



UNIVERZA  
V LJUBLJANI

FMF

Fakulteta za matematiko  
in fiziko

---

# Metoda maksimalne entropije

---

*Modelska analiza, vaja 12*

**Avtor:** Urban Pečoler  
**Vpisna številka:** 28242019

Januar 2025

# Kazalo

<b>1</b>	<b>Uvod</b>	<b>1</b>
<b>2</b>	<b>Ločljivost modela</b>	<b>2</b>
<b>3</b>	<b>Frekvenčni spekter</b>	<b>3</b>
<b>4</b>	<b>Linearna napoved</b>	<b>6</b>
4.1	Sončni cikel . . . . .	7
4.2	Lunine efemeride . . . . .	8
4.3	Borza . . . . .	10

# 1 Uvod

Metoda maksimalne entropije (Maximum Entropy Method, MEM) je metoda za ocenjevanje spektralne gostote časovnih vrst na podlagi omejenega števila informacij, kot so avtokorelacijski koeficienti. Temelji na principu maksimalne entropije, ki zagotavlja, da je izračunana spektralna gostota čim manj pristranska in vsebuje le informacije, ki jih določajo dani podatki.

Naj bo signal  $s_n$  opisan z avtokorelacijsko funkcijo  $R(k)$ , ki je definirana kot

$$R(k) = \frac{1}{N - k + 1} \sum_{n=0}^{N-k} s_n s_{n-k},$$

kjer  $k$  predstavlja zamik. Na podlagi avtokorelacijske funkcije določimo parametre avtokorelacijskega modela (AR modela), ki opisujejo signal kot

$$s_n = - \sum_{k=1}^p a_k s_{n-k} + G u_n,$$

pri čemer  $a_k$  označujejo avtokorelacijske parametre,  $G$  pa ojačenje šuma  $u_n$ . Parametre  $a_k$  izračunamo z reševanjem sistema Yule-Walkerjevih enačb

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p-1) \\ R(1) & R(0) & \cdots & R(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R(p-1) & R(p-2) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix}.$$

Spektralna gostota moči  $P(\omega)$  je nato določena kot

$$P(\omega) = \frac{G^2}{\left| 1 + \sum_{k=1}^p a_k e^{-i\omega k} \right|^2}.$$

Prednost metode maksimalne entropije je v tem, da omogoča izračun spektralne gostote tudi za relativno kratke časovne vrste, kjer klasične metode, kot sta Fourierjeva transformacija ali periodogram, pogosto odpovejo zaradi nezadostne ločljivosti. Poleg tega zagotavlja spektralno oceno z največjo možno gladkostjo, skladno z razpoložljivimi podatki.

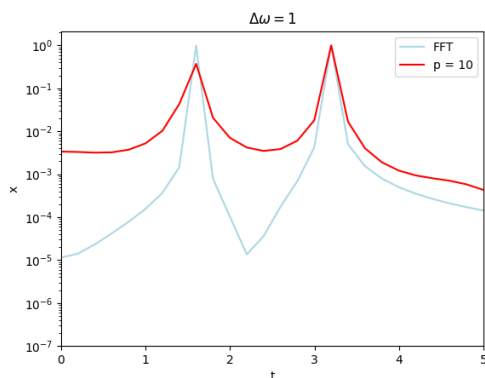
Pri nalogi sem si pomagal s Pytonsko knjižnjico `spectrum` in funkcijama `aryule` in `arma2psd`.

## 2 Ločljivost modela

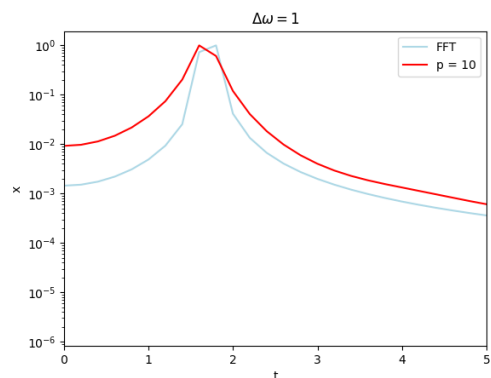
Najprej sem pogledal ločljivost modela. Ločljivost metode maksimalne entropije sem primerjal z Fourierjevo transformacijo. Kot testni signal sem si izbral neko vsoto dveh sinusov,

$$f(y) = \sin(\omega t) + \sin[(\Delta\omega + \omega)t]$$

In preveril pri kakšni  $\Delta\omega$  še vedno dobimo dva vrhova.



