

## 6. naloga – Varčni modeli

Jože Zobec

### 1 Uvod

Teoretične parametre se v fiziki iz eksperimentov (slednji so lahko tudi numerični) po navadi prebere iz prilagajanja modelske krivulje na merske točke. Krivuljo prilagajamo tako, da za določen nabor parametrov zadostimo arbitrarnim (po navadi minimizacijskim) zahtevam. Najbolj popularen kriterij je ti.  $\chi^2$ , ki parametre minimizira tako, da je vsota napak za posamezne merske točke čim manjša. Tudi za prilagajanje imamo na voljo dovolj širok nabor metod, med najbolj znanimi je linearna regresija, katero bomo uporabili v tej nalogi.

Linearna regresije prek  $\chi^2$  skuša prilagoditi vektor parametrov  $\beta$  tako, da bo enačba  $y = X\beta + \varepsilon$  imela čim manjši vektor napak  $\varepsilon$ . To naredimo prek minimizacije  $\chi^2$ , ki je kar kvadratna forma

$$\chi^2 = 2J = (y - X\beta)^T R^{-1}(y - X\beta). \quad (1)$$

Zahteva, da je  $\chi^2$  minimalen vrne, da so idealni parametri

$$\beta = (X^T R^{-1} X)^{-1} X^T R^{-1} y, \quad P = (X^T R^{-1} X)^{-1}. \quad (2)$$

kjer je  $\beta$  vektor parametrov našega modela,  $X$  je strukturna matrika, ki  $\beta$  povezuje z  $y$ , matrika  $R$  je kovariančna matrika merskih točk (kadar so slednje nekorelirane je to diagonalna matrika kvadratov napak posameznih točk), matrika  $P$  pa je kovariančna matrika optimalnih parametrov  $\beta$ . Enačbi (1) in (2) sta primer splošne linearne regresije (zato, ker dopuščamo, da so merske točke lahko korelirane in vsaka pri drugačni napaki – tj. matrika  $R$  ima lahko karseda splošno obliko).

Kaj pa, kadar nam teorija ne napoveduje oblike krivulje? V takem primeru enostavno poskušamo, vendar se vsej ko prej lahko zgodi, da bo naša modelska krivulja imela parametre, ki nimajo fizikalnega pomena – red našega modela je prevelik. Hočemo seveda, da bo naš model karseda varčen – tj. imel čim manj parametrov. To dosežemo s pomočjo singularnega razcepa matrike  $X$ . Singularne vrednosti, ki so zelo majhne, namigujejo na to, da imamo preveč prostih parametrov v modelu in da niso vsi parametri linearno neodvisni.

### 2 Razmislek

V prvem primeru (tj. toplotna prevodnost) prilagajamo testno „krivuljo“  $\lambda(T, P)$  in skušamo poiskati tako, ki bo imela najmanj parametrov in najboljši  $\chi^2$ . Naš model bo največ tretjega reda, tj.

$$\lambda = \sum_{i,j=0}^3 \beta_{ij} T^i P^j.$$

Parametrov  $\beta_{ij}$  je največ 10, nekateri izmed teh parametrov bodo nič (oz. jih zavržemo), torej lahko strukturno matriko  $X$  računamo le enkrat (ta matrika je dimenzije  $n \times 10$ ,  $n$  je število merskih točk) nato pa zavračamo stolpce. Največ lahko zavržemo 9 stolpcev (če zavržemo vseh 9, potem nima smisla), najmanj pa nobenega, torej dobimo  $G$  načinov, kako lahko to storimo:

$$G = \sum_{k=0}^9 \binom{10}{k} = 1023.$$

To ni tako veliko. V tej nalogi bom skušaj izračunati vse kombinacije (točk je malo in  $\sim 10^3$  množenj majnih matrik ni tako potratno) in potem nadaljeval z najbolj zanimivimi.

Sedaj pridemo k drugemu problemu – kako poiskati oz. predstaviti in klasificirati vse možne kombinacije? To je čisto preprosto. Vsako izbiro lahko predstavimo z binarnim številom – vsak stolpec iz matrike  $X$  bodisi izberemo, bodisi zavržemo. Ker je teh stolpcev 10 dobimo dvojiško število dolžine 10. Od tod tudi pride, da je število kombinacij 1023, saj

$$G = 2^n - 1,$$

kjer ‘ $-1$ ’ pride zato, ker kombinacije  $(00 \dots 0)$  ne upoštevamo, saj ta nima smisla. Vendar ta dvojiška števila lahko enolično pretvorimo nazaj v desetiška in dobimo indeks izbire  $X$ , ki je ravno med 1 in 1023. Ta preslikava je bijektivna, tj. smo s tem posredno poiskali vse možne kombinacije.

Kadar uporabimo  $\chi^2$  za kvantitativno mero ujemanja našega modela danimi eksperimentalnimi točkami, je predvsem prikladno uporabiti  $\overline{\chi^2}$ , tj. reduciran  $\chi^2$ , ki je definiran kot

$$\overline{\chi^2} \equiv \frac{\chi^2}{n - k - 1}, \quad (3)$$

kjer je  $n$  število merskih točk,  $k$  pa je število parametrov  $\beta_{ij}$  ki so različni od 0 (spet torej koliko stolpcev smo sprejeli). Količina  $\overline{\chi^2}$  ima predvsem to lepo lastnost, da je za krivulje, ki se boljše ujemajo s točkami, bližje 1.

Ko bomo izračunali za vse možne primere izbire  $X$  vse možne  $\chi^2$ , bomo izluščili primere, ki imajo  $\overline{\chi^2} \sim 1$  in napravili dodatno analizo, tj. preverili število prostih parametrov krivulje in napravili singularni razcep matrike  $X$  ter pod drobnogled vzeli kovariančno matriko  $P$ .

Isto lahko uporabimo za kiralnost, le da imamo tam drugačno totalno matriko  $X$  in drugačno matriko  $R$ .

## 3 Rezultati

### 3.1 Toplotna prevodnost jekla Armco

Tabela 1 prikazuje toplotno prevodnost  $\lambda$  jekla Armco v odvisnosti od temperature in moči grelca. Temperaturo kot moč bi lahko reskalirali tako, da bi bila med 0 in 1, vendar tu niti ni take potrebe (brezdimenzijsko transformacijo se po navadi uvaža v sisteme takrat, kadar hočemo dobiti rešitev neodvisno od skale problema, te pa v tem primeru ni).

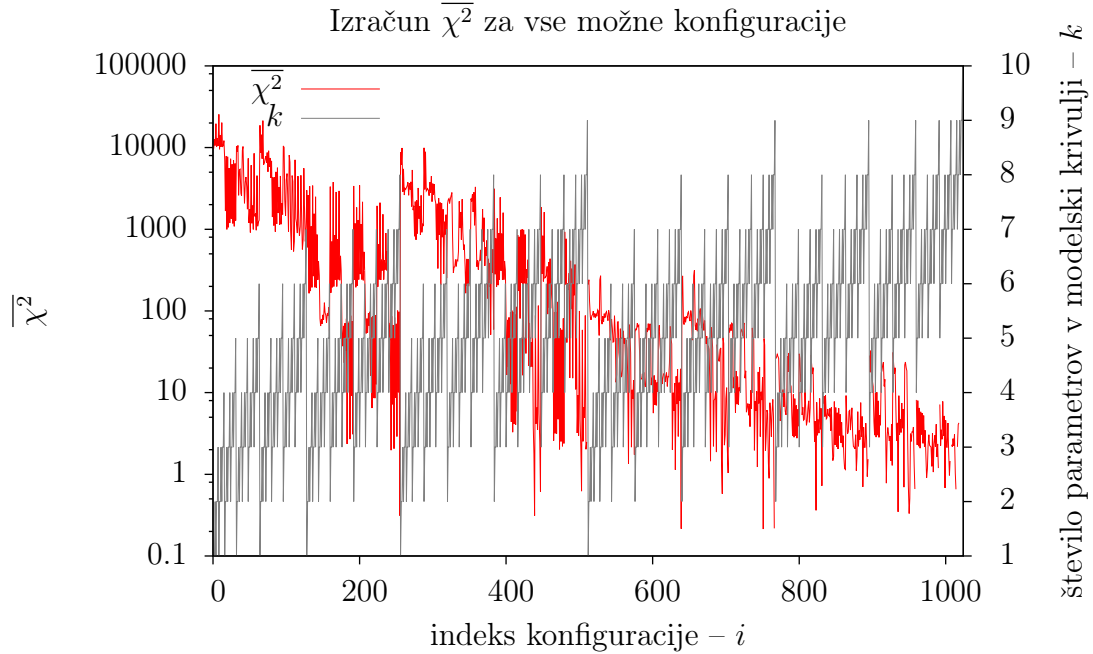
Tabela 1: Tabela meritev za toplotno prevodnost jekla Armco.

| $T[^\circ\text{F}]$ | $P[\text{W}]$ | $\lambda[\text{Btu/h ft } ^\circ\text{F}]$ | $\sigma_\lambda$ |
|---------------------|---------------|--|------------------|
| 90                  | 276           | 42.345                                     | 0.28             |
| 100                 | 545           | 41.60                                      | 0.16             |
| 149                 | 275           | 39.5375                                    | 0.28             |
| 161                 | 602           | 37.7875                                    | 0.16             |
| 206                 | 274           | 37.3525                                    | 0.28             |
| 227                 | 538           | 36.4975                                    | 0.16             |
| 247                 | 274           | 36.36                                      | 0.28             |
| 270                 | 550           | 35.785                                     | 0.16             |
| 352                 | 272           | 33.915                                     | 0.28             |
| 362                 | 522           | 34.53                                      | 0.16             |

