data set has one generating process. The word model is more associated with the estimation based on given observations. Different types of models can be calculated for measured data, for which the generating process might be unknown. In the literature, the difference between the words process and

model is not always used very strictly. This does not lead to much confusion because the actual meaning 60 Automatic Autocorrelation and Spectral Analysis follows from the context. Petrus)

4.4. Bijeli šum

Petrus:

White Noise: Page 60,

Priestley:

Purely Random Processes; "White Noise": Page 114,

5. Autoregresivni modeli

Prilikom prognoziranja vrijednosti autokovarijančne funkcije s koeficijentima kašnjenja nailazi se na problem stalnog porasta količine članova funkcije koji su u direktnoj korelaciji s porastom učestalosti uzorkovanja. Takav prikaz za velike količine članova pokazivati će rezultate s velikom preciznošću, no zbog ograničenosti resursa navedeno rješenje nije svrsihodno. Budući da ne postoje konačna saznanja ili koncenzus o konačnom broju koji će reprezentirati najprecizniji prognostički model autokovarijančne funkcije s koeficijentima kašnjenja, u temi ovoga rada razmatrati ćemo prognostičke modele koji kao uvjet imaju neprestan dotok najnovijih podataka uz ograničenje ukupnog broja članova prognostičke funkcije.

Navedene prognostičke modele možemo opisati pomoću triju osnovnih modela stacionarnih stohastičkih procesa [2]:

1. Autoregresijski model (AR),
2. Model pomičnog prosjeka (MA),
3. Mješoviti modeli:
 - a) Autoregresijski model i model pomičnog prosjeka (ARMA),
 - b) Integrirani autoregresijski model i model pomičnog prosjeka (ARIMA).

Kroz nadolazeća potpoglavlja detaljnije ćemo opisati svojstva i način korištenja prethodno navedenih osnovnih modela na nekom skupu podataka – što će se u suštini sastojati od sljedećih postupaka:

1. Izrada procjene *parametara modela*,
2. Određivanje *reda modela*.

Intuitivniji redoslijed tih postupaka bio bi određivanje reda modela, po čemu bi slijedila procjena parametara modela navedenoga reda. Međutim, razvijene metode određivanja reda modela predodređene su inicijalnom izradom procjene parametara modela, te ćemo ih kao takve proučavati poštujući gore navedeni redoslijed postupaka. [1]

5.1. Autoregresijski model (AR)

Za neki određeni statistički model konstatiramo da je autoregresijski ukoliko prognozira buduće vrijednosti temeljem prethodno ostvarenih vrijednosti. Pri analizi financijskih tržišta, autoregresijski model koristi se za predviđanje buduće cijene vrijednosnih papira na temelju njihovih

vrijednosti iz prethodnog razdoblja. Već u samome početku možemo zaključiti da navedeni model ima jedan praktičan nedostatak, a to je određeni stupanj nepouzdanosti temeljem implikativne predodređenosti budućeg ponašanja sustava u ovisnosti o prošlom ponašanju – što je posebice izraženo prilikom nepredvidivih pojava u sustavu. Iako primjera za neočekivane pojave u svakodnevnim sustavima ne nedostaje, najintuitivniji su pojave financijskih kriza ili nagli razvoj određene industrije.

Primjerice, investitor koji koristi autoregresijski model za prognoziranje vrijednosti neke vrijednosnice, pretpostaviti će da su novi kupci i prodavači navedene vrijednosnice prilikom određivanja cijena akvizicije ili prodaje pod psihološki značajnim utjecajem nedavnih događanja na financijskim tržištima. Iako se navedeni model pokazao investitorima kao vjerodostojan, te im omogućio pozitivan priljev financijskih sredstava u godinama koje su prethodile 2008. godini, model se dokazao kao nepouzdan u slučaju kada većina investitora nije bila upoznata s rizikom koji su predstavljali vrijednosni papiri osigurani nekom vrstom hipoteke. Povećanje rizičnosti takvih vrijednosnih papira ponukan je sumnjom nekolicine iskusnih investitora u svjetski renomirane rejting agencije – čija je zadaća bila sa što većom preciznošću evaluirati i ocijeniti vrijednost nekretnina, pokretnina te financijskih sredstava pod hipotekom. Ne ulazeći dublje u praktičnu problematiku, navedena sumnja ispostavila se opravdanom, a dospijećem informacije o neispravnom postupanju prilikom ocjenjivanja uzrokovalo je krah financijskih tržišta i početak nove ekonomske krize 2008. godine. Ukoliko je u novonastalom razdoblju investitor bio predan korištenju autoregresijskog modela, na temelju današnjih saznanja možemo sa velikom sigurnošću konstatirati kako je otvario značajni negativni priljev financijskih sredstava – koji je potencijalno mogao rezultirati i proglašenjem osobnog bankrota. Također, važno je naglasiti snažan utjecaj kratkotrajnih neočekivanih pojava na sve buduće prognozirane vrijednosti autoregresijskog modela, što bi značilo da je navedena financijska kriza iz 2008. godine i dalje "prisutna" prilikom korištenja današnjeg korištenja autoregresijskog modela.

Sada kada smo ustanovili prednosti i mane korištenja autoregresijskog modela u praksi, možemo započeti sa dubinskom matematičkom analizom istoga.

Autoregresijski $AR(p)$ modeli opisuju nasumične stacionarne procese čija izlazna vrijednost u trenutku x_t linearno ovisi o vlastitim vrijednostima koji su prethodili tom trenutku (x_{t-p}). Navedeno opće načelo možemo definirati izrazom

$$\epsilon_t = x_t - \mu_x + \phi_1(x_{t-1} - \mu_x) + \phi_2(x_{t-2} - \mu_x) + \dots + \phi_p(x_{t-p} - \mu_x) + \theta_0, \quad (5.1)$$

gdje su:

ϵ_t	bijeli šum; nasumični stacionarni proces s pripadajućim statističkim parametrima μ_x i σ_x^2
x_t	nezavisna varijabla promatranog skupa
μ_x	očekivana vrijednost slučajne varijable
ϕ_p	autoregresijski parametar
θ_0	konstantni član
p	red autoregresijskog modela

Ukoliko iz izraza 5.1 želimo definirati vrijednost člana procesa x_n , dolazimo do izvedenice

$$(x_t - \mu_x) = \epsilon_t - \phi_1(x_{t-1} - \mu_x) - \phi_2(x_{t-2} - \mu_x) - \dots - \phi_p(x_{t-p} - \mu_x) + \theta_0. \quad (5.2)$$

Izraz 5.2 predstavlja opći eksplicitni prikaz autoregresijskog modela reda p .

Neke od općih pretpostavki kojima se ne gubi općenitost prethodno opisanih izraza, a koje će nam biti od u budućim raspisima, su:

- autoregresijski parametri zaprimaju isključivo ne-nulte vrijednosti ($\phi_p \neq 0$),
- utjecaj konstantnog člana možemo zanemariti ($\theta_0 = 0$),
- u budućim razmatranjima, vršiti ćemo analizu centriranog procesa opisanog jednadžbom:

$$Z_t = x_t - E(x_t) = x_t - \mu_x. \quad (5.3)$$

Ukoliko s B označimo *operator pomaka*, te primijenimo li pretpostavku definiranu jednadžbom 5.3, dolazimo do definicijskog izraza

$$\begin{aligned} BZ_t &= Z_{t-1}, \\ B^2Z_t &= B(BZ_t) = BZ_{t-1} = Z_{t-2}, \\ &\vdots \\ B^pZ_t &= Z_{t-p}. \end{aligned} \quad (5.4)$$

Daljnjim uvrštavanjem izraza 5.4 u izraz 5.1 dobivamo

$$\begin{aligned} \epsilon_t &= Z_t + \phi_1 BZ_t + \phi_2 B^2Z_t + \dots + \phi_p B^pZ_t, \\ \epsilon_t &= Z_t(1 + \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p). \end{aligned} \quad (5.5)$$

Iz izraza 5.5 možemo definirati razvoj općeg polinoma koji opisuje autoregresijske parametre kako slijedi:

$$\phi(B) = 1 + \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p. \quad (5.6)$$

Zaključno, iz prethodnih izraza slijedi sažetiji opći oblik $AR(p)$ modela:

$$\phi(B)Z_t = \epsilon_t. \quad (5.7)$$

5.1.1. Dodatne pretpostavke

U svrhu lakšeg pronalaska vrijednosti autoregresijskih koeficijenata, potrebno je definirati dodatne pretpostavke koje će nam biti korisne u budućim razmatranjima. **Opisati zašto $\mu_x = 0$ i važnost svojstva stacionarnosti te povezivanje sa MA(q)**

Detaljnija pojašnjenja spomenutih pretpostavki detaljnije ćemo opisati kroz primjere $AR(1)$ i $AR(2)$ u potpoglavljima 5.1.3. i 5.1.4..

Dodati intro u sljedeća potpoglavlja s nastojanjem pojašnjenja konkretnih svojstava autoregresijskih modela na primjeru modela nižeg reda

5.1.2. Metode određivanja vrijednosti nepoznatih koeficijenata

Metoda najmanjih kvadrata

Jedna od metoda dobivanja vrijednosti parametara **ovaj shit nadopuniti**. Prisjetimo li metode najmanjih kvadrata u potpoglavlju 3.2.2., njezinom direktnom primjenom slijedi odabir vrijednosti parametara $\mu, \phi_1, \dots, \phi_p$ koji će zadovoljavati uvjet minimizacije funkcije srednje kvadratne greške

$$\begin{aligned} \mathbb{E}(\mu, \phi_1, \dots, \phi_p) &= \sum_{t=p+1}^N \epsilon_t^2 \\ &= \sum_{t=p+1}^N \{x_t - \mu_x + \phi_1 (x_{t-1} - \mu_x) + \dots + \phi_p (x_{t-p} - \mu_x) + \omega_t\}^2. \end{aligned} \quad (5.8)$$

Minimizacijom izraza 5.5 dolazimo do izraza

$$\sum_{t=p+1}^N \{x_t - \mu_x + \phi_1 (x_{t-1} - \mu_x) + \dots + \phi_p (x_{t-p} - \mu_x) + \omega_t\} = 0, \quad (5.9)$$

iz kojega konačno slijedi izraz

5.1.3. Autoregresijski model 1. reda

5.1.4. Autoregresijski model 2. reda

5.2. Pomična srednja vrijednost (MA)

5.3. ARMA (p,q) proces

5.4. ARIMA proces

6. Pokus

7. Zaključak

(Tu ide rezime core teme rada. Dodatno se može iskoristiti "Ken Holden, David A. Peel, John L. Thompson - Economic Forecasting: An Introduction-Cambridge University Press (1990)" kao outro; nadograditi vrlo dobrim dodatnim zaključcima koji su dati u konkretnom primjeru, ali poopćiti budući da nakon uvoda treba održati elektrotehniku dominantnom temom.) ...

Temeljem stečenih saznanja, zaključujemo da se konstrukcija modela i njihovo korištenje u prognostici može sažeti u sljedeće korake:

1. Odabir valjane teorije koja najbolje opisuje ponašanje promatranih podataka i relevantnih faktora – koje mogu biti klasificirani kao endogeni faktori, čije je ponašanje predmet konstruiranoga modela, ili egzogeni faktori, čije je ponašanje promatrano deterministički i neovisno o egzaktnom skupu parametara određenih konstruiranim modelom.
2. Matematički zapis korištene teorije, koji povezuje prethodno navedene faktore pomoću složene matematičke jednadžbe. Poseban naglasak pri određivanju međusobnog odnosa parametara ima rezultantni odnos prethođenja ili kašnjenja konstruiranoga modela spram varijable matematičkoga očekivanja i diskretnih mjernih podataka.
3. Pronalazak najreprezentativnijega skupa podataka, čije su vrijednosti u skladu s mjeriteljskim normama koje podržavaju korištenu opisnu teoriju.
4. Korištenje prikladnih kvantitativnih metoda za procjenu numeričkih vrijednosti nepoznatih parametara matematičke jednadžbe koji će najbolje aproksimirati vrijednosti u području korištenog skupa podataka.
5. Prognoziranje budućega ponašanja predmeta konstruiranoga modela temeljem prognoziranih vrijednosti kauzalnog modela egzogenih faktora – iz kojih nadalje proizlaze prognozirane vrijednosti endogenih faktora. Vjerodostojnost navedenoga postupka proizlazi iz poznavanja budućega ponašanja egzogenih faktora, što sprječava neželjeni efekt projekcije prošlih vrijednosti na budućnost.

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

Bibliografija

- [1] Priestley, M.B.: "Spectral Analysis and Time Series (Vols. 1 & 2)", Academic Press, London, 1981.
- [2] Broersen, Piet M.T.: "Automatic Autocorrelation and Spectral Analysis", Springer, London, 2006.
- [3] F.M. Dekking i dr.: "A Modern Introduction to Probability and Statistics: Understanding Why and How", Springer, London, 2006.
- [4] Holden, K.; Peel, D.A.; Thompson, J.L.: "Economic forecasting: an introduction", Cambridge University Press, Cambridge, 1990.
- [5] Presidential Commission: "Report of the Presidential Commission on the Space Shuttle Challenger Accident", U.S. Government Printing Office, Washington, D.C., 1986.