Modelska analiza I
2011/12

6. naloga – Varčni modeli

Jože Zobec

1 Uvod

Teoretične parametre se v fiziki iz eksperimentov (slednji so lahko tudi numerični) ponavadi 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. χ^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 χ^2 skuša prilagoditi vektor parametrov β tako, da bo enačba $y = X\beta + \varepsilon$ imela čim manjši vektor napak ϵ. To naredimo prek minimizacije χ^2, ki je kar kvadratna forma

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

Zahteva, da je χ^2 minimalen vrne, da so idealni parametri

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

kjer je β vektor parametrov našega modela, X je strukturna matrika, ki β 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 β. 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 χ^2. Naš model bo največ tretjega reda, tj.

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

Parametrov β_{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 števio kombinacij 1023, saj

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

kjer '-1' pride zato, ker kombinacije $(00\ldots0)$ 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 χ^2 za kvantitativno mero ujemanja našega modela danimi eksperimentalnimi točkami, je predvsem prikladno uporabiti $\overline{\chi^2}$, tj. reduciran χ^2, ki je definiran kot

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

kjer je n število merskih točk, k pa je število parametrov β_{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 bolje ujemajo s točkami, bližje 1.

Ko bomo izračunali za vse možne primere izbire X vse možne χ^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.

3 Rezultati

Tabela 1 prikazuje toplotno prevodnost λ 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 uvaja v sisteme takrat, kadar hočemo dobiti rešitev neodvisno od skale problema, te pa v tem primeru ni).

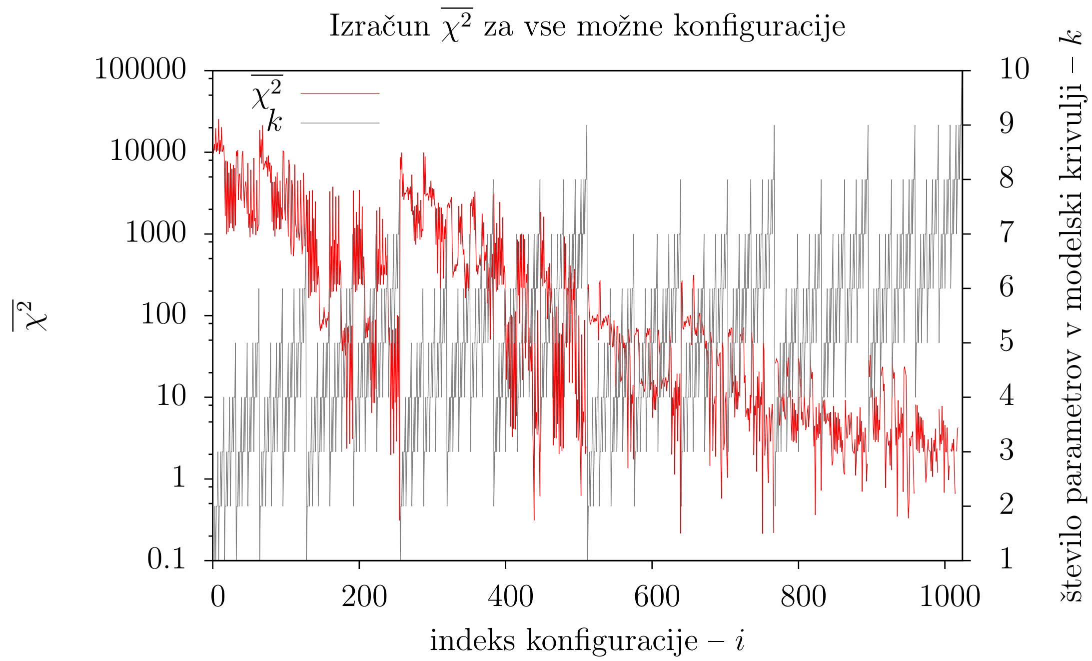
Tabela 1: Tabela meritev za toplotno prevodnost jekla Armco.

T[°F]	P[W]	λ[Btu/h ft °F]	σ_λ
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

$$X = \left[1 \mid T \mid P \mid T^2 \mid TP \mid P^2 \mid T^3 \mid T^2P \mid TP^2 \mid P^3 \right].$$

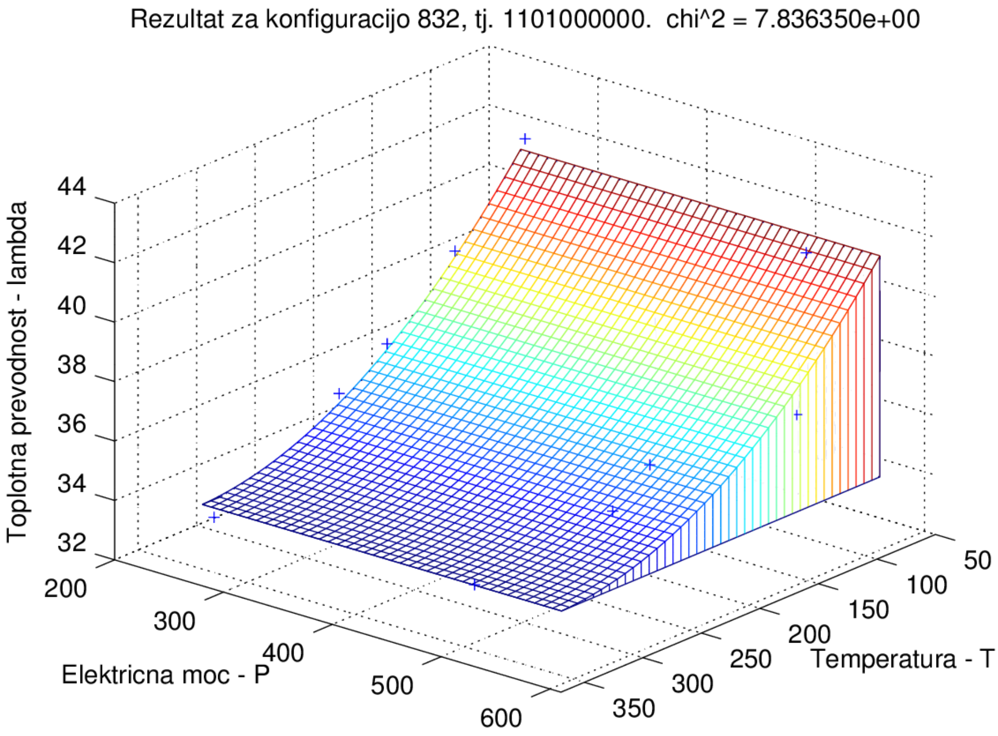
Konfiguracija (1001001000) predstavlja da v X obdržimo konstanto, T^2 in T^3, tj. da za model vzamemo $\lambda \approx \beta_{00} + \beta_{20}T^2 + \beta_{30}T^3$ in je oštevilčena z indeksom $2^9 + 2^6 + 2^3 = 584$. Matrika R je diagonalna, $R = \mathrm{diag}(\sigma^2_{\lambda_1}, \sigma^2_{\lambda_2}, \ldots, \sigma^2_{\lambda_{10}})$, vektor y pa vsebuje λ_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 $\overline{\chi^2}$ takoj nižji za velikostni red.

