

# 12. naloga - Metoda maksimalne entropije in linearna napoved

Tina Klobas

3. avgust 2022

## 1 Opis problema

V nalogi z uporabo metode maksimalne entropije za določanje spektrov podanih signalov, ki jih primerjamo s tistimi pridobljenimi z metodo Fourierove transformacije.

## 2 Določanje frekvenčnega spektra

Določi frekvenčni spekter signalov iz datotek `val2.dat`, `val3.dat` in `co2.dat` z metodo maksimalne entropije. V slednji datoteki s koncentracijo CO<sub>2</sub> v zraku upoštevaj splošen (linearen) letni trend. Preizkusi delovanje metode v odvisnosti od števila polov in od gostote prikaza. Pri CO<sub>2</sub> si oglej še njihovo lego. Primerjaj natančnost metode z rezultati, ki jih da FFT ob uporabi filtrov. Sestavi tudi testni signal iz vsote sinusnih členov z bližnjimi frekvencami ter razišči ločljivost metode.

Za linearno napoved signala uporabimo zvezo:

$$S_n = \sum_{k=1}^p a_k S_{n-k} + Gb_0. \quad (1)$$

Želimo določiti koeficiente  $a_k$  s katerimi bomo lahko določili spekter signala. Da to dosežemo, moramo minimizirati kvadrat napake

$$\xi^{(p)} = E[e_n] = E[(S_n + \sum_{k=1}^p a_k S_{n-k})^2], \quad (2)$$

oziroma

$$\xi_{\min(p)} = E[S_n] + \sum_{k=1}^p a_k E[S_n S_{n-k}]. \quad (3)$$

$E[S_n]$  izračunamo z avtokorelacijsko funkcijo:

$$R(k) = \frac{1}{N-k} \sum_n S_n S_{n-k}. \quad (4)$$

Dobimo sistem:

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p-1) \\ R(1) & R(0) & \cdots & R(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R(p-1) & R(p-2) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix}. \quad (5)$$

Z dobljenimi koeficienti  $a_k$  lahko nato določimo gostoto spektra:

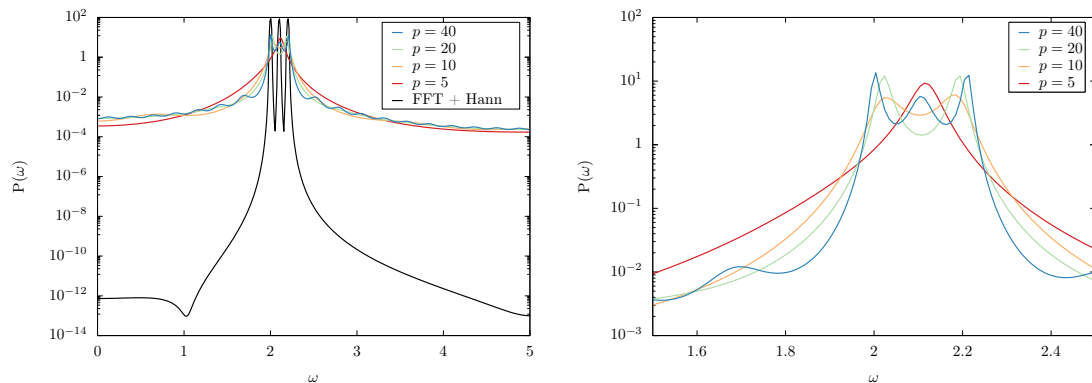
$$P(\omega) = \frac{G^2}{|1 + \sum_{k=1}^p a_k e^{-i\omega k}|^2}, \quad (6)$$

ki jo lahko primerjamo s tisto dobljeno z metodo Fourierove transformacije.

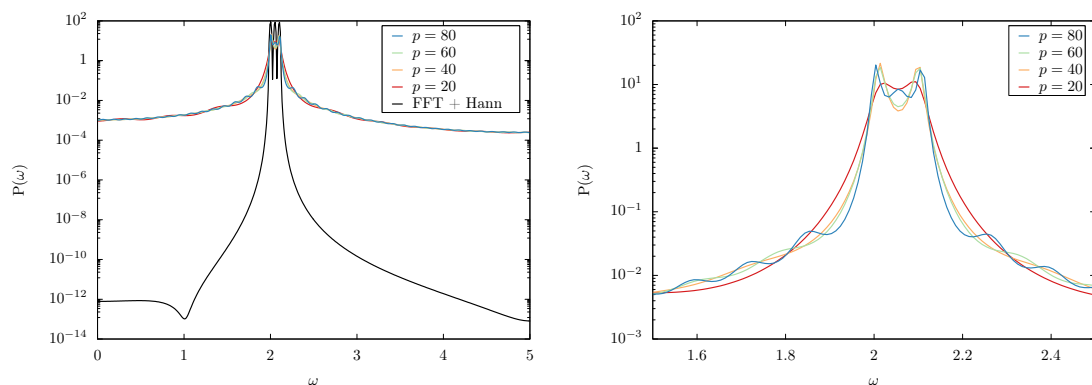
## 2.1 Reševanje in rezultati

### Testni signal

Za začetek primerjajmo ločljivost metode maksimalne entropije s Fourierovo transformacijo, kjer uporabimo Hannovo okensko funkcijo. Za testni signal uporabimo vsoto treh sinusov in preverimo, koliko blizu so si lahko njihove frekvence, da jih bo metoda še našla.



Slika 2.1: Primerjava obeh metod na testnem signalu  $\sin(4\pi x) + \sin(4, 2\pi x) + \sin(4, 4\pi x)$ . Na desni je prikazana primerjava MEM pri različnih številih polov  $p$ .

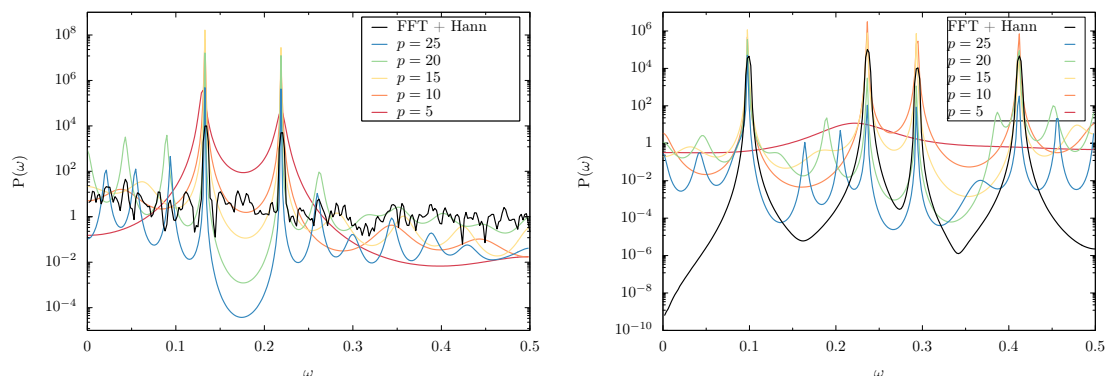


Slika 2.2: Primerjava obeh metod na testnem signalu  $\sin(4\pi x) + \sin(4, 1\pi x) + \sin(4, 2\pi x)$ . Na desni je prikazana primerjava MEM pri različnih številih polov  $p$ .

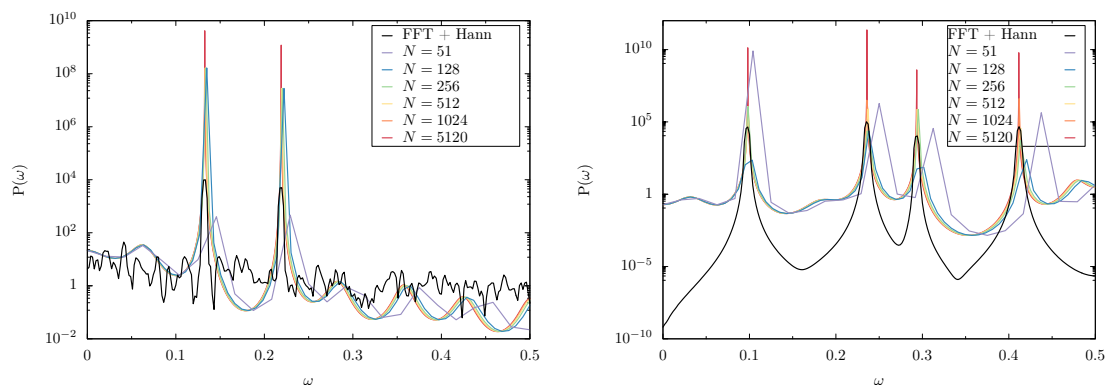
Na grafih 2.1 in 2.2 vidimo, da metoda s Fourierovo transformacijo dobro zadane obe testni funkciji. Pri primeru 2.1 vidimo, da metoda maksimalne entropije zadane vse tri vrhove šele pri zelo velikem številu  $p = 40$ , pri testni funkciji s še bližjimi frekvencami pa vse tri zadane šele pri  $p = 80$ . Pri nadaljnjem večanju parametra  $p$  izboljšamo ločljivost, vendar postajajo vrhovi manj izraziti, pojavijo se pa tudi dodatni manjši vrhovi.

#### 2.1.1 Datoteki val2.dat in val3.dat

Najprej si pogledjmo spektre za signale iz datotek `val2.dat` ter `val3.dat` pri različnih številih polov  $p$  in gostote prikaza. Spektre izračunane z metodo maksimalne entropije primerjamo s pristopom preko Fourierove transformacije z uporabo Hannove okenske funkcije.



Slika 2.3: Na levi je spekter datoteke `val2.dat` na desni pa `val3.dat` pri različnih številih polov  $p$  in pri številu točk  $N = 512$ .



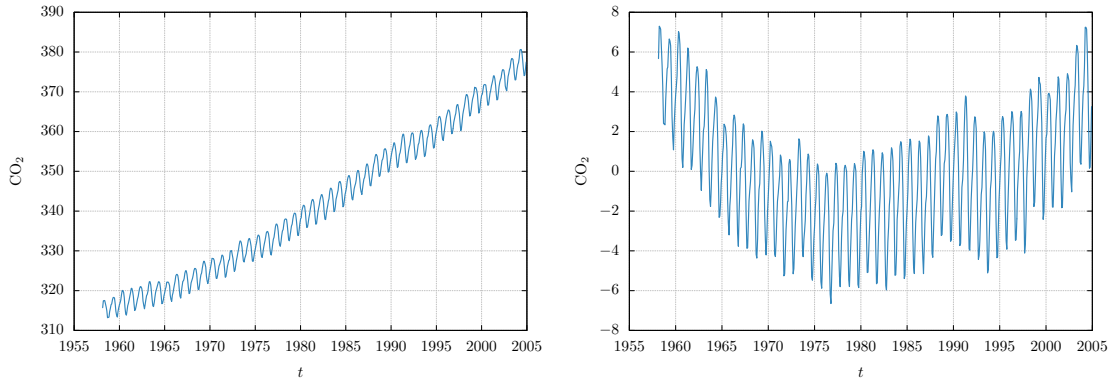
Slika 2.4: Na levi je spekter datoteke `val2.dat` na desni pa `val3.dat` pri različnih številu točk  $N$  in pri številu polov  $p = 15$ .

Na grafu 2.3 vidimo, da se lahko zgodi, da s premajhnim številom polov  $p$  vrhov signala sploh ne najdemo. Število polov  $p = 5$  je za oba signala premajhno, da bi lahko metoda podala zanesljivo informacijo o signalu. V primerjavi z metodo *FFT* je za signal datoteke `val2.dat` metoda *MEM* bolj izrazito našla vrhove in zmanjšala šum med glavnima vrhoma. Z večanjem števila polov nižamo intenziteto ozadja, vendar dosežemo neko število, ki preveč poudari nekatere frekvence, kar lahko zavede pri interpretaciji signala. Pri signalu datoteke `val3.dat` pa še opazimo, da s pretiranim večanjem števila polov izgubljam intenziteto glavnih frekvenc.