Slika 1: Frekvenčni spekter testnega signala določen s hitro Fourierjevo transformacijo (FFT) in metodo maksimalne entropije (MEM) za  $\Delta\omega = 10$



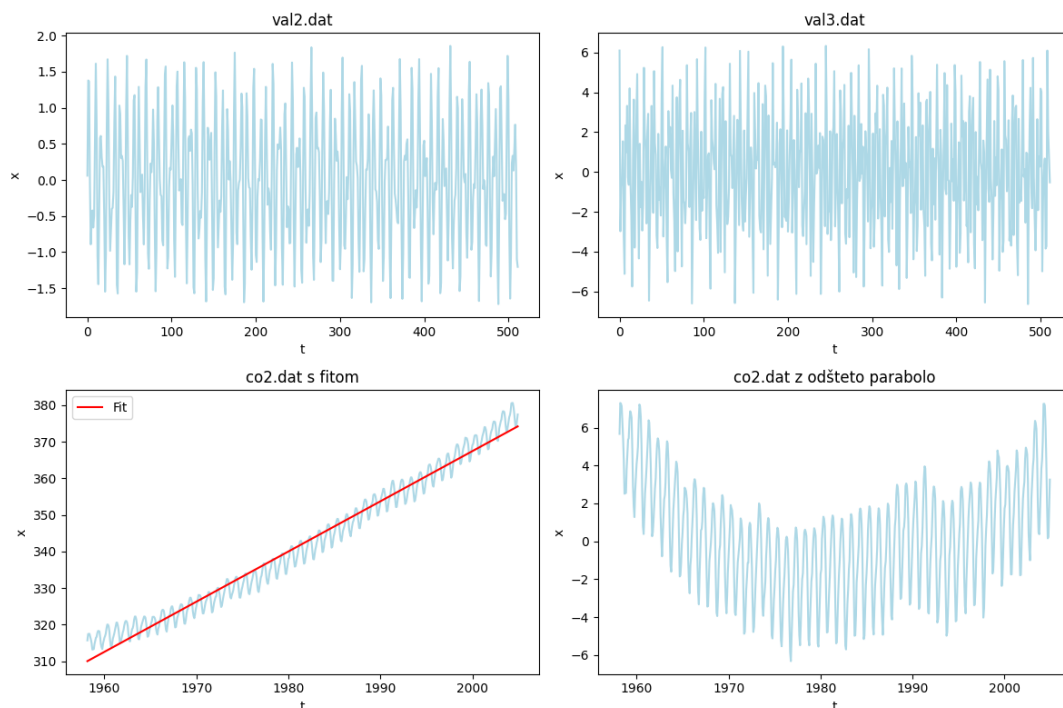
Slika 2: Frekvenčni spekter testnega signala določen s hitro Fourierjevo transformacijo (FFT) in metodo maksimalne entropije (MEM) za  $\Delta\omega = 1$

Vidimo da  $f(y) = \sin \omega t + \sin 10 + \omega t$  vrne bistveno drugačne rezultate kot  $f(y) = \sin \omega t + \sin (1 + \omega)t$ .

### 3 Frekvenčni spekter

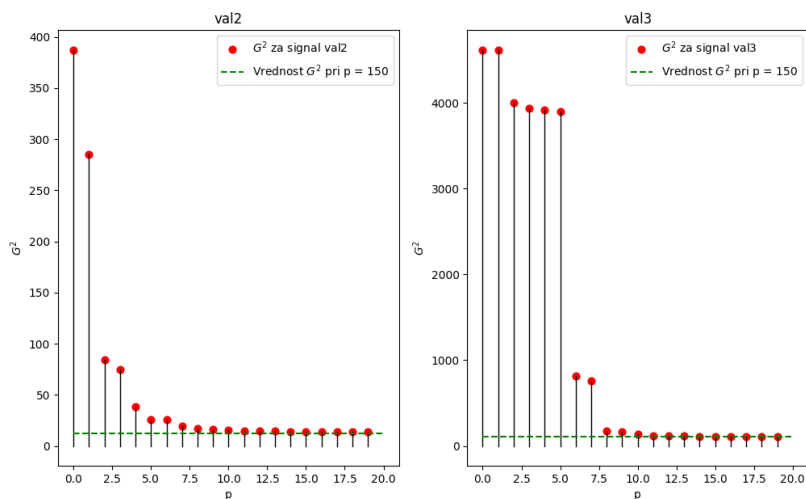
Na sliki 2 je prikazana časovna slika treh različnih signalov katerih frekvenčni spekter bom analiziral v prvem delu naloge. Datoteki "val2" in "val3" sta mi že znani, saj smo jih analizirali že pri eni od prejšnjih nalog pri tem predmetu. Tretji signal "co2" pa prikazuje letno spreminjanje  $CO_2$  v atmosferi. Pri analizi bom rezultate primerjal tudi s spektrom hitre Fourierjeve transformacije (FFT).

Za začetek sem si podatke narisal, da vidim s čim imam opravka. Signal  $CO_2$  pa sem že kar na začetku obdelal in sicer tako, da sem odštel kvadratični trend.



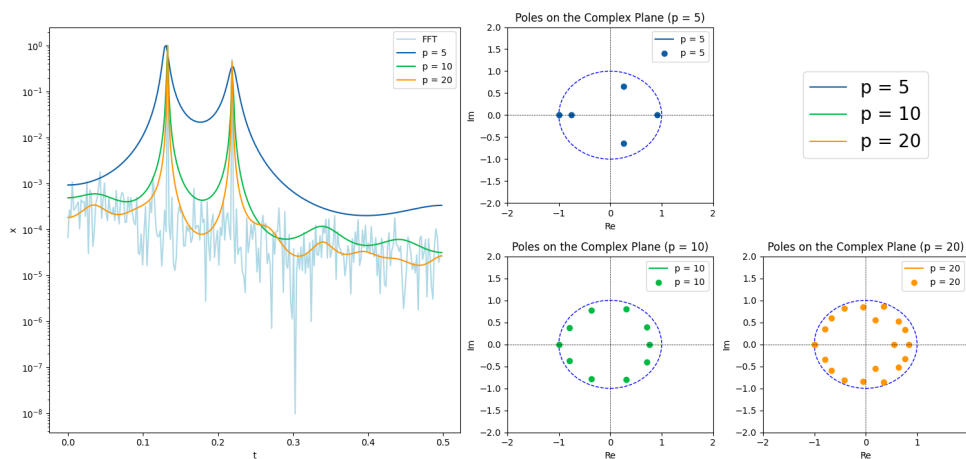
Slika 3: Narisani podatki za datoteke `val2dat`, `val3dat`, `co2dat`. SPODNJI GRAF LEVO: Fittana parabola  $ax^2 + bx + c$ , DESNO podatki `co2dat` z odštetim kvadratičnim trendom

Preden sem šel risati spektre z metodo MEM, sem pogledal minimalno število polov, ki jih potrebujem, da dobro ujamem frekvenčni spekter signala. Na sliki 4 vidimo, da se pri določeni vrednosti  $p$  vrednost  $G^2$  ustali, tako da dodtano povečanje števila polov  $p$  nima več takšne teže. Z črtkano črto sem označil nekakšno kvazi limito, ki sem jo izračunal pri  $p = 150$ . To sem naredil za signala "val2" in "val3".

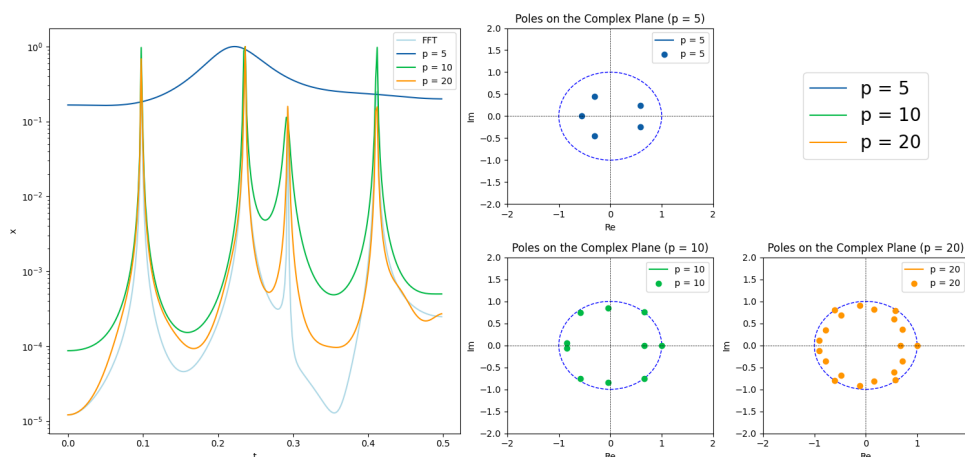


Slika 4:  $G^2$  v odvisnosti od števila polov  $p$  za datoteki **val2.dat** in **val3.dat**

Na slikah 5 in 6 je prikazana analiza signalov **val2** in **val3**. Narisal sem frekvenčni spekter in vrednosti polov za različne rede  $p$ . Vidimo, da z metodo minimalne entropije lepo zadanemo frekvence, ki so v signalu.



Slika 5: Frekvenčni spekter datoteke **val2.dat** z FFT metodo in MEM pri 3 različnih redih singularnosti. Na desni narisane vrednosti polov na kompleksni ravnini

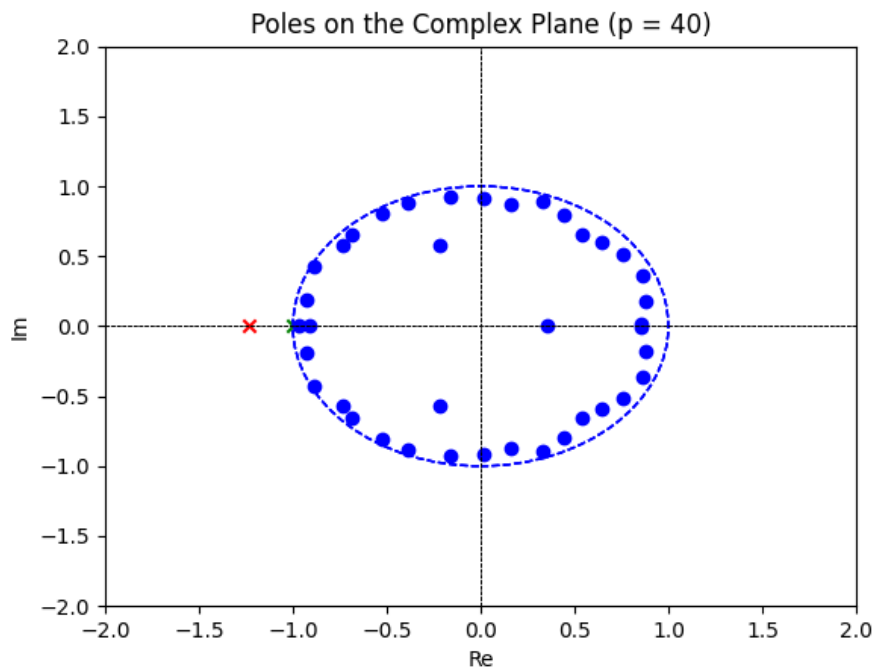


Slika 6: Frekvenčni spekter datoteke `val3dat` z FFT metodo in MEM pri 3 različnih redih singularnosti. Na desni narisane vrednosti polov na kompleksni ravnini

Pri metodi MEM je pomembno pogledati tudi, če je filter stabilen. Če vsi poli ležijo znotraj enotske kompleksne krožnice, potem je filter stabilen. Če pa je kakšen pol zunaj omenjene krožnice pa ga je treba preslikati z preslikavo

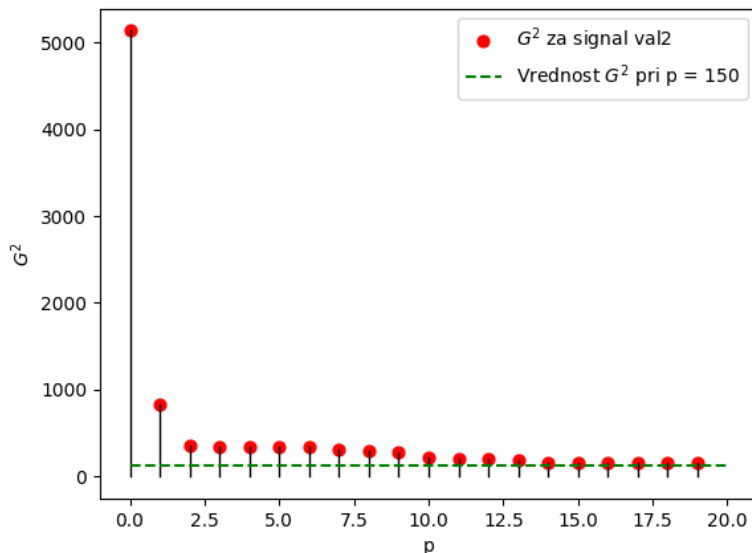
$$z_i \rightarrow \frac{z_i}{|z_i|}$$

Na sliki 7 sem to preveril in z preslikavo spremenil pole pri signalu "val2". Z rdečo je označen pol, ki je bil zunaj krožnice, z zeleno preslikava tega pola v enotsko krožnico, ostali "že stabilni" poli pa so narisani z modro barvo

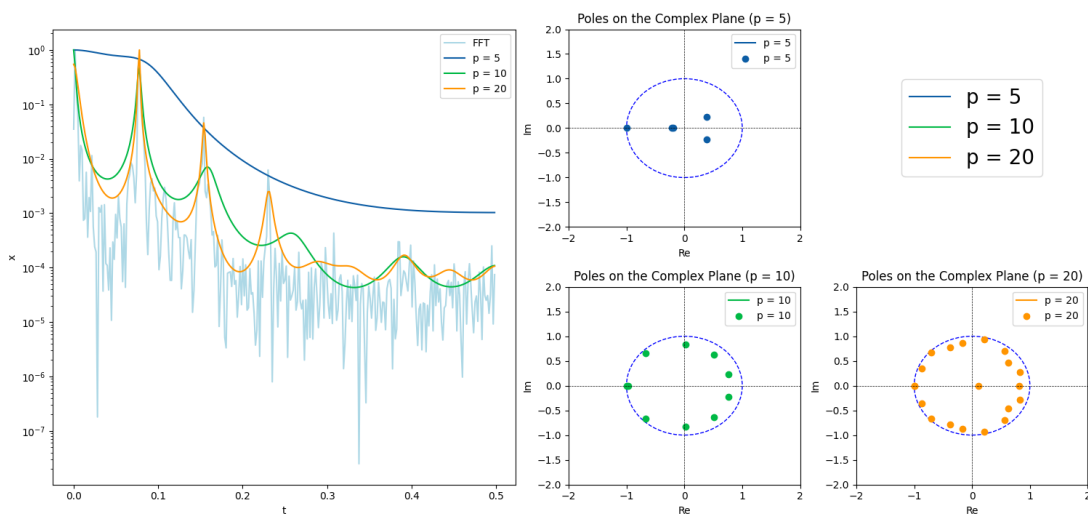


Slika 7: Narisani poli datoteke `val2dat` pri  $p = 40$ . Z rdečo barvo označen pol, ki je zunaj enotske krožnice, z zeleno transformacija tega pola v enotsko krožnico, ostali "že stabilni" poli pa so narisani z modro barvo

Enako analizo sem naredil za do sedaj še neznanе podatke  $CO_2$ . Naprej sem pogledal funkcijo  $G^2$  v odvisnosti od  $p$ , nato pa še frekvenčne spektre in koordinate polov za tri različne rede. Pri celotni analizi sem delal s podatki `co2.dat`, katerim sem odštel prej prilagojeni kvadratični trend, ki sem ga narisal na sliki 3.