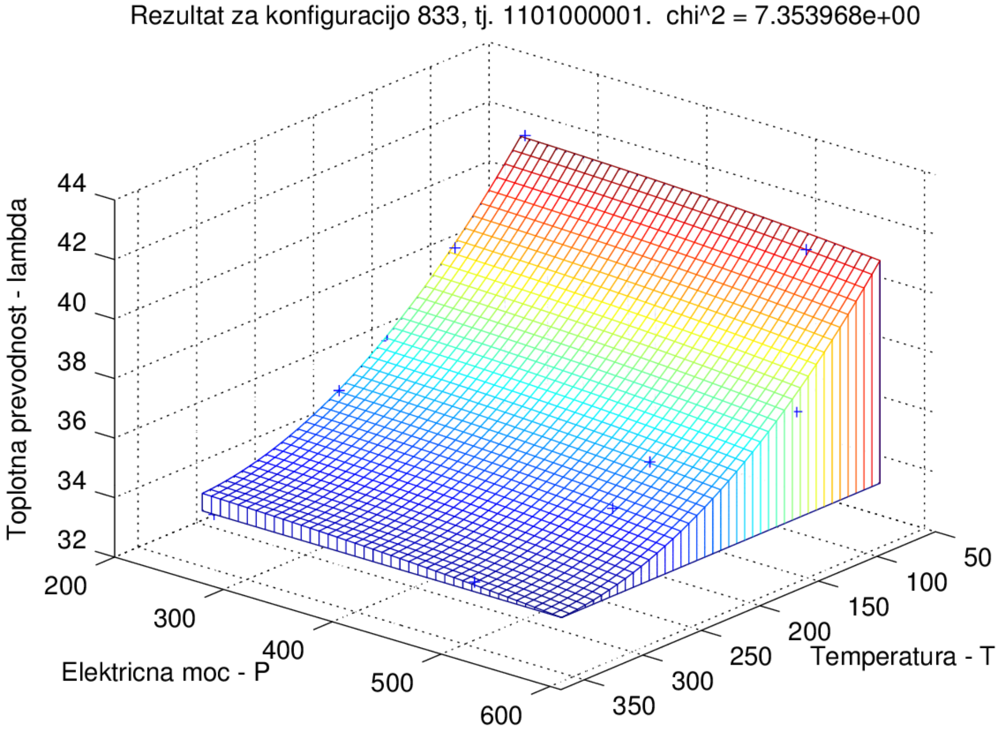
Izmed vseh teh možnosti sem se nato odločilo, da bom zavrgel vse, za katere je $\overline{\chi^2} < 0.5$, $\overline{\chi^2} > 8$, $k > 7$.

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.



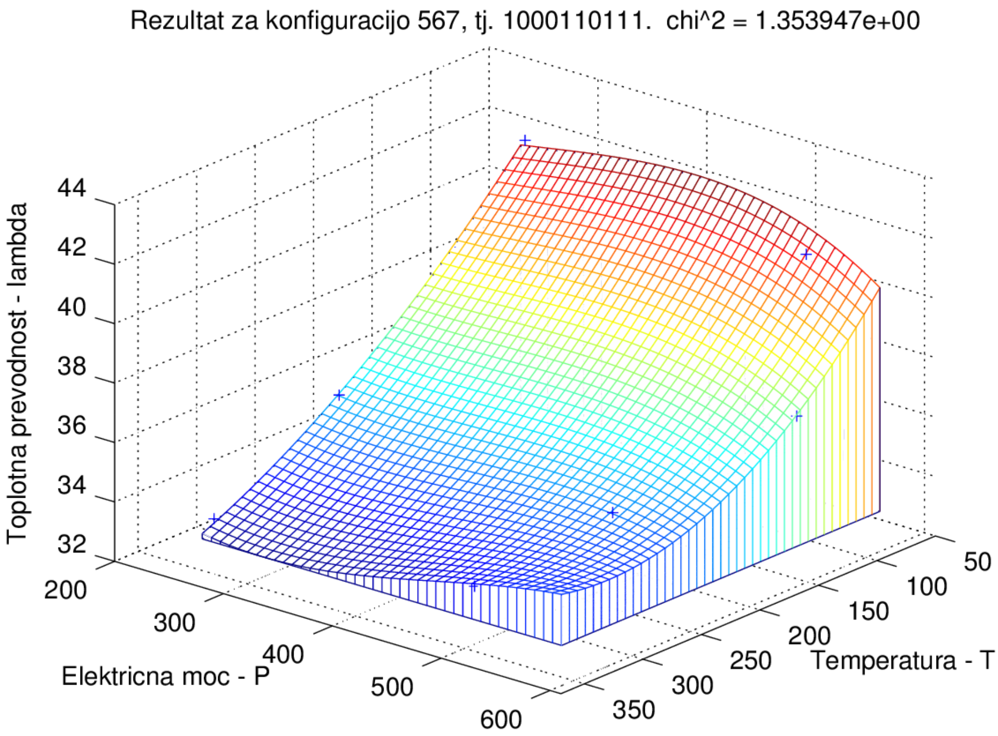
Slika 2: Odvisnost od P v tem primeru zanemarimo, in upoštevamo parabolično odvisnost od T.