Imamo 10 merskih točk, torej imamo največ 10 prostih parametrov za naš model, kar je v skladu s prejšnjim razmislekom. Stolpci maksimalne matrike  $X$  so

$$X = \begin{bmatrix} 1 & | & T & | & P & | & T^2 & | & TP & | & P^2 & | & T^3 & | & T^2P & | & TP^2 & | & P^3 \end{bmatrix}.$$

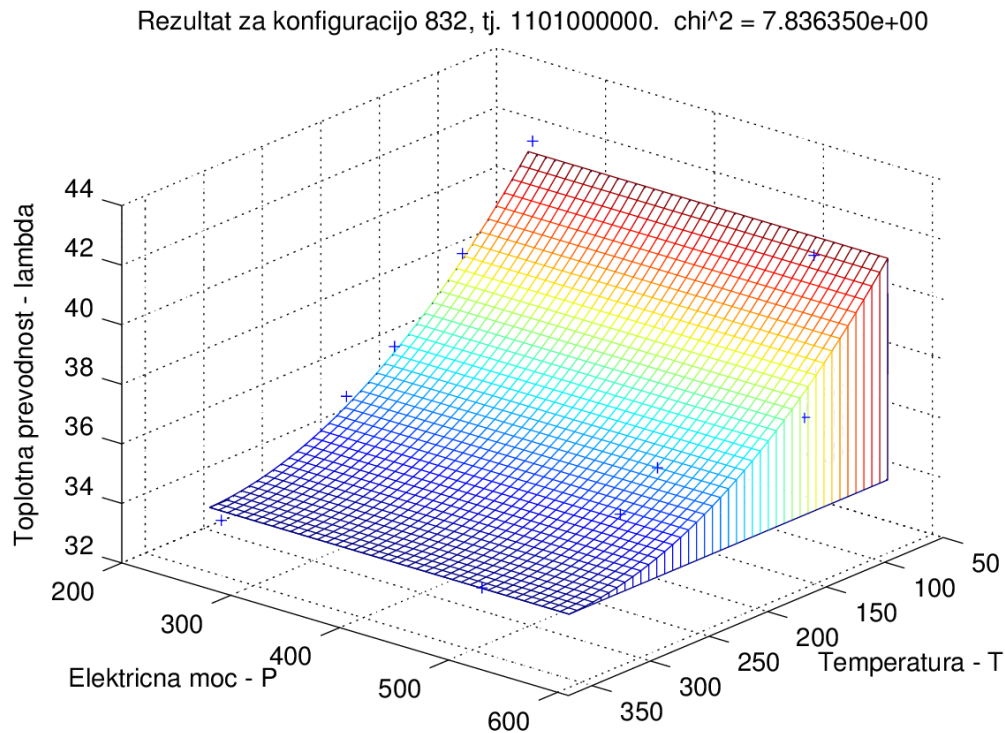
Konfiguracija (1001001000) predstavlja da v  $X$  obdržimo konstanto,  $T^2$  in  $T^3$ , tj. da za model vzamemo  $\lambda \approx \beta_{00} + \beta_{20}T^2 + \beta_{30}T^3$  in je oštevilčena z indeksom  $2^9 + 2^6 + 2^3 = 584$ . Matrika  $R$  je diagonalna,  $R = \text{diag}(\sigma_{\lambda_1}^2, \sigma_{\lambda_2}^2, \dots, \sigma_{\lambda_{10}}^2)$ , vektor  $y$  pa vsebuje  $\lambda_i$ , kjer je  $i$  indeks meritve. Sedaj imamo vse kar potrebujemo za izračun.



Slika 1: Konstantni člen vstopi v konfiguraciji 512. Vidimo, da lahko tudi brez konstantnega člena dobimo dobre rezultate, vendar so redki. S konstantnim členom je  $\chi^2$  takoj nižji za velikostni red.

Izmed vseh teh možnosti sem se nato odločilo, da bom zavrgel vse, za katere je  $\overline{\chi^2} < 0.5$ ,  $\overline{\chi^2} > 8$ ,  $k > 7$ .

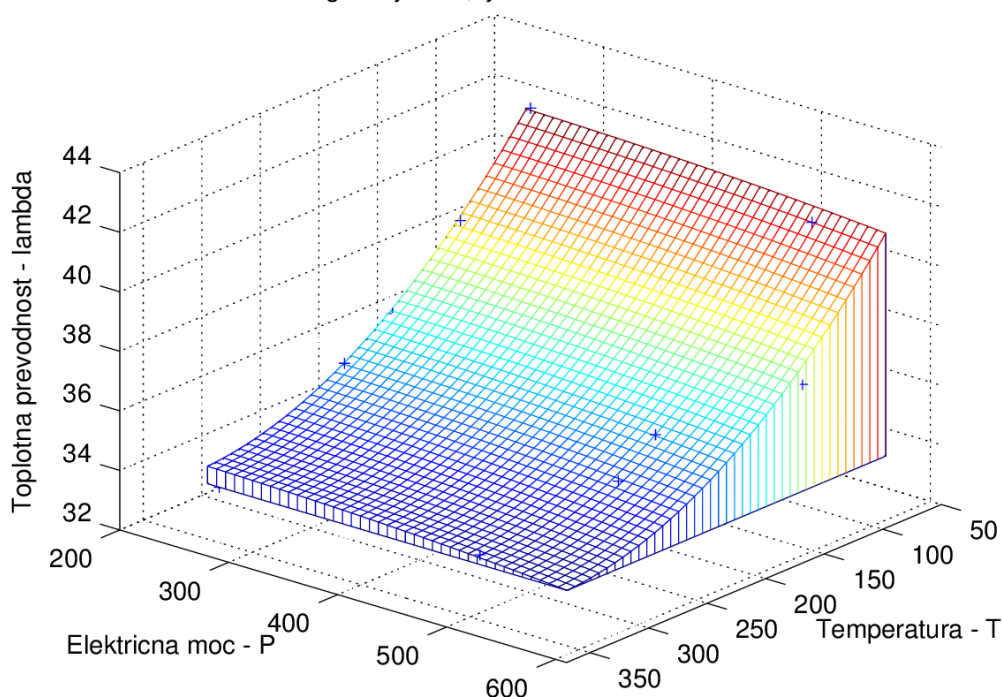
Model z najmanjšim  $k$ , ki ustreza temu kriteriju ima  $k = 3$  in nosi indeks  $i = 832$ . Model je  $\lambda \approx \beta_{00} + \beta_{10}T + \beta_{20}T^2$  in je na sliki 2.



Slika 2: Odvisnost od  $P$  v tem primeru zanemarimo, in upoštevamo parabolično odvisnost od  $T$ .

Temu soroden je model 833, ki predpostavlja  $\lambda \approx \beta_{00} + \beta_{10}T + \beta_{20}T^2 + \beta_{03}P^3$ . Vidimo ga na sliki 3.

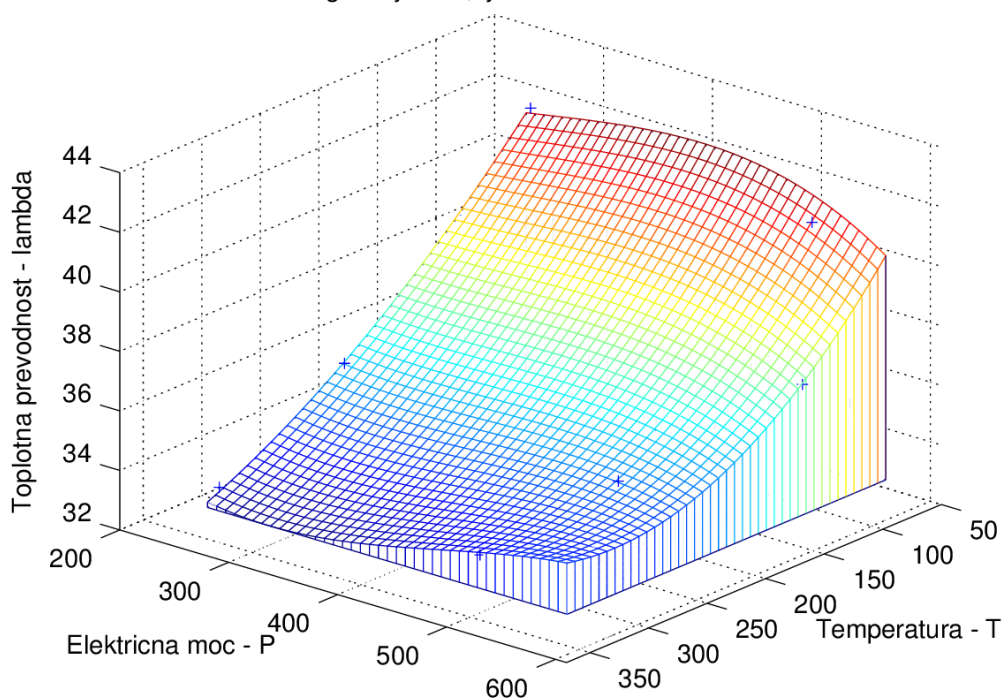
Rezultat za konfiguracijo 833, tj. 1101000001.  $\chi^2 = 7.353968e+00$



Slika 3: V tem modelu imamo splošno parabolično odvisnost od temperature in strogo kubično od moči grelca. Ujemanje je kar solidno, vendar se ni za prehvaliti.

Graf, ki ima  $\overline{\chi^2}$  najbližje 1, je  $i = 567$  s  $\overline{\chi^2} = 1.35$ . Tu imamo model  $\lambda \approx \beta_{00} + \beta_{11}TP + \beta_{02}P^2 + \beta_{21}T^2P + \beta_{12}TP^2 + \beta_{03}P^3$ , vendar ima eno izmed singularnih vrednosti precej majhno. Singularne vrednosti  $X$  za ta model so  $S = \text{diag}(4.19 \cdot 10^8, 8.56 \cdot 10^7, 1.25 \cdot 10^7, 8.48 \cdot 10^4, 1.82 \cdot 10^4, 2.86 \cdot 10^{-1})$ . Rezultat je na sliki 4.

Rezultat za konfiguracijo 567, tj. 1000110111.  $\chi^2 = 1.353947e+00$



Slika 4: Čeprav ima zelo dobro ujemanje v  $\overline{\chi^2}$  ta graf mogoče ni najboljši sodeč po singularnih vrednostih.

Grafov, ki so ustrezajo prejšnjemu kriteriju je 152, kar je preveč. Uvedel sem nov pogoj: vse singularne vrednosti morajo biti večje od 1. To nam vrne 18 grafov, od katerih so vsi (zanimivo) brez konstantnega člana in imajo bodisi 5, bodisi 6 elementov. Te konfiguracije so v tabeli 2.

Tabela 2: Tu vidimo modele, ki so bili na koncu sprejeti. Zadnji stolpec predstavlja najmanjšo singularno vrednost.

| konf. | $\overline{\chi^2}$ | $k$ | $s_{\min}$ |
|-------|---------------------|-----|------------|
| 183   | 2.36                | 6   | 40.15      |
| 187   | 2.69                | 6   | 40.66      |
| 217   | 6.48                | 5   | 56.38      |
| 243   | 1.98                | 6   | 38.93      |
| 245   | 7.19                | 6   | 41.96      |
| 407   | 4.38                | 6   | 35.55      |
| 409   | 3.28                | 5   | 46.16      |
| 411   | 4.13                | 6   | 35.04      |
| 413   | 4.05                | 6   | 35.31      |
| 437   | 3.83                | 6   | 35.66      |
| 438   | 1.78                | 6   | 14.95      |
| 441   | 4.17                | 6   | 32.54      |
| 442   | 6.55                | 6   | 16.75      |
| 465   | 2.58                | 5   | 35.23      |
| 467   | 3.17                | 6   | 29.45      |
| 469   | 3.15                | 6   | 32.08      |
| 473   | 2.02                | 6   | 7.81       |
| 497   | 3.21                | 6   | 27.64      |

Iz tabele 2 nato izberemo graf, ki ima najmanj parametrov in je najbližlje  $\overline{\chi^2} = 1$  in dobimo konfiguracijo 465. Ta najbolje zadošča kriteriju po čim manjšem številu parametrov  $k$ . Ta model je

$$\lambda = \beta_{10}T + \beta_{01}P + \beta_{20}T^2 - \beta_{02}P^2 + \beta_{03}P^3, \quad (4)$$

parametri z napakami (ki smo jih prebrali iz kovariančne matrike) so v tabeli 3, graf pa je prikazan na sliki 5

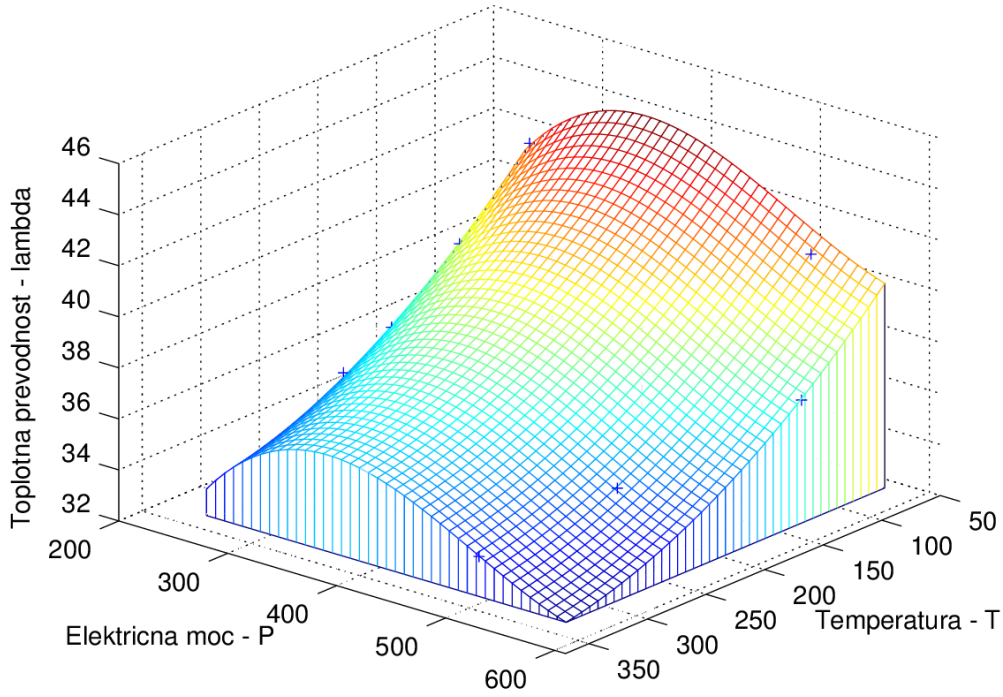
Tabela 3: Parametri najbolj ugodnega modela.

|              | $\beta_{ij}$         | $\sigma_{\beta_{ij}}$   |
|--------------|----------------------|-------------------------|
| $\beta_{10}$ | $-6.3 \cdot 10^{-2}$ | $\pm 3.9 \cdot 10^{-3}$ |
| $\beta_{01}$ | $3.2 \cdot 10^{-1}$  | $\pm 4.5 \cdot 10^{-3}$ |
| $\beta_{20}$ | $7.3 \cdot 10^{-5}$  | $\pm 8.6 \cdot 10^{-6}$ |
| $\beta_{02}$ | $-6.7 \cdot 10^{-4}$ | $\pm 1.9 \cdot 10^{-5}$ |
| $\beta_{03}$ | $4.3 \cdot 10^{-7}$  | $\pm 2.0 \cdot 10^{-8}$ |

Matrika singularnih vrednosti  $X$  je

$$S = \text{diag}(3.85 \cdot 10^8, 1.85 \cdot 10^5, 6.16 \cdot 10^4, 89.57, 35.23). \quad (5)$$

Rezultat za konfiguracijo 465, tj. 0111010001.  $\chi^2 = 2.584468e+00$



Slika 5: Točkam prilagojen polinomski nastavek za  $\lambda(T, P)$ . To je naša zadnja izbira, ki je dovolj nizka v  $\chi^2$ , ima vse singularne vrednosti večje od 10 in hkrati najmanj koeficientov v razvoju.

V vseh primerih je kovariančna matrika  $P$  skoraj diagonalna.

### 3.2 Dolžine disklinacij tekočih kristalov

Spet bomo uporabili kar nastavek, ki bo dolžinam disklinacij (označili jih bomo z  $\ell$ ) tekočih kristalov prilagodil polinom največ desetega reda v parametru kiralnosti (označili ga bomo z  $x$ ). Naš nastavek je torej

$$\lambda \approx \sum_{j=0}^9 \beta_j x^j. \quad (6)$$

Tako kot prej, smo tudi sedaj naredili vse možne izbire polinomskih modelov, in analogno kot prej, zavrgli vse, kjer je  $\bar{\chi}^2 \notin [0.6, 1.4]$ ,  $k > 7$  in  $s_{\min} \leq 1$ . Vsaka datoteka nam da seznam konfiguracij, ki preživijo sečnjo. Iz njih potem preberemo najbolj zanimive kandidate. Datoteka `wrboost_eight.dat` jih ima 8, `wrboost_omega.dat` jih ima 21, datoteka `wrboost_theta.dat` pa jih ima 15.

Prvo, kar sem opazil je to, da so indeksi vseh kandidatov večji od 512, kar pomeni, da vsi vsebujejo konstantni člen. Ko sem naprej pregledoval tabele sem poiskal tiste modele, ki imajo  $\bar{\chi}^2$  čim bližje 1, hkrati pa čim večje singularne vrednosti, čim manj parametrov  $\beta$ , hkrati pa mora biti tudi red polinoma čim manjši.