Slika 8:  $G^2$  v odvisnosti od števila polov  $p$  za datoteko `co2.dat`



Slika 9: Frekvenčni spekter datoteke `co2.dat` z FFT metodo in MEM pri 3 različnih redih singularnosti. Na desni narisane vrednosti polov na kompleksni ravnini

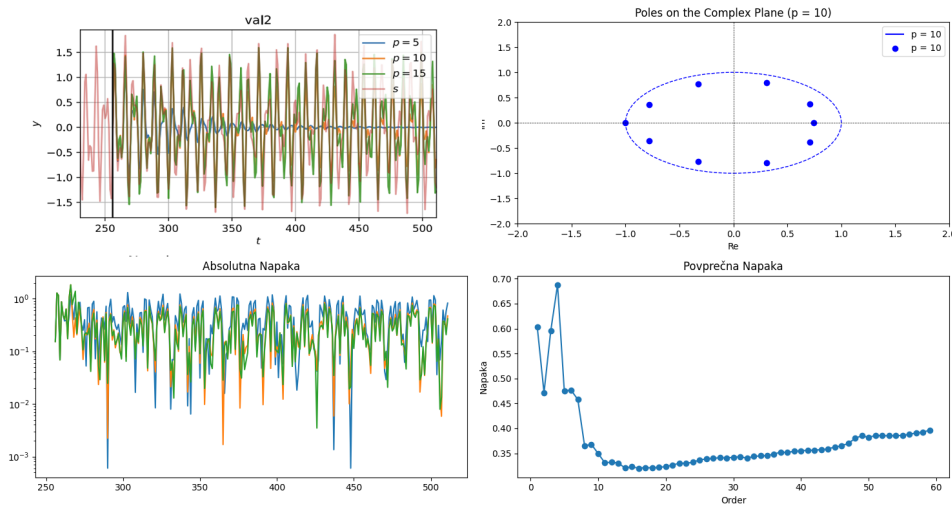
## 4 Linearna napoved

MEM metodo lahko uporabimo tudi za linearno napoved signala. Za to uporabimo enačbo

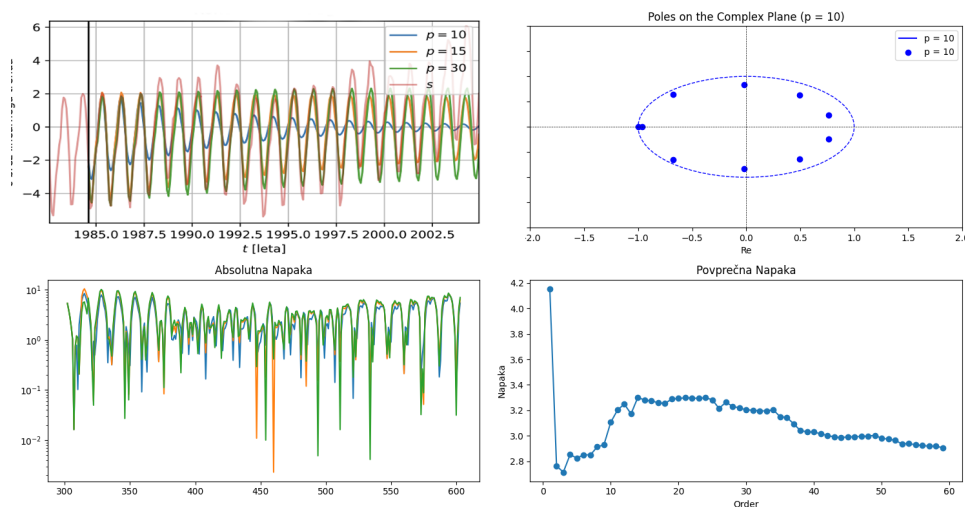
$$S_n = - \sum_{k=1}^p a_k S_{n-k} \quad (1)$$



Pri analizi sem signal po dolžini razdelil na dve polovici. Prvo polovico sem uporabil kot `train_signal`, na katerem sem izračunal koeficiente  $a_k$ , s katerimi sem opravil linearno napoved, drugo polovico signala pa sem uporabil za primerjavo in izračun natančnosti napovedi. Najprej sem pogledal napoved za že analizirana signala "val2" in "co2".



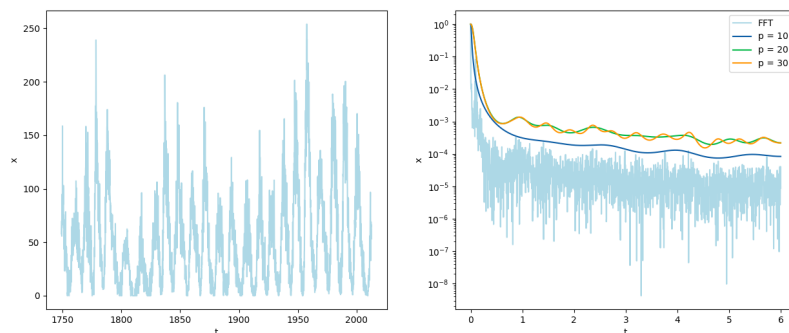
Slika 10: [LEVO ZGORAJ] Napoved druge polovice signala datoteke `val2dat`, [DESN0 ZGORAJ] narisani poli reda  $p = 10$ , [LEVO SPODAJ] Absolutna napaka napovedanega signala od pravih vrednosti za enake rede polov, kot so narisani na zgornjem grafu [DESN0 SPODAJ] Povprečna napaka napovedi v odvisnosti od reda polov



Slika 11: LEVO ZGORAJ] Napoved druge polovice signala datoteke `co2`, [DESN0 ZGORAJ] narisani poli reda  $p = 10$ , [LEVO SPODAJ] Absolutna napaka napovedanega signala od pravih vrednosti za enake rede polov, kot so narisani na zgornjem grafu [DESN0 SPODAJ] Povprečna napaka napovedi v odvisnosti od reda polov

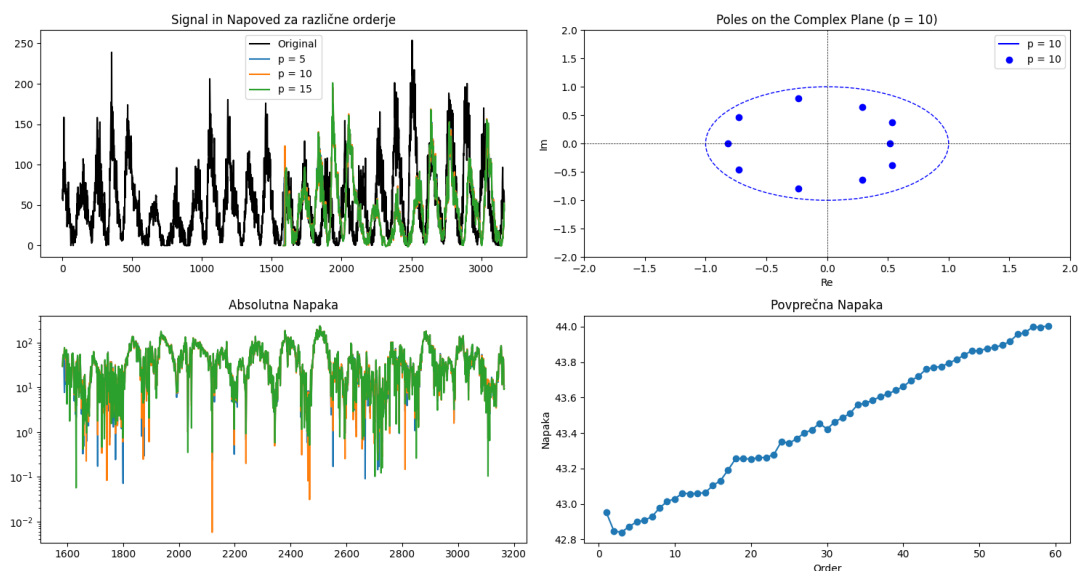
#### 4.1 Sončni cikel

V podatkih `Wolf_numberdat` smo imeli cikle Sončeve aktivnosti, ki jih kvantiziramo z Wolfovim številom. Podatki so bili podani za časovno obdobje med letoma 1750 in 2012. Vidimo, da je



Slika 12: [LEVO] Vizualizacija podatkov datoteke `Wolf_number.dat` [DESNO] Frekvenčni spekter datoteke

signal precej bolj kompliciran in zašumljen od signalov ki smo jih obdelovali v poglavju 2. Na sliki 12 sem izrisal podatke in prikazal frekvenčno sliko s pomočjo FFT ter MEM metode za različno število polov. Na sliki 13 pa sem naredil linearno napoved in analizo napak.

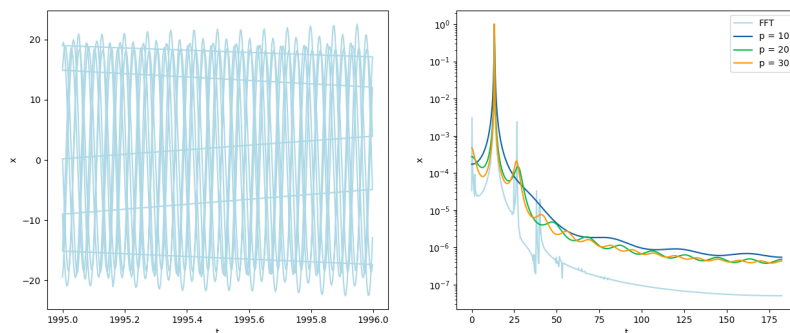


Slika 13: [LEVO ZGORAJ] Napoved druge polovice signala datoteke `Wolf_number.dat`, [DESNO ZGORAJ] narisani poli reda  $p = 10$ , [LEVO SPODAJ] Absolutna napaka napovedanega signala od pravih vrednosti za enake rede polov, kot so narisani na zgornjem grafu [DESNO SPODAJ] Povprečna napaka napovedi v odvisnosti od reda polov

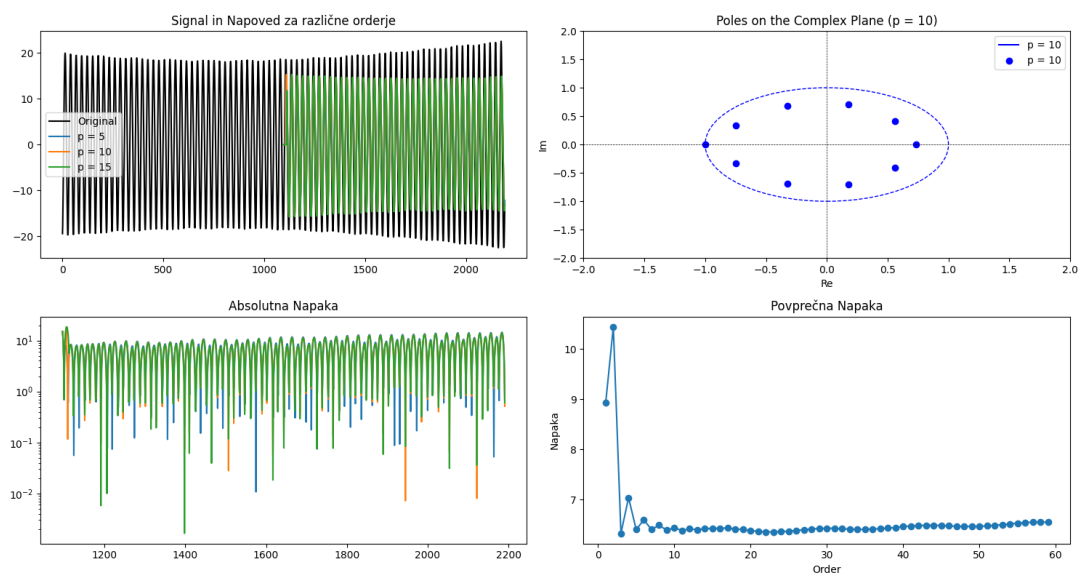
Vidimo, da se napaka z številom polov veča, kar je pričakovano, saj se tudi linearna napoved slabša z številom polov. Že iz slike 12 je razvidno, da ne FFT ne MEM ne data nobenih vidnih vrhov, kar pomeni da obe metodi vidita le šum in nobene "prave frekvence".

## 4.2 Lunine efemeride

Naslednji problem in podan data set, ki sem ga pogledal je bil problem napovedi luninih efemerid. Podatki so bili razdeljeni na datum, deklinacijo in rektascenzijo, tako da sem analiziral posebej vsako od slednjih dveh. Najprej sem pogledal deklinacijo.

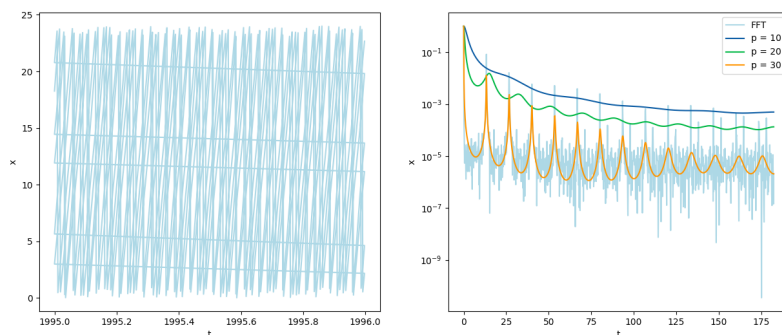


Slika 14: [LEVO] Vizualizacija podatkov datoteke `lunadat` za primer deklinacije [DESNO] Frekvenčni spekter datoteke

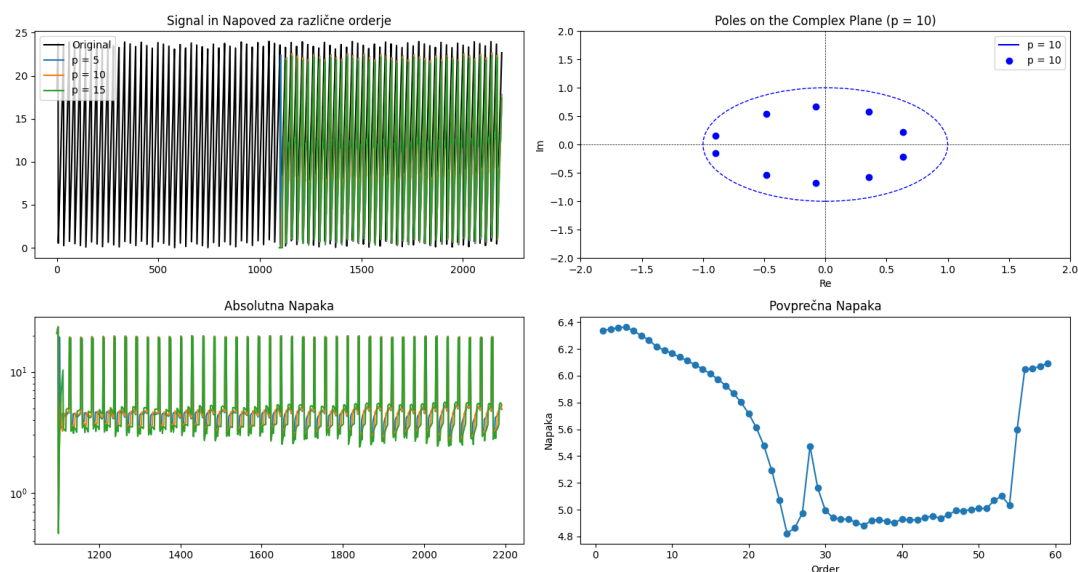


Slika 15: [LEVO ZGORAJ] Napoved druge polovice signala datoteke `lunadat`, [DESNO ZGORAJ] narisani poli reda  $p = 10$ , [LEVO SPODAJ] Absolutna napaka napovedanega signala od pravih vrednosti za enake rede polov, kot so narisani na zgornjem grafu [DESNO SPODAJ] Povprečna napaka napovedi v odvisnosti od reda polov

Nato pa še podatke za rektascenzijo.



Slika 16: [LEVO]Vizualizacija podatkov datoteke `lunadat` za primer rektascenzije [DESNO] Frekvenčni spekter datoteke

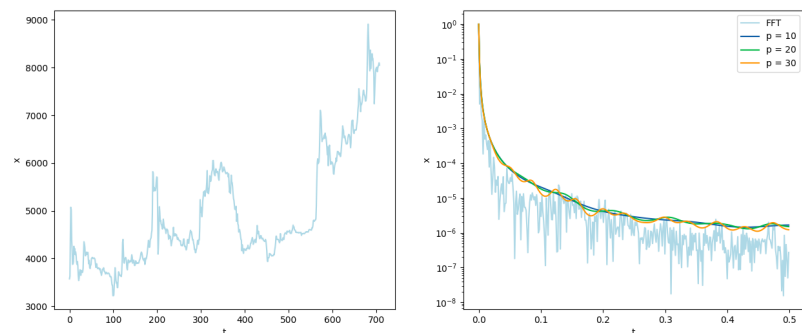


Slika 17: [LEVO ZGORAJ]Napoved druge polovice signala datoteke `lunadat`, [DESNO ZGORAJ] narisani poli reda  $p = 10$ , [LEVO SPODAJ] Absolutna napaka napovedanega signala od pravih vrednosti za enake rede polov, kot so narisani na zgornjem grafu [DESNO SPODAJ] Povprečna napaka napovedi v odvisnosti od reda polov

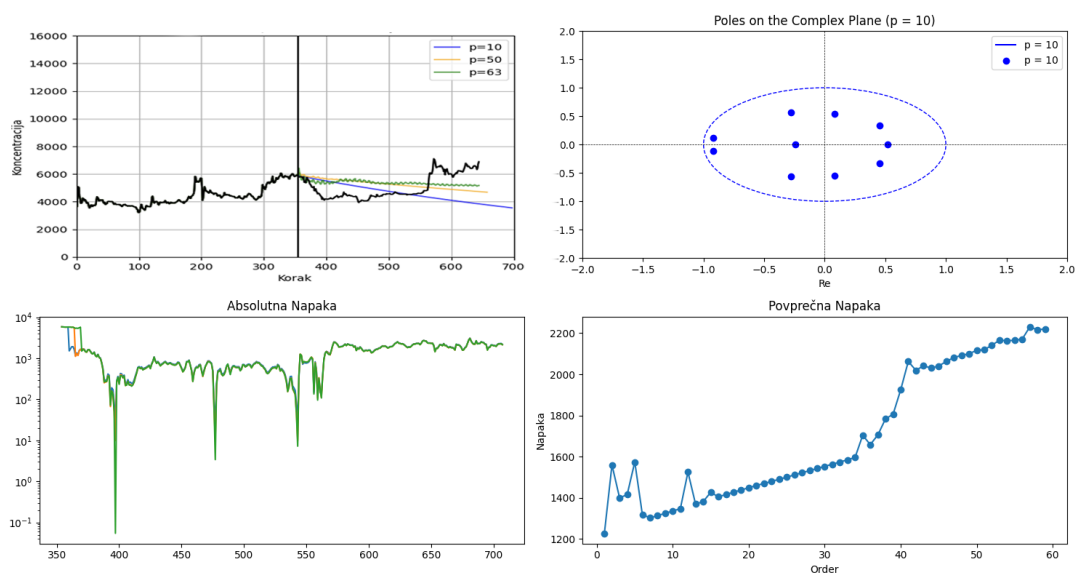
Vidimo, da pri obeh primerih z zmerno visokim  $p$ -jem dobro ujamemo frekvence, kot jih najde FFT. Napaka se z večjimi  $p$ -ji manjša, v primeru rektascenzije pa se okrog  $p = 50$  spet poveča.

### 4.3 Borza

Kot zadnji primer pa sem pogledal še podatke iz borze. Od vseh primerov gre tukaj za najbolj naključne podatke, tako da je pričakovati da bodo naše metode povsem odpovedale.



Slika 18: [LEVO] Vizualizacija podatkov datoteke **borzadat** [DESNO] Frekvenčni spekter datoteke



Slika 19: [LEVO ZGORAJ] Napoved druge polovice signala datoteke **borzadat**, [DESNO ZGORAJ] narisani poli reda  $p = 10$ , [LEVO SPODAJ] Absolutna napaka napovedanega signala od pravih vrednosti za enake rede polov, kot so narisani na zgornjem grafu [DESNO SPODAJ] Povprečna napaka napovedi v odvisnosti od reda polov

Pričakovano ne ujamemo nobene frekvence, saj ni nič periodičnega, prav tako linearna napoved povsem odpove.