Varčni modeli

Miha Čančula

23. november 2011

1 Splošno

1.1 Naloga

Naloge sem reševal z razcepom matrike $\mathbf{S}$ na singularne vrednosti, tako da sem minimiziral izraz

$$\chi^2_{red} = \frac{1}{m-k}\sum_{i=1}^{m}\left(\frac{(\mathbf{Sa})_i - \mathbf{y}_i}{\sigma_i}\right)^2 \tag{1}$$

$$= \frac{1}{m-k}\left|\mathbf{CSa} - \mathbf{Cy}\right|^2 \tag{2}$$

Tu je $\mathbf{a}$ vektor parametrov, $\mathbf{y}$ vektor meritev odvisne spremenljivke, $\mathbf{S}$ pa modelska matrika ki izmerjene vrednosti neodvisnih spremenljivk povezuje s parametri. m je število meritev, k pa število uporabljenih parametrov. Matrika $\mathbf{C}$ je matrika uteži, s katero dosežemo, da imajo posamezne meritve različen vpliv na optimalno vrednost parametrov. V našem primeru, ko nimamo korelacij med posameznimi meritvami, je ta matrika diagonalna in podana z $\mathbf{C}_{ii} = \sigma_i^{-1}$.

Da sem lahko singularne vrednosti matrike $\mathbf{S}$ primerjal med seboj, sem najprej vse spremenljivke normiral, tako da je bