

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
ODDELEK ZA FIZIKO

MODELSKA ANALIZA I

6. naloga: Luščenje modelskih parametrov in razdelčni modeli

Žiga Šinigoj, 28222025

Ljubljana, november 2022

1 Uvod

Parametre modela s katerim opišemo izmerke določim tako, da bo funkcional χ^2 minimalen

$$\chi^2 = \sum_{i=1}^N \left(\frac{y_i - y(x_i, \mathbf{a})}{\sigma_i} \right)^2, \quad (1)$$

kjer je \mathbf{a} vektor modelskih parametrov. Če je funkcijska zveza linearna v parametrih \mathbf{a} , lahko funkcional minimiziram z reševanjem normalnega ali razširjenega sistema, ali uporabe SVD razcepa. Nelinearne probleme lahko rešim z metodo *Levenberg – Marquardt*.

2 Farmakologija

V farmakologiji merijo odziv tkiv na različne reagente. Za večino teh pojavov lahko privzamemo, da gre za reakcijo, kjer spremljamo vezavo molekul reagenta X na receptorje Y v tkivu.



V stacionarnem stanju velja

$$y = \frac{y_0 x}{x + a}, \quad (3)$$

ker je zveza linearna v parametrih, jo lahko lineariziram kot

$$\frac{1}{y} = a_1 + a_2 \frac{1}{x}, \quad (4)$$

kjer je $a_1 = 1/y_0$, $a_2 = a/y_0$. Napake σ_i se pri tem transformirajo kot $\tilde{\sigma}_i = \frac{\sigma_i}{y_i^2}$. Funkcijo minimiziramo, ko velja $\frac{\partial \chi^2}{\partial a_j} = 0$, kar se privede na reševanje sistema

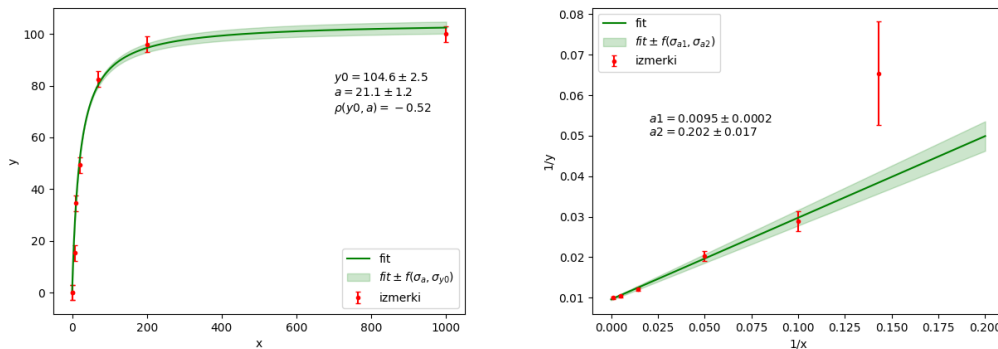
$$A\mathbf{a} = \mathbf{b} . \quad (5)$$

Poleg parametrov \mathbf{a} dobimo lahko tudi kovariančno matriko iz inverza A

$$\text{var}[a_j] = (A^{-1})_{jj}, \quad \text{cov}[a_j, a_k] = (A^{-1})_{jk}, \quad (6)$$

in korelacijski koeficient kot $\rho = \frac{\text{cov}[a_j, a_k]}{\sqrt{\text{var}[a_j]\text{var}[a_k]}}$.

Optimalne parametre sem dobil z reševanjem sistema in z funkcijo *curve_fit* iz paketa *scipy.optimize*, ki uporablja metodo *Levenberg – Marquardt*.



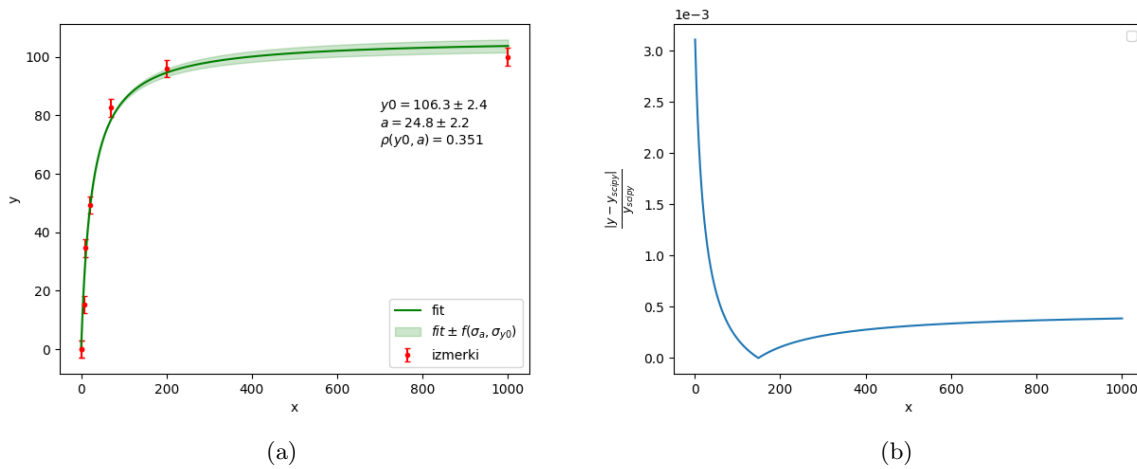
Slika 1: a) Optimalen fit v prostoru x-y. b) Optimalen fit po enačbi 4, brez zadnjih dveh točk. Zelen pas predstavlja družino krivulj v intervalu napak.

Najprej sem poiskal optimalna parametra a_1 in a_2 z reševanjem sistema velikosti 2×2 , družino krivulj, ki so v intervalu napak ($\pm f(\sigma_i, \sigma_j)$) prikazuje slika 1 v prostoru, kjer je problem linearen in v prostoru x-y. Ko dobim rešitev za $a_1 \pm \sigma_1$ in $a_2 \pm \sigma_2$, lahko iz enačb, ki definirajo a_1 in a_2 dobim iskana parametra y_0 in a . Pri tem pazim, da transformiram tudi njihovo napako

$$(y_0 \pm \sigma_0) = \frac{1}{a_1 + \sigma_{a1}} = \frac{1}{a} (1 \mp \frac{\sigma_{a1}}{a_1}) \rightarrow \sigma_0 = \frac{\sigma_{a1}}{a_1^2}, \quad (7)$$

$$\sigma_a = \frac{\sigma_{a2} - \sigma_{y0} \frac{a_2}{y_0} a_1}{a_1 (1 + \frac{\sigma_{y0}}{y_0})}. \quad (8)$$

Te napake so rezultat v 1(a), $\chi^2 = 6.24$. Ker imata zadnji točki zelo veliko napako, ne prispevata veliko k χ^2 . V ta namen sem poiskal rešitev še z nelinearno minimizacijo (slika 2(a)). Vrednosti parametrov se malo razlikujejo, funkcional $\chi^2 = 22.04$ je precej večji kot pri reševanju sistema. Rešitvi sta na pogled identični, zato je smiselno pogledati relativno razliko (slika 2(b)), ki potrjuje natančnost metode z reševanjem sistema enačb.



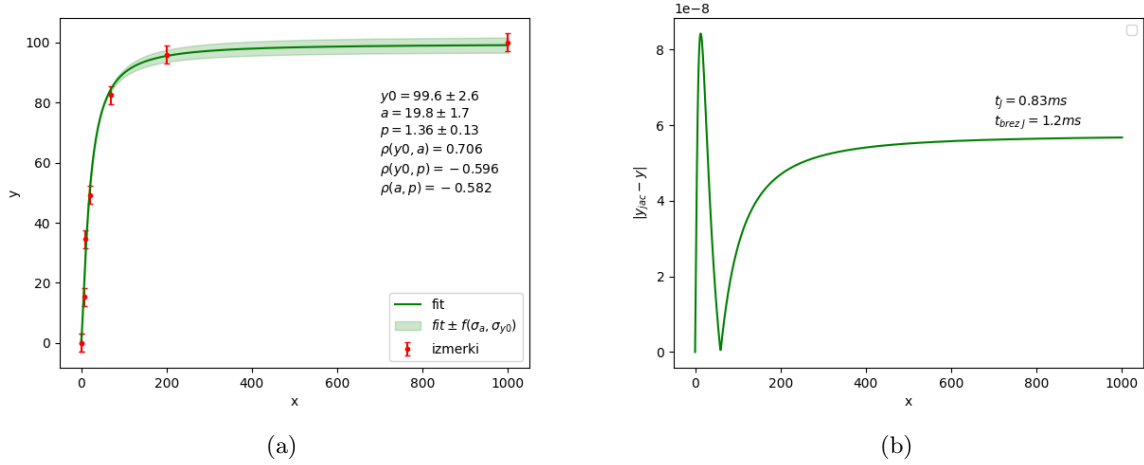
Slika 2: a) Optimalen fit v prostoru x-y z uporabo funkcije *curve_fit*. Zelen pas predstavlja družino krivulj v intervalu napak. b) Relativna razlika med pridobljeno modelsko rešitvijo z reševanjem sistema in s *Scipyjevo* funkcijo *curve_fit*.

2.1 Razširjen model

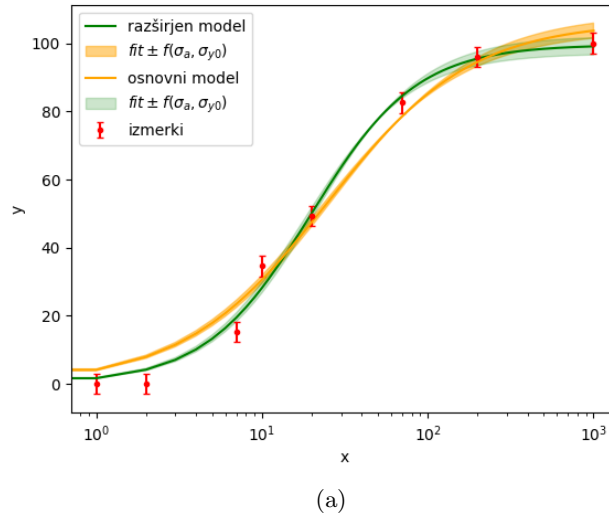
Velika vrednost χ^2 v primeru z nelinearnim reševanjem bi lahko napeljala na idejo, da obstaja boljši model. Osnovni model nadgradimo z modelom

$$y = \frac{y_0 x^p}{x^p + a^p}, \quad (9)$$

ki ni več linearen v parametrih. Sedaj lahko optimiziram parametre z metodo *Levenberg – Marquardt*, ki je privzeta metoda funkcije *curve_fit*. V splošnem moramo metodi podati tudi odvode po parametrih. Če odvodov ne podamo, potem bi algoritem sam poiskal lokalne odvode s končnimi diferencami. Rezultat s podanimi odvodi je prikazan na sliki 4. Zanimalo me je, kako podani analitični odvodi prispevajo k natančnosti in časovni zahtevnosti (slika 4(b)), kot vidim je algoritem zelo natančen tudi brez podanih odvodov. Časovna zahtevnost postane večja če mu ne podamo odvodov po parametrih, v našem preprostem primeru se podaljša čas izvajanja za 0.4 ms. Vrednost funkcionala $\chi^2 = 9.357$ je manjša kot v primeru enostavnega modela. Primerjavo obeh modelov prikazuje slika 5, že iz slike v logaritemski skali je razvidno, da se razširjen model bolje prilega podatkom.



Slika 3: a) Prilagajanje nelinearnega modela. Zelen pas predstavlja družino krivulj v intervalu napak. b) Absolutna razlika med rešitvami, katerim podamo odvode po parametrih in rešitvami brez podanih odvodov funkciji *curve_fit*.



Slika 4: Enostavni in razširjeni model z intervali napak v logaritemski skali.

3 Čistilnost ledvic

Čistilnost ledvic lahko modeliramo z različnimi modeli. Najprej si pogledjmo prilagajanje enodelčnega modela podatkom.

3.1 Enorazdelčni model

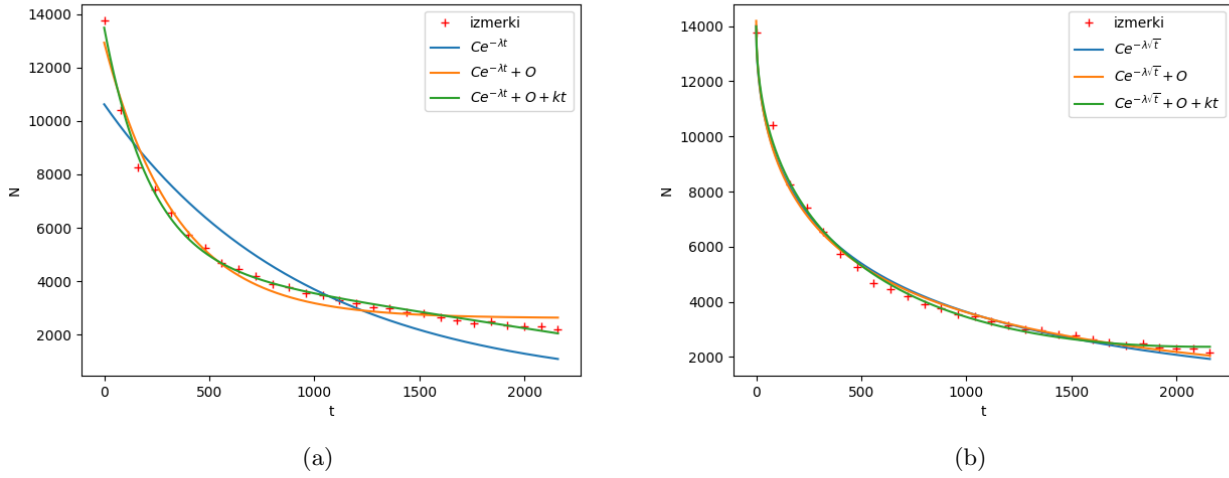
Podatkom sem prilagajal modele, kjer je čas teče linearно in korensko v eksponentu. Vse sem prilagajal s funkcijo *curve_fit*, saj se lahko linearizira samo eksponentno funkcijo brez ozadja. Najprej sem prilagajal modele

$$Ce^{-\lambda t}, \quad Ce^{-\lambda t} + O, \quad Ce^{-\lambda t} + kt + O, \quad (10)$$

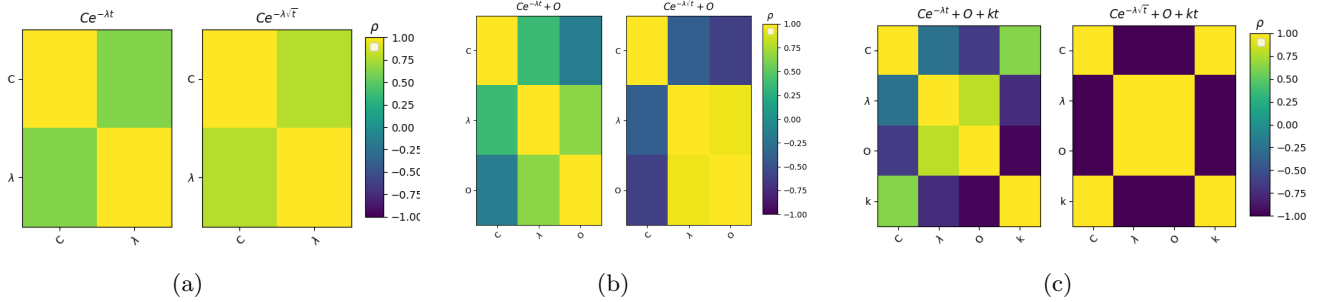
ki so prikazani na sliki 5(a). Model z dvema prostima parametroma se slabo prilega podatkov. Če dodamo prost parameter za ozadje dobimo precej dobro ujemanje. Po navodilih sem prilagajal še modele

$$Ce^{-\lambda \sqrt{t}}, \quad Ce^{-\lambda \sqrt{t}} + O, \quad Ce^{-\lambda \sqrt{t}} + kt + O. \quad (11)$$

Modeli s korensko funkcijo časa v eksponentu se bolje prilagajajo (slika 5(b)) kot prejšnji modeli, sploh model z dvema parametroma. Korelacije med parametri prikazuje slika 6. Korelacija med parametri modela je večja v primeru ko imamo korensko odvisnost časa v eksponentu (slika 6). Največjo korelacijo med parametri kaže korenski model s štirimi prostimi parametri.



Slika 5: Prilagajanje enorazdelčnih modelov k podatkom.



Slika 6: Korelacije med parametri. a) Korelacije pri dvoparametričnih modelih. b) Korelacije pri troparametričnih modelih. c) Korelacije pri štiriparametričnih modelih

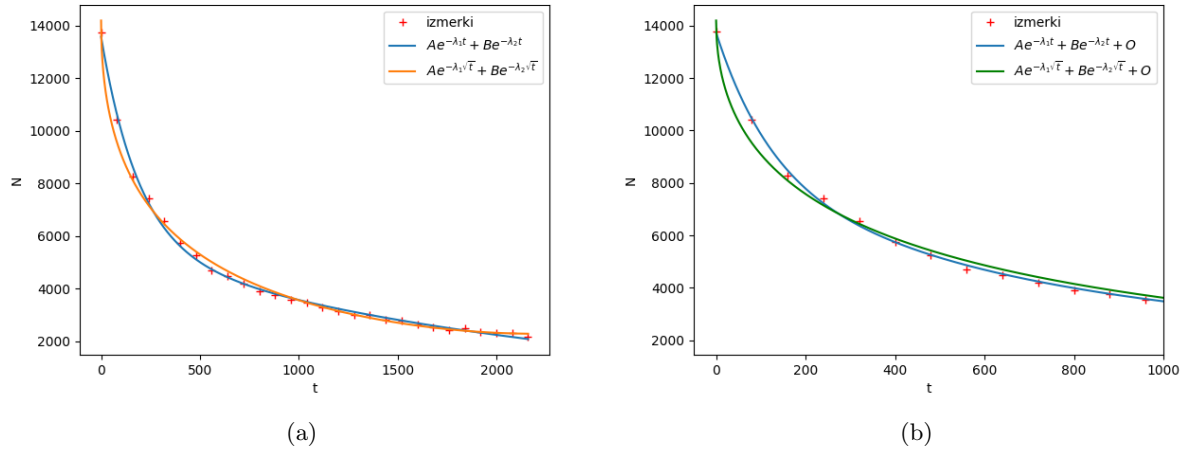
3.2 Dvorazdelčni model

Za boljši opis čistilnosti ledvic uporabim dvodelčni model. Funkcije ki sem jih prilagajal podatkom so 4 in 5 parametrične

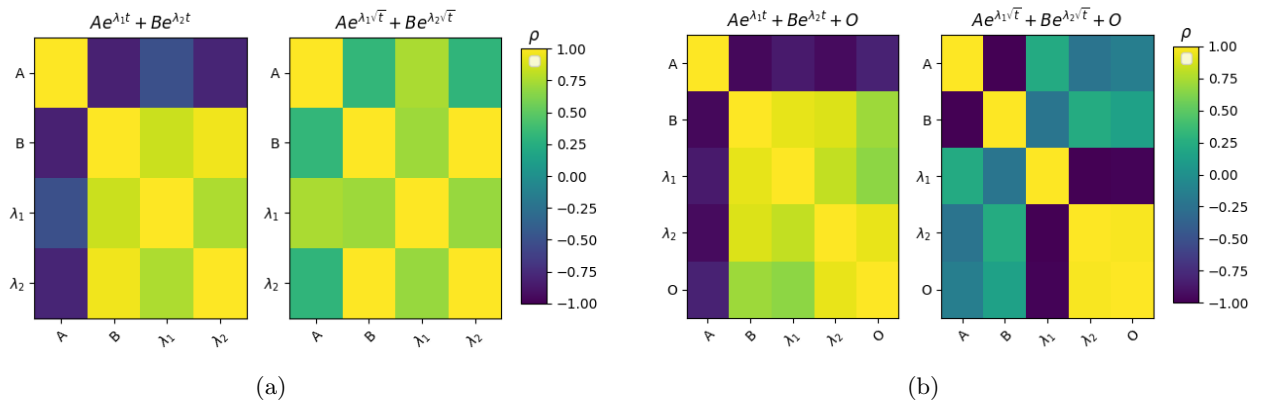
$$Ae^{-\lambda_1 t} + Be^{-\lambda_2 t}, \quad Ae^{-\lambda_1 \sqrt{t}} + Be^{-\lambda_2 \sqrt{t}}, \quad (12)$$

$$Ae^{-\lambda_1 t} + Be^{-\lambda_2 t} + O, \quad Ae^{-\lambda_1 \sqrt{t}} + Be^{-\lambda_2 \sqrt{t}} + O. \quad (13)$$

Prileganje funkcij podatkom prikazuje slika 7. Na pogled se zdi, da se 4 parametrične funkcije bolje prilagajajo kot 5 parametrične. Slika 7(b) je približana k začetnim meritvam, saj lahko tam vidim odstopanja modela s korensko odvisnostjo časa, ki namiguje na to, da ima model preveč parametrov. Korelacije med prostimi parametri so prikazane na sliki 8. Korelacije so v dvorazdelčnih modelih večje v primeru ko je čas linearna funkcija v eksponentu, kar ni veljalo pri enodelčnih modelih. Za določitev najboljšega približka je najbolje, da primerjam vrednosti χ^2 (tabeli 1,2). Najmanjšo vrednosti funkcionala ima model dvorazdelčni model $Ae^{-\lambda_1 t} + Be^{-\lambda_2 t} + O$ in se najboljše prilega podatkom.



Slika 7: Prilagajanje dvorazdelčnih modelov k podatkom.



Slika 8: Korelacije med parametri. a)Korelacije pri štiriparametričnih modelih. b)Korelacije pri petparametričnih modelih.

$Ce^{-\lambda t}$	$Ce^{-\lambda t} + O$	$Ce^{-\lambda t} + kt + O$	$Ce^{-\lambda \sqrt{t}}$	$Ce^{-\lambda \sqrt{t}} + O$	$Ce^{-\lambda \sqrt{t}} + kt + O$
$26.7 \cdot 10^6$	$3.02 \cdot 10^6$	$622.4 \cdot 10^3$	$1.8 \cdot 10^6$	$1.6 \cdot 10^6$	$866 \cdot 10^3$

Tabela 1: Vrednosti χ^2 pri različnih modelih.

$Ae^{-\lambda_1 t} + Be^{-\lambda_2 t}$	$Ae^{-\lambda_1 \sqrt{t}} + Be^{-\lambda_2 \sqrt{t}}$	$Ae^{-\lambda_1 t} + Be^{-\lambda_2 t} + O$	$Ae^{-\lambda_1 \sqrt{t}} + Be^{-\lambda_2 \sqrt{t}} + O$
$442 \cdot 10^3$	$1.3 \cdot 10^6$	$208 \cdot 10^3$	$1.6 \cdot 10^6$

Tabela 2: Vrednosti χ^2 pri različnih modelih.

4 Magnetni spektrometer

Za uporabo visokoločljivostnega magnetnega spektrometra potrebujem preslikavo, ki iz izmerjenih količin rekonstruira parametre trajektorije delcev, potrebne za izračun energije in drugih kinematičnih količin. Vhodni podatki so kot detektorja ϑ_{det} , položaj na detektorju x_{det} in kot na tarči ϑ_{tg} . Najprej sem za varčni model vzel preslikavo v kateri nastopata ϑ_{det} , x_{det} linearno

$$\vartheta_{tg} = a_0 + a_1 x_{det} + a_2 \vartheta_{det} + a_3 x_{det} \vartheta_{det} , \quad (14)$$

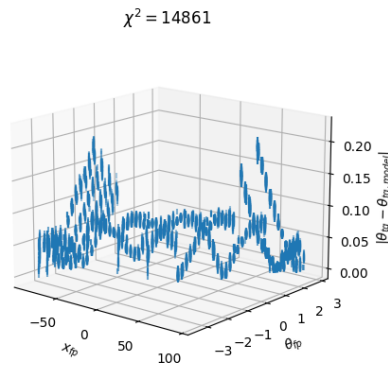
ker je sistem v tem primeru predoločen, algoritem *Levenberg – Marquardt* ne deluje več dobro, zato uporabim SVD razcep. Vektor iskanih parametrov $\mathbf{a} = (a_0, a_1, a_2, a_3)$ dobim kot

$$\mathbf{a} = \sum_{i=0}^{M-1} \left(\frac{\mathbf{U}_i \mathbf{b}}{w_i} \right) \mathbf{V}_i \quad (15)$$

, kjer je M število prostih parametrov, \mathbf{U} in \mathbf{V} matriki iz SVD razcepa, w_i singularne vrednosti razcepa in \mathbf{b} vektor izmerkov v tarči. Velja še

$$var[a_j]^2 = \sum_{i=0}^{M-1} \frac{V_{ij}^2}{w_i^2}, \quad cov[a_j, a_k] = \sum_{i=0}^{M-1} \frac{V_{ij} V_{ik}}{w_i^2} \quad (16)$$

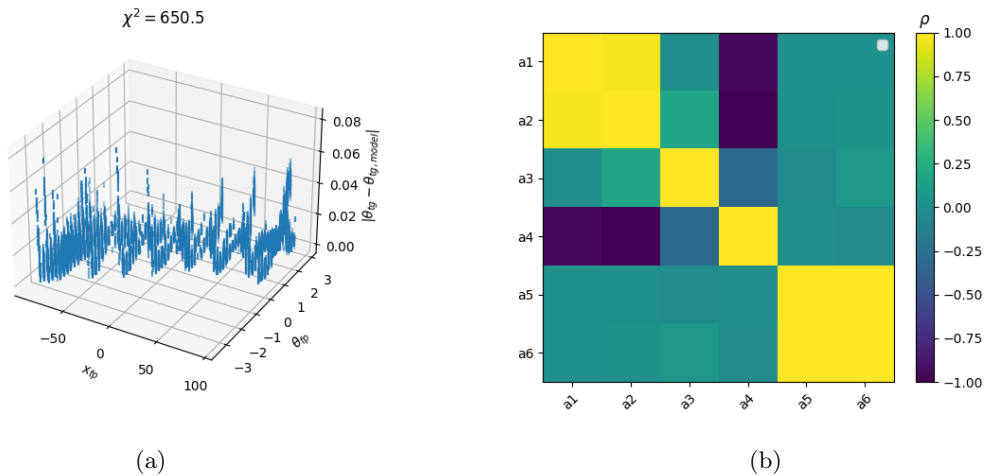
Odstopanje kotov modela 14 od pravih kotov na tarči prikazuje slika 9, odstopanja niso velika, največja so okrog 3%. Že preprost model s 4 parametri naredi dobro preslikavo. Vseeno je vrednost funkcionala χ^2 precej visoka, zato je smiselno nadgraditi model.



Slika 9: Absolutna odstopanja modelskih kotov od pravih.

Kot nadgradnjo sem modelu dodal še člene, kjer nastopa x_{det}^2 .

$$\vartheta_{tg} = a_0 + a_1 x_{det} + a_2 \vartheta_{det} + a_3 x_{det} \vartheta_{det} + a_4 x_{det}^2 + a_5 x_{det}^2 \vartheta_{det}, \quad (17)$$



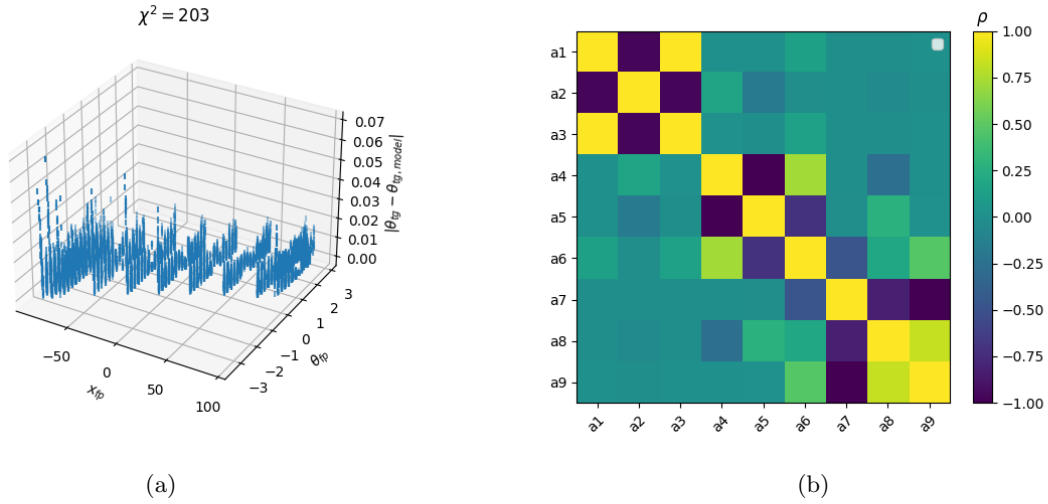
Slika 10: a) Absolutna odstopanja modelskih kotov od pravih. b) Korelacije med modelskimi parametri.

Odstopanja in korelacije modela 17 prikazuje slika 10. Z dodatkom kvadratnega člena sem dobil bolj natančen model, največja odstopanja so okrog 1%, vrednost funkcionala se je zmanjšala za skoraj 23-krat. Dodajamo

lahko poljubno členov, vendar ni nujno da bomo dobili boljši model, če dodamo več prostih parametrov. Ker je korelacija med nekaterimi členi velika, sem naivno poskusil odstraniti enega od členov in namesto njega dodati višji red, a to ne privede do boljše rešitve. V izrazu 17 lahko opazim, da potence x_{det} in ϑ_{det} ne nastopajo simetrično, v ta namen sem dodal še člene z ϑ_{det}^2 in s tem simetriziral izraz

$$\vartheta_{tg} = a_0 + a_1 x_{det} + a_2 \vartheta_{det} + a_3 x_{det} \vartheta_{det} + a_4 x_{det}^2 + a_5 x_{det}^2 \vartheta_{det} + a_6 \vartheta_{det}^2 + a_7 \vartheta_{det}^2 x_{det} + a_8 x_{det}^2 \vartheta_{det}^2, \quad (18)$$

Odstopanja in korelacije modela 18 prikazuje slika 11. Vrednost χ^2 je manjša kot v prejšnjem modelu, kar nakazuje na to, da je model boljši. Odstopanja od pravih vrednosti kotov so manjša (ni več toliko posameznih ekstremnih odstopanj (šipce)). V tem modelu so parametri bolj korelirani.



Slika 11: a) Absolutna odstopanja modelskih kotov od pravih. b) Korelacije med modelskimi parametri.

Seveda bi lahko dodajal še člene in iskal boljši model, vendar je pri dodajanju pomembno da upoštevamo simetrije, če te izboljšajo model. Če bi samo dodajal potence ϑ_{det} in x_{det} v model, ne bi dobil boljše rešitve. Mogoče bi bilo smiselno pogledati modele, ki jih dobim iz vsote izrazov n-tih potenc dvočlenika $(x_{det} + \vartheta_{det})^n$, kjer bi pred vsak člen dodal prosti parameter.

5 Zaključek

V nalogi sem prilagajal različne modele na sete podatkov. Uporabil sem 3 različne metode: eksaktno reševanje sistema, metodo lm in SVD razcep. Pri vsakem prilagajanju modela sem pogledal tudi korelacije in vrednosti funkcionala χ^2 . Težava pri prilagajanju modela je vsekakor optimalno število parametrov, ki najbolj opišejo model. Velikokrat se zdi, da bo več parametrov boljše opisalo realni problem, vendar ni vedno tako.