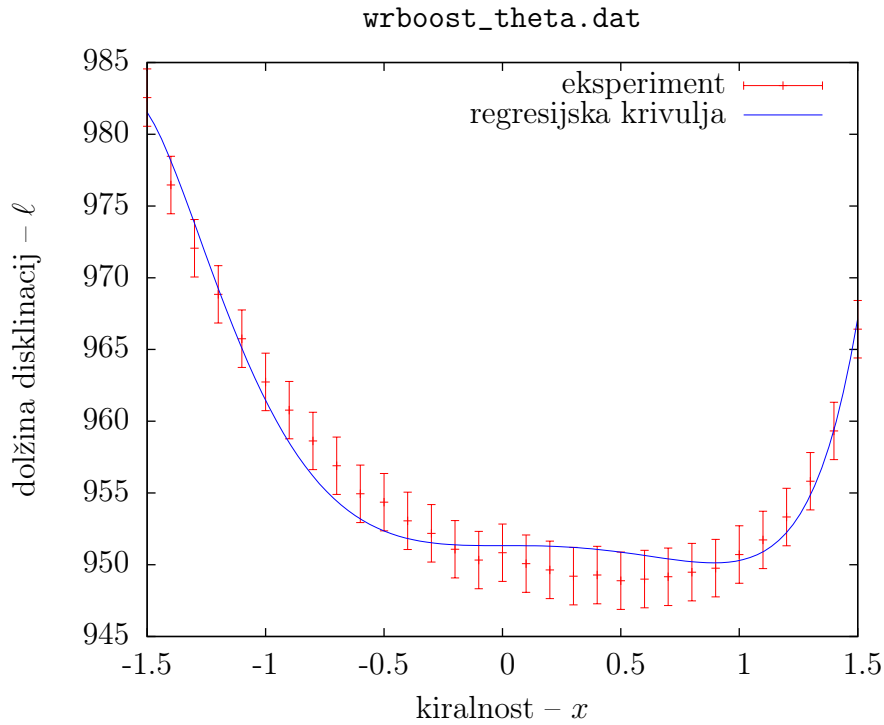


Za `wrboost_eight.dat` sem dobil rešitev  $i = 609$ , za `wrboost_omega.dat` sem dobil  $i = 522$ , za `wrboost_eight.dat` pa sem dobil tri možne rešitve, ki so se ujemale v  $k$ ,  $s_{\min}$  in v  $\overline{\chi^2}$ . To so  $i = 521, 524$  in  $536$ . Izbral sem  $i = 536$ , ker ima najmanjši red polinoma.

Edina potenca kiralnosti,  $x$ , ki je skupna vsem modelom, je  $x^0$ . Potem imata pa še `wrboost_omega.dat` in `wrboost_theta.dat` skupno potenco  $x^6$ . V vseh primerih sem dobil kovariančno matriko  $P$  skoraj diagonalno (nima tipične šahovnične strukture). Rezultati parametrov so v spodnjih tabelah 4, 5, 6 in grafih 6, 7 in 8.

Tabela 4: Po moji presoji najboljši kandidat za datoteko `wrboost_eight.dat`. Model je  $\ell \approx \beta_0 + \beta_3 x^3 + \beta_4 x^4 + \beta_9 x^9$ , tj.  $609_{10} = 1001100001_2$ .

|           | $\beta_i$                    | $\sigma_{\beta_i}$ |
|-----------|------------------------------|--------------------|
| $\beta_0$ | 951.33                       | $\pm 0.32$         |
| $\beta_3$ | -5.91                        | $\pm 0.40$         |
| $\beta_4$ | 4.54                         | $\pm 0.17$         |
| $\beta_9$ | 0.33                         | $\pm 0.05$         |
|           | $\overline{\chi^2} = 1.1681$ | $s_{\min} = 3.50$  |

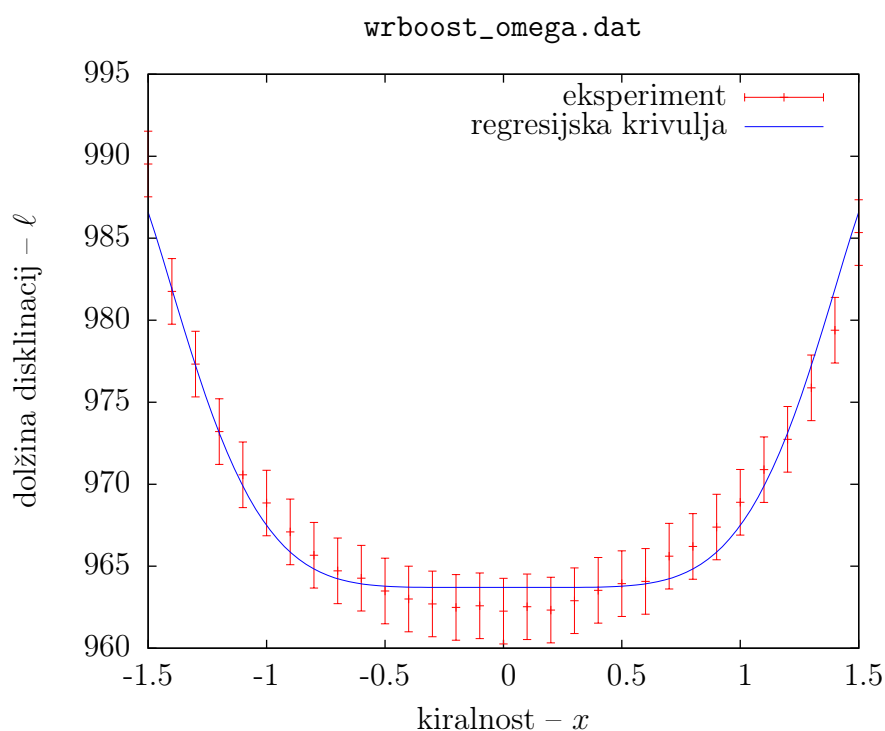


Slika 6: Ta graf ima polinom s štirimi parametri –  $\beta_0$ ,  $\beta_3$ ,  $\beta_4$  in  $\beta_9$ . Je sicer res, da je nekam nenavadno, da bi preprost fizikalni pojav opisovali s tako velikim redom, kot je 9, vendar analiza  $P$  in singularne matrike kaže v preferenčnost tega modela. Zanimal sem se tudi za model, ki je imel malo boljši  $\overline{\chi^2}$ , vendar sem pristal pri temu, ker je imel manj koeficientov v razvoju.

Rezultati za primer „omega“ so nekoliko drugačni, ker imamo v bistvu krivuljo, ki je soda, kar se tudi kaže v napovedani izbiri koeficientov  $\beta$ .

Tabela 5: Po moji presoji najboljši kandidat za datoteko `wrboost_omega.dat`. Model je  $\ell \approx \beta_0 + \beta_6 x^6 + \beta_8 x^8$ , tj.  $522_{10} = 1000001010_2$ .

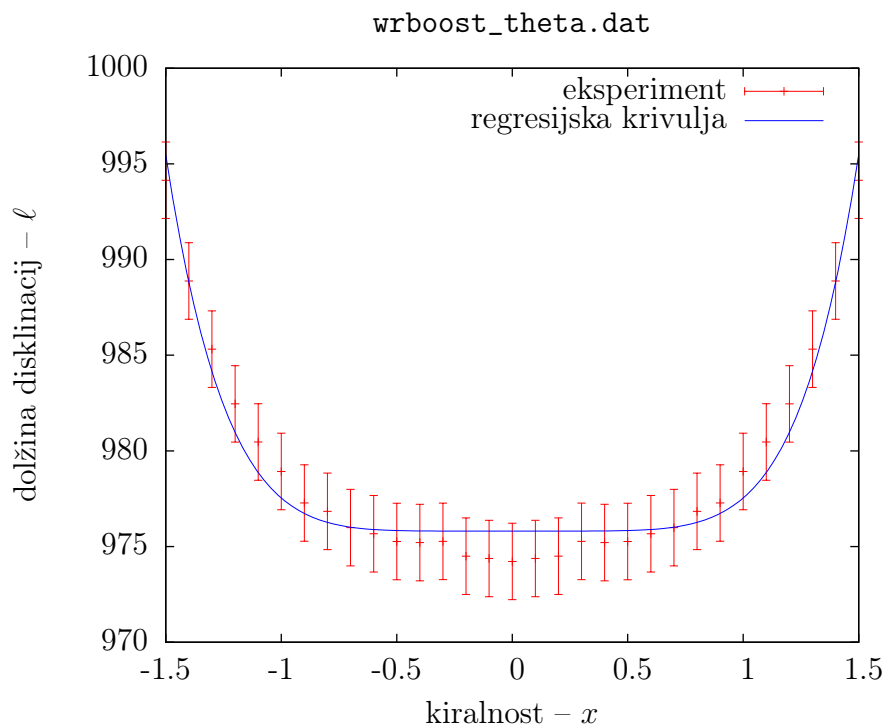
|           | $\beta_i$       | $\sigma_{\beta_i}$ |
|-----------|-----------------|--------------------|
| $\beta_0$ | 963.71          | $\pm 0.34$         |
| $\beta_6$ | 5.20            | $\pm 0.40$         |
| $\beta_8$ | -1.42           | $\pm 0.28$         |
|           | $\chi^2 = 0.80$ | $s_{\min} = 2.03$  |



Slika 7: Ujemanje je znotraj napake. Izgleda, da je izbrana napaka za merske točke ustrezna. Polinom ima parametre  $\beta_0$ ,  $\beta_6$  in  $\beta_8$ .

Tabela 6: Po moji presoji najboljši kandidat za datoteko `wrboost_theta.dat`. Model je  $\ell \approx \beta_0 + \beta_5 x^5 + \beta_6 x^6$ , tj.  $536_{10} = 1000011000_2$ .

|           | $\beta_i$       | $\sigma_{\beta_i}$ |
|-----------|-----------------|--------------------|
| $\beta_0$ | 975.81          | $\pm 0.30$         |
| $\beta_5$ | -3.55           | $\pm 0.10$         |
| $\beta_6$ | 1.73            | $\pm 0.08$         |
|           | $\chi^2 = 0.61$ | $s_{\min} = 4.72$  |



Slika 8: Spet vidimo, da je ujemanje ravno znotraj napake, ki je ustrezno napovedana. Polinom ima parametre  $\beta_0$ ,  $\beta_5$  in  $\beta_6$ .

## 4 Zaključek

Nekaj brskanja po internetu razkrije, da obstajajo avtomatizirani načini pravičnega določanja modelov, s katerimi se mi v tej nalogi nismo srečali, smo pa kljub temu izvedeli dovolj, da smo lahko s pomočjo singularnega razcepa ugotovili katere modele lahko zavrnemo in nato, kot se mnogokrat zgodi v analizi podatkov, zavrgli modele po opravljenem „reзу“ (ang. *cut*). Sicer je moj način nekoliko „špartanski“, (izračunali smo vse možne modele do v bistvu arbitrarnega reda v polinomskem razvoji in šele potem naredili primerjavo) in bi pričakovali, da bi skripta za `Octave` tekla počasi (`Octave` ima res počasne zanke) je program vseh 1023 modelov izračunal v kaki sekundi, tj. s tem nismo izgubili niti na času, niti v splošnosti.