Na grafu 2.4 primerjamo ali se nam splača varčevati z gostoto prikaza. S prenizkim številom točk sicer ne zgrešimo popolnoma glavnih frekvenc in nam že tako lahko metoda poda neko informacijo o zaznanem signalu. S povečevanjem števila točk postajajo vrhovi vedno ožji in višji kot metoda *FFT*, vendar za primer signala `val3.dat` slednja veliko bolj zatire vmesne frekvence.

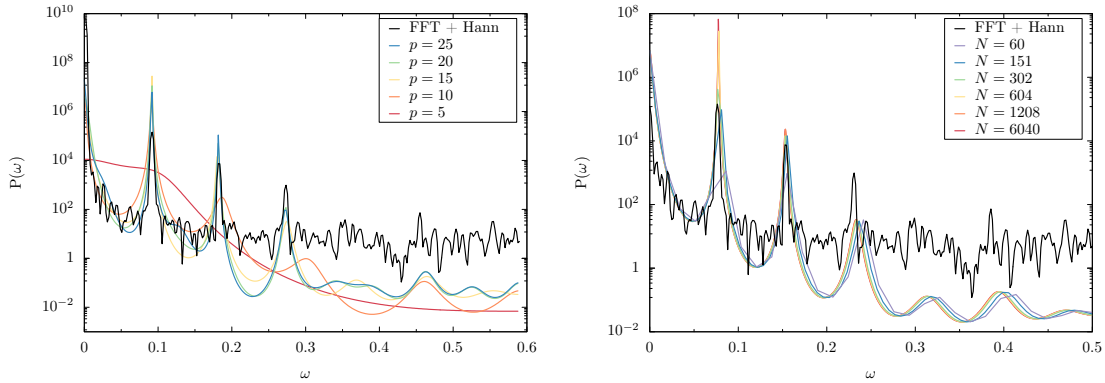
### 2.1.2 Letna količina $\text{CO}_2$ v zraku

Imamo datoteko z letnimi merjenimi podatki o koncentraciji  $\text{CO}_2$  v zraku, prikazanim na grafih 2.5.



Slika 2.5: Na levi je koncentracija  $\text{CO}_2$  v zraku, na desni pa enaka koncentracija z odštetim linearno naraščujočim trendom.

Tudi za te podatke si pogledjmo kako se spreminja informacija o spektru s spreminjanjem parametrov  $p$  ter  $N$  na grafih 2.6.

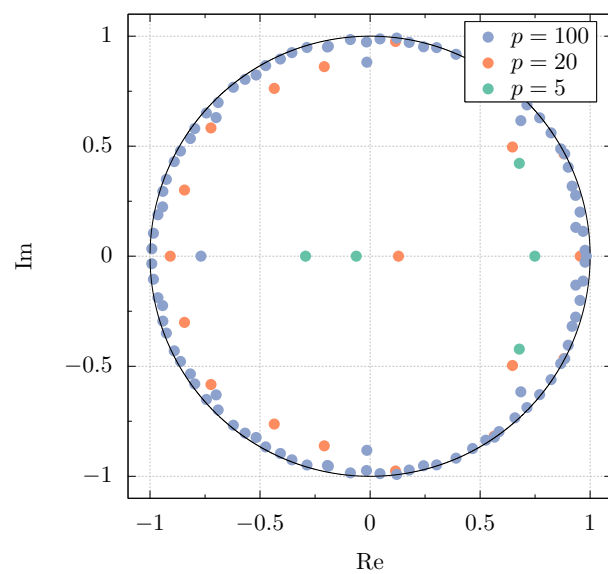


Slika 2.6: Na levi je spekter koncentracije  $\text{CO}_2$  v zraku pri številu točk  $N = 604$  in različnih parametrih  $p$  na desni pa pri vrednosti  $p = 15$  in različnih gostotah  $N$ .

Podobno – s premajhnim številom polov povsem zgrešimo vrhove spektra. V tem primeru je metoda *MEM* vselej boljša od *FFT*. Vrhovi so ožji in bolj izraziti, hkrati pa tudi zmanjšamo intenziteto ozadja in znotraj šuma se skrivata še dva vrhova, ki sta bolj opazna na desnem grafu pri številu polov  $p = 15$ . Več o zanesljivosti filtra lahko izvemo iz grafa polov v kompleksni ravnini. Izračunamo ničle polinoma:

$$1 + \sum_{k=1}^p a_k z^k = 0 \quad (7)$$

in jih narišemo na graf 2.7. Lahko opazimo, da se pri večanju števila polov približujemo enotski krožnici, pri  $p = 100$  pa že vsi bolj ali manj ležijo na njej.



Slika 2.7: Poli spektra `co2.dat`.