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.



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 $\overline{\chi^2}$ najbljižje 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 $S = \mathrm{diag}(4.19 \cdot 10^8, 8.56 \cdot 10^7, 1.25 \cdot 10^7, 8.48 \cdot 10^4, 1.82 \cdot 10^4, 2.86 \cdot 10^{-1})$. Rezultat je na sliki 4.



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 so vsi brez konstantnega člena 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 najbljižlje $\overline{\chi^2} = 1$ in dobimo konfiguracijo 465. Ta najbolje zadošča kriteriju po čim manjšem številu parametrov k. Ta model je

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

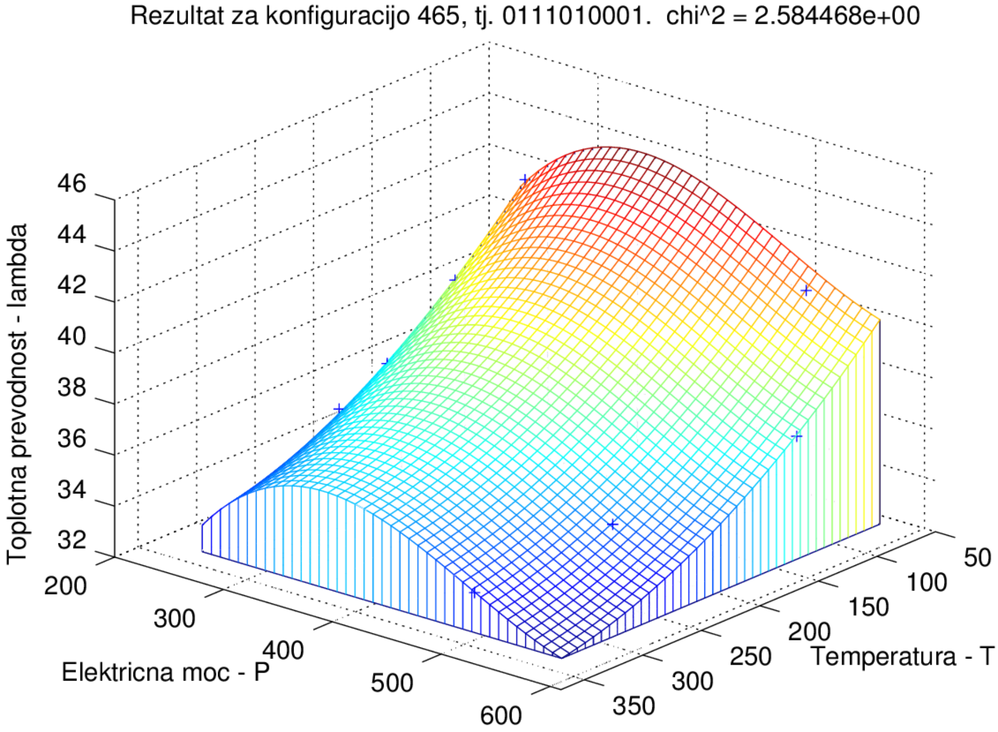
parametri z napakami (ki smo jih prebrali iz kovariančne matrike) so v tabeli 3, graf pa je prikazan na sliki 5

Tabela 3: Parametri najbolj ugodnega modela.

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

Matrika singularnih vrednosti X je

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



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