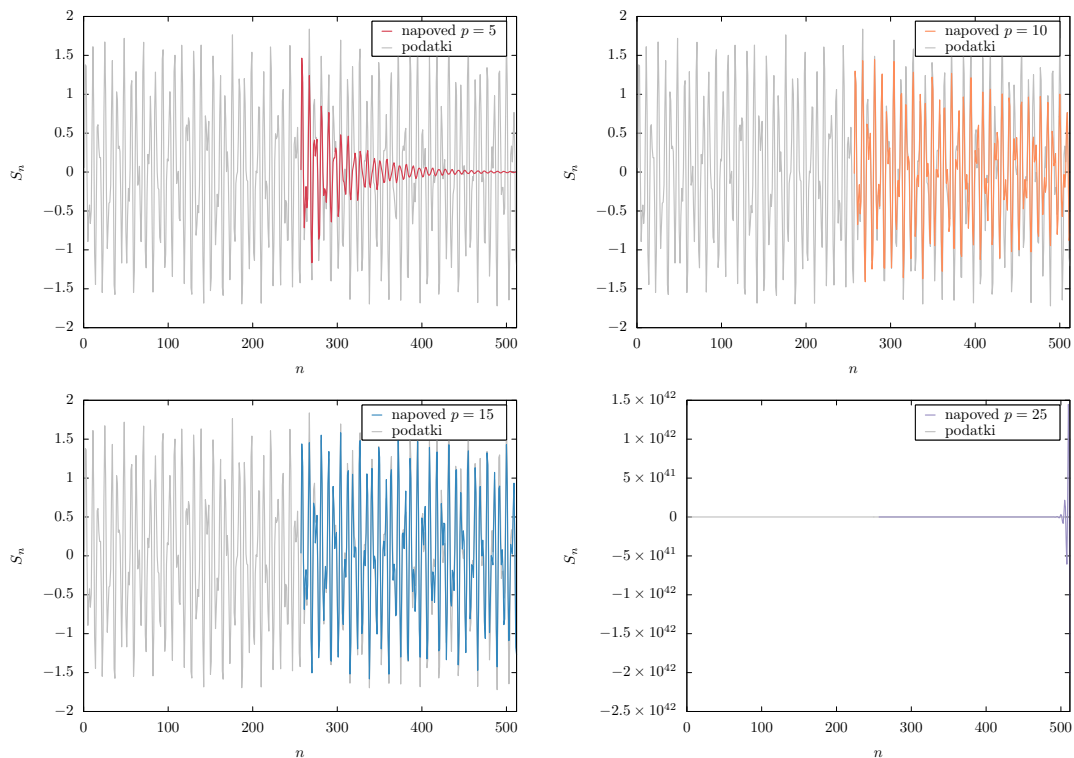
### 3 Linearna napoved

Koeficiente metode maksimalne entropije na signalih iz prve naloge uporabi za linearno napoved: koeficiente določi iz prve polovice zapisa signala, drugo polovico pa uporabi za primerjavo z napovedjo. Priporočljivo je majhno število polov. Ugotovi, kako vpliva na napoved šum! Podobno preuči datoteko `borza.dat` in podatke za lunine efemeride v `luna.dat`, oziroma sončni cikel `Wolf_number.dat`. Preveri, ali so vsi koreni preslikavne funkcije znotraj enotnega kroga, in jih po potrebi preslikaj vanj.

#### 3.1 Reševanje in rezultati

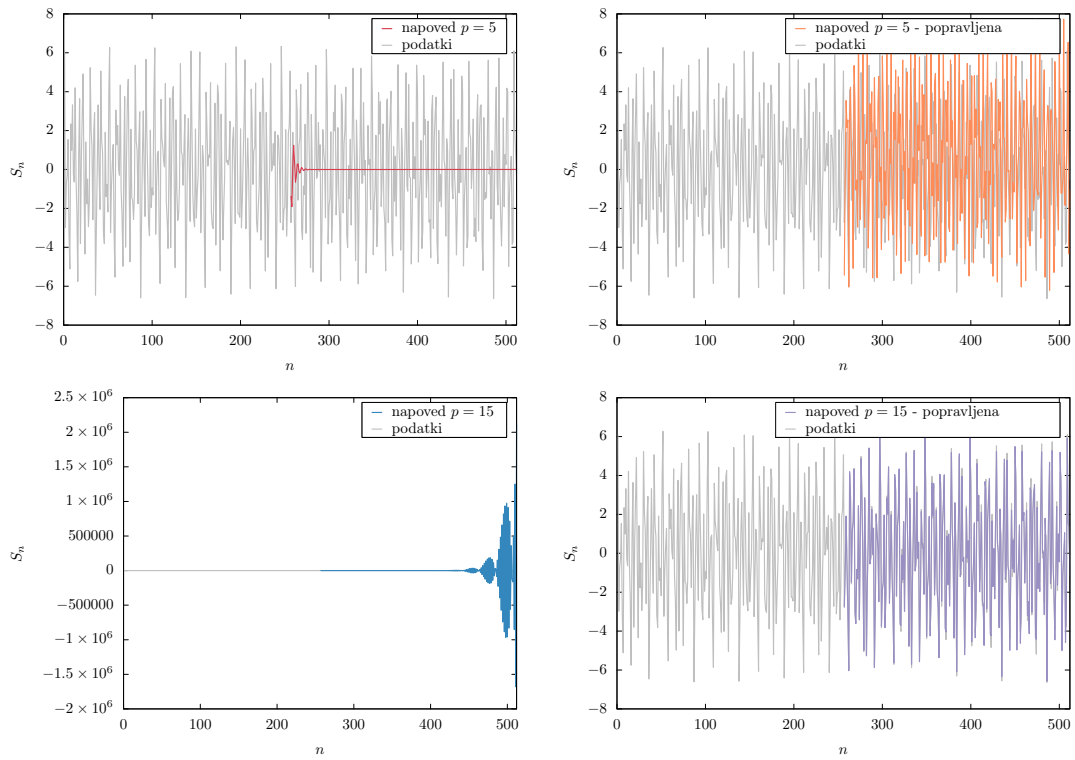
##### 3.1.1 Datoteki `val2.dat` in `val3.dat`

Preizkusimo teorijo in z uporabo enačbe 1 napovemo obnašanje signala datotek `val2.dat` in `val3.dat`.

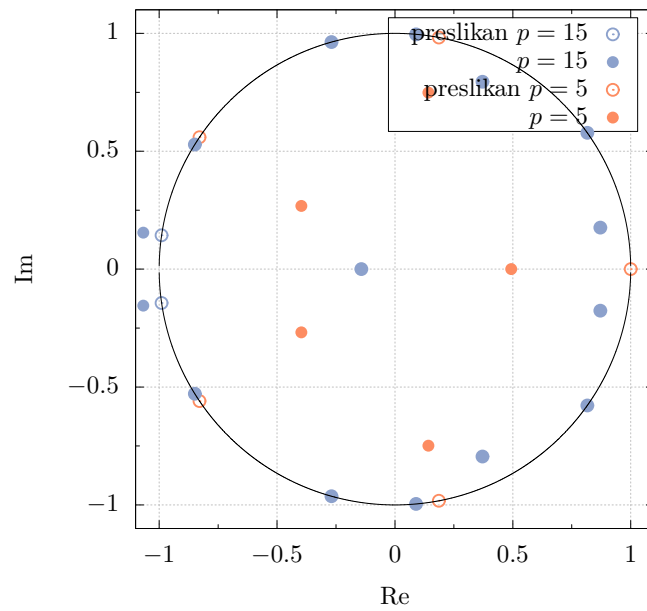


Slika 3.1: Napoved signala `val2.dat` pri različnih številih polov  $p$ .

Na grafih 3.1 opazimo, da pri premajhnem številu polov ( $p = 5$ ) napoved zelo hitro zamre, z večanjem števila polov pa se natančnost napovedi izboljšuje, dokler ne pridemo do prevelikega števila polov, kjer napoved zelo naraste. Pri podatkih iz datoteke 3.2 vidimo podoben trend. Razlog za poslabšanje napovedi leži v ničlah polov na enotski krožnici. Na grafu 3.3 vidimo, da sta za oba primera – ko signal hitro zamre in ko zelo naraste, lege polov odmaknjene od enotske krožnice. Napoved popravimo tako, da pole preslikamo na krožnico in tako popravimo koeficiente  $a$ , izboljšanje lahko opazimo na grafih 3.2.

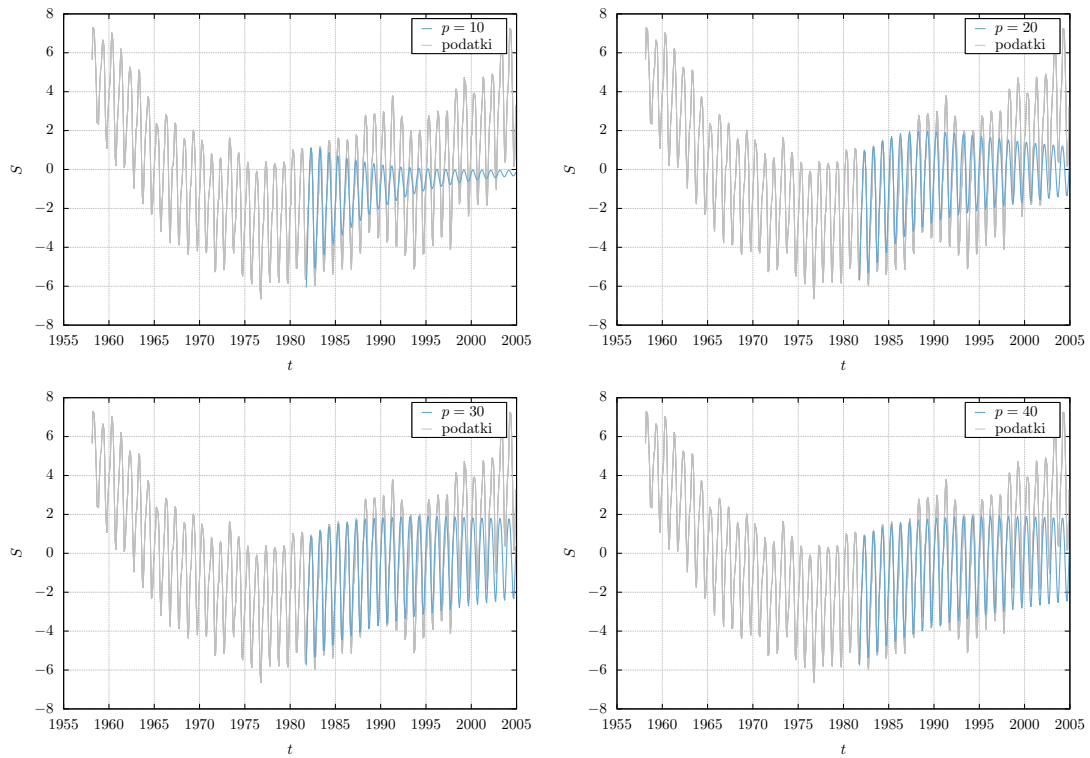


Slika 3.2: Napoved signala `val3.dat` pri različnih številih polov  $p$  in pri popravljenih koeficientih signala, pri katerih smo preslikali ničle polov na enotsko krožnico (desno).

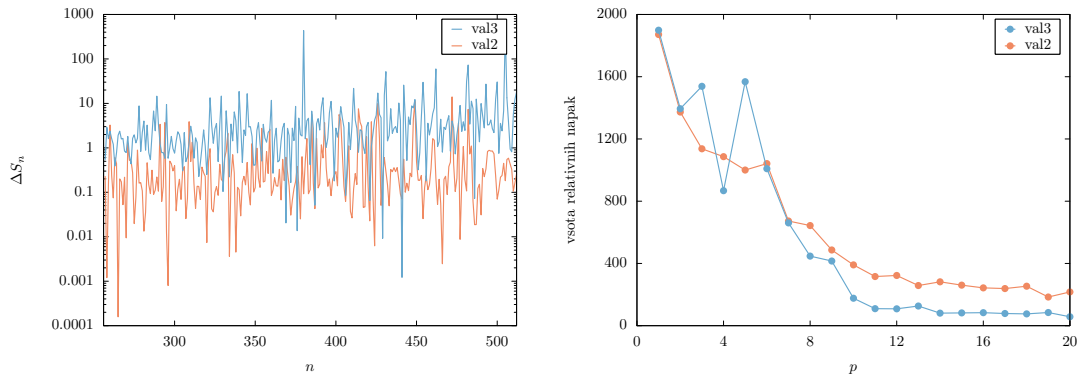


Slika 3.3: Ničle polov pri signalu `val3.dat` in ničle polov preslikane na krožnico.

Poglejmo si kaj se dogaja z natančnostjo napovedi.



Slika 3.5: Linearna napoved signal `co2.dat` pri različnih  $p$ .



Slika 3.4: Spreminjanje relativne napake napovedi signala na vsakem koraku (levo) in vsota napake napovedi v odvisnosti od števila polov  $p$  (desno) za signala `val2.dat` in `val3.dat`.

Na levem grafu 3.4 vidimo, kako se na vsakem koraku napovedi spreminja natančnost, vendar ker napaka močno fluktuira ne moremo zagotovo trditi kaj se dogaja, opazimo pa rahlo linearno naraščanje. Na desnem grafu 3.4 opazimo kako se izboljšuje napaka napovedi z večanjem števila polov  $p$ . Nekje okoli vrednosti 12 se manjšanje napake ustali in se ne izboljšuje več z večanjem števila polov.

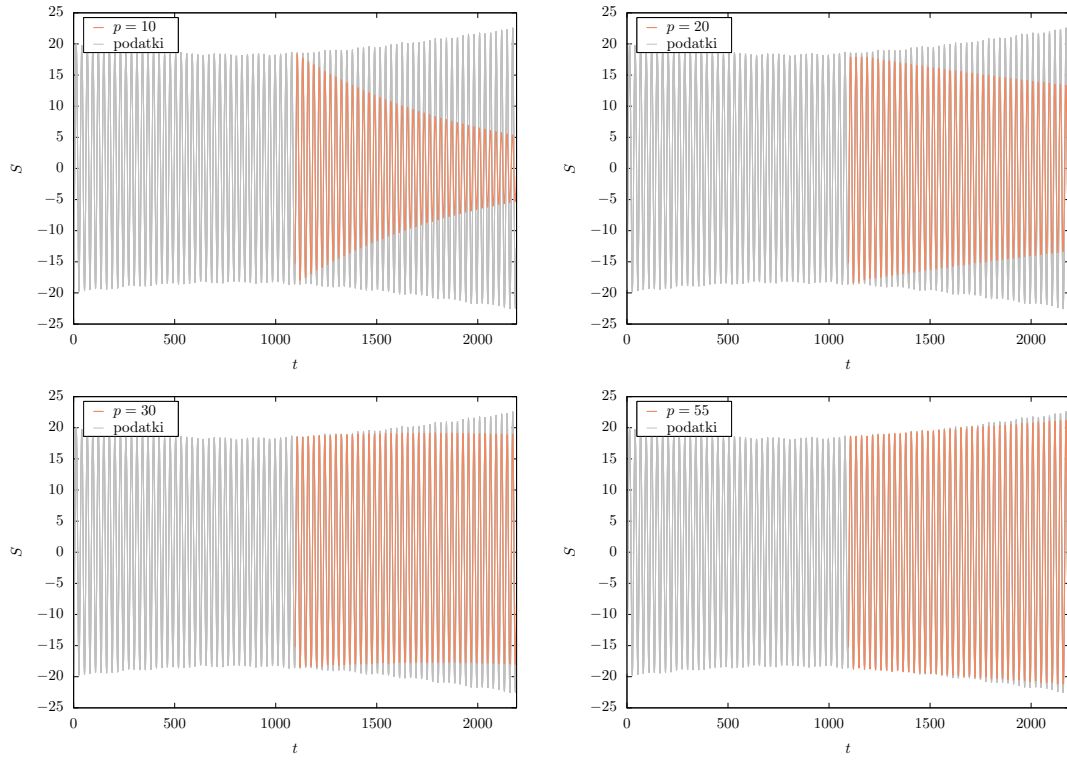
### 3.1.2 CO<sub>2</sub>

Na grafu 3.6, da vrednost  $p = 10$  premajhna, da bi zadela obnašanje signala, z večanjem števila polov pa dobro ujamemo splošno obnašanje in frekvenco signala, ne zadenemo pa amplitude.

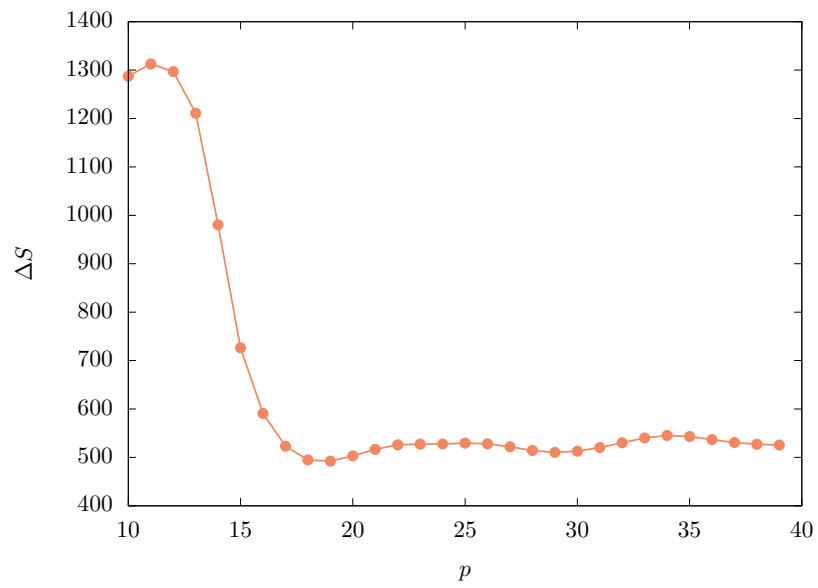
### 3.1.3 Lunine efemeride

Na grafu 3.6 lahko opazimo, da z večanjem števila polov  $p$  pri napovedi lunine deklinacije se pri velikem številu približamo amplitudi signala. Na grafu 3.7 vidimo, da tudi v tem primeru z večanjem števila polov  $p$  izboljšujemo natančnost napovedi.

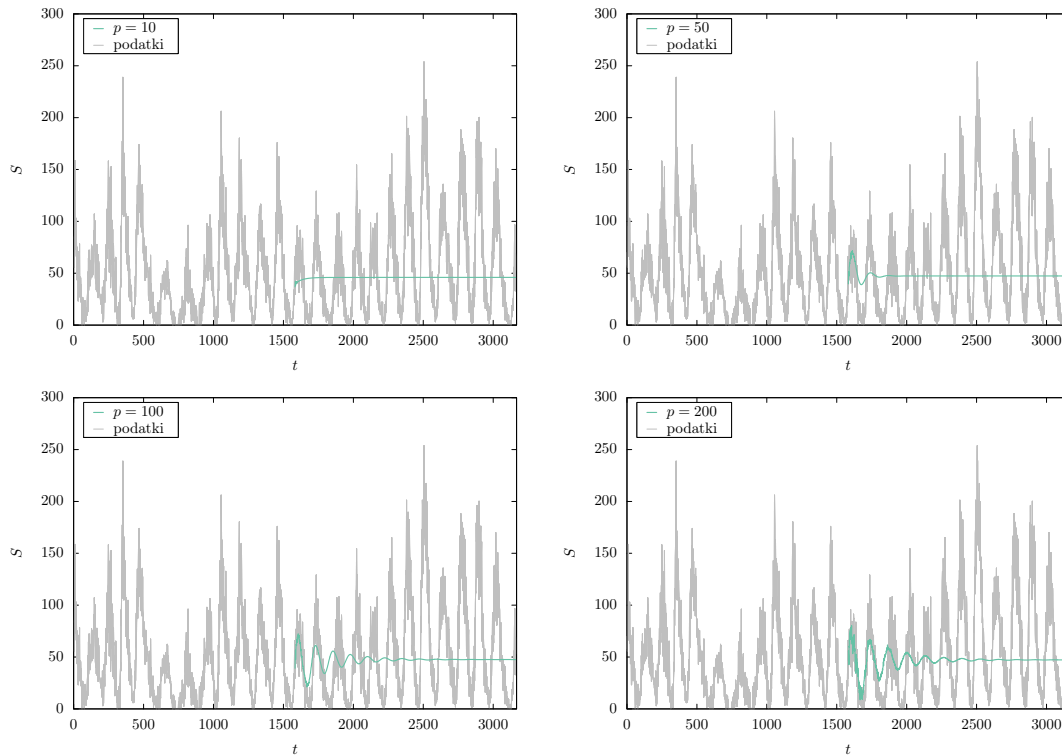




Slika 3.6: Linearna napoved lunine deklinacije pri različnih  $p$ .



Slika 3.7: Napaka linearne napovedi lunine deklinacije v odvisnosti od števila polov.



Slika 3.8: Linearna napoved Wolfovega števila pri različnih  $p$ .

### 3.1.4 Sončne pege

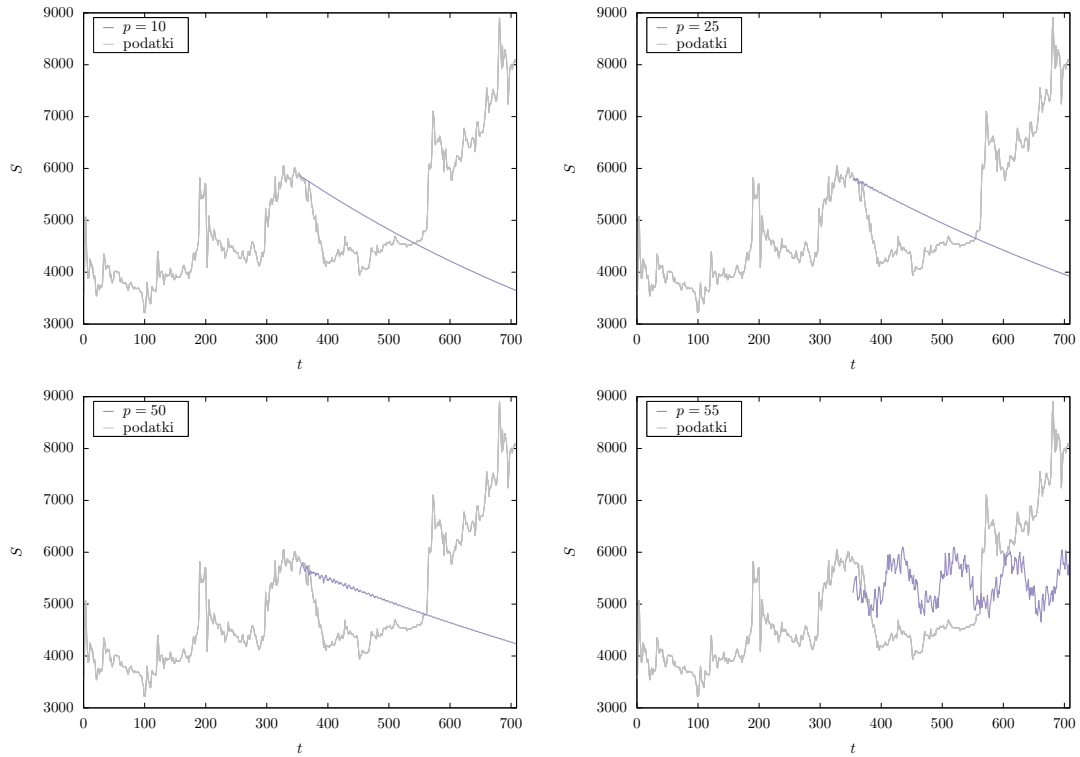
Na grafih 3.8 lahko opazimo, da pri Wolfovem številu ne uspemo dobro napovedati obnašanja signala, čeprav imamo zelo veliko število polov. Pri večjem številu polov bi zadeli frekvenco signala, vendar bi potrebovali veliko več polov in točk signala, da bi se približali njegovi amplitudi.

### 3.1.5 Borza

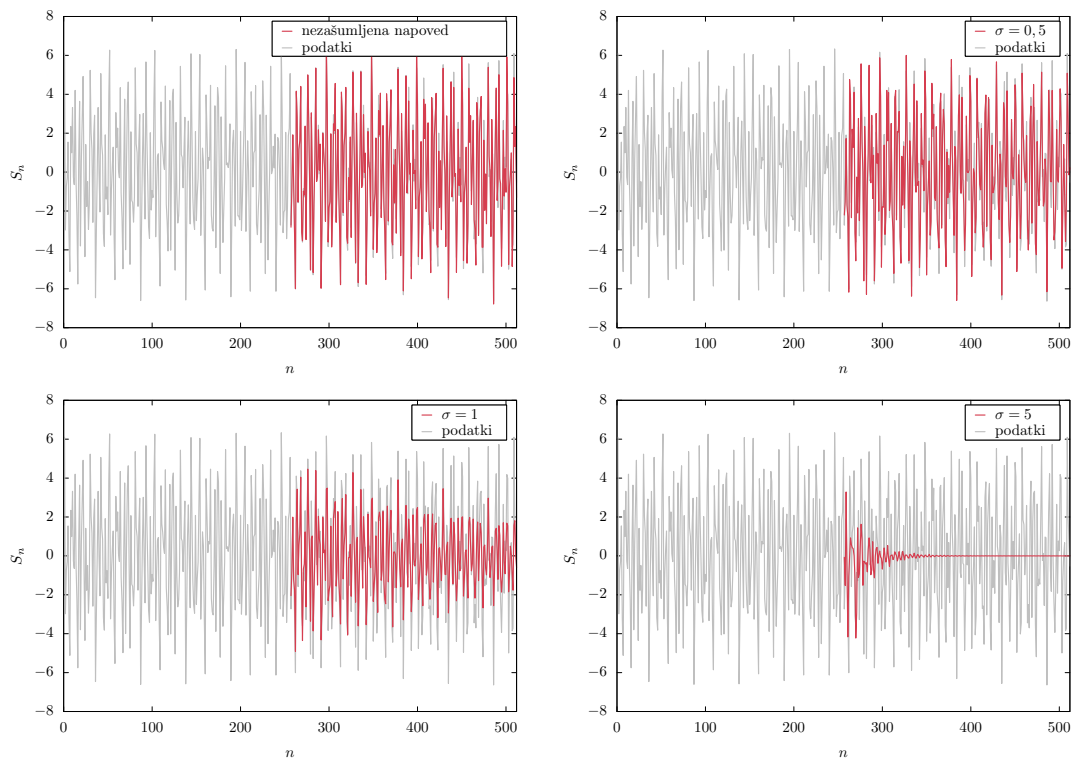
Na grafih 3.9 lahko vidimo, da metoda z linearno napovedjo ne deluje na takih podatkih, morda je bolje, da se za hiter zaslužek poslužimo drugačnih metod.

## 3.2 Vpliv šuma na napoved

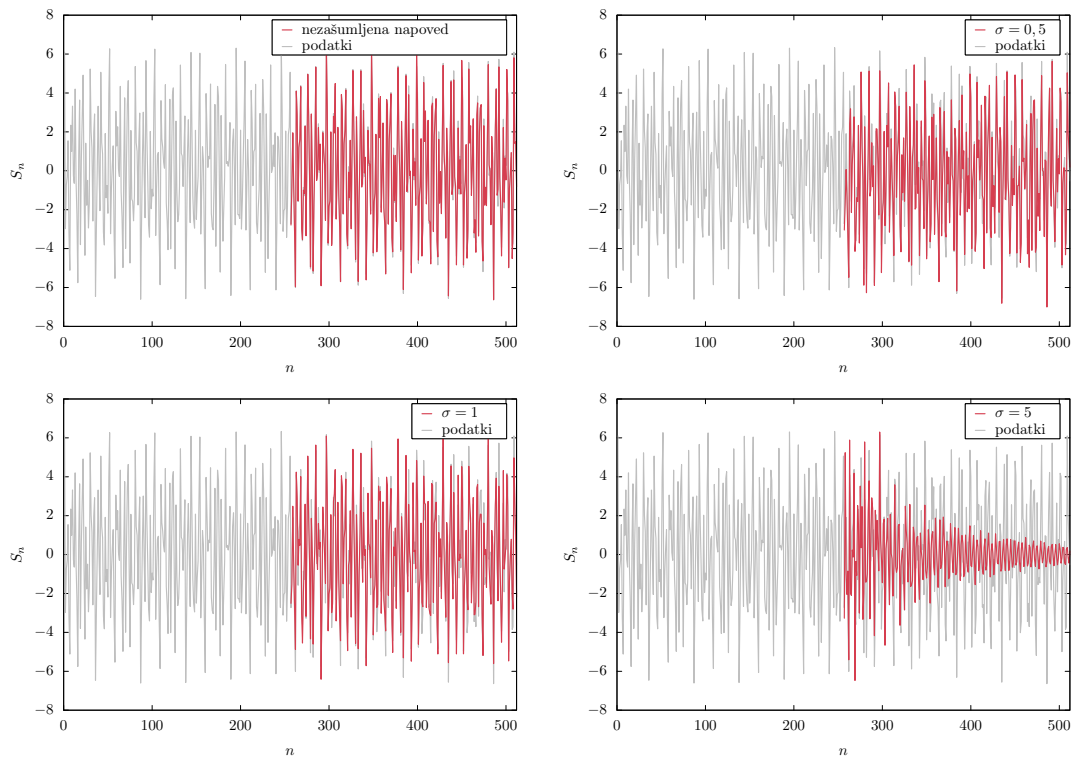
Poglejmo na primeru datoteke `val3.dat` kako vpliva dodatek šuma na linearno napoved. Na grafih 3.10 lahko vidimo, da začne napoved pri sigma 1 precej odstopati od pravih vrednosti, če jo povečujemo pa postane to še bolj izrazito. Poskusimo še povečati število polov  $p$ . Na grafih 3.11 lahko opazimo, da s povečanjem števila polov  $p$  izboljšamo amplitudo napovedi.



Slika 3.9: Linearna napoved podatkov iz borze pri različnih  $p$ .



Slika 3.10: Linearna napoved nezašumljenih podatkov val3 in zašumljenih z Gaussovskim šumom z različnimi  $\sigma$  pri  $p = 20$ .



Slika 3.11: Linearna napoved nezašumljenih podatkov val3 in zašumljenih z Gaussovskim šumom z različnimi  $\sigma$  pri  $p = 50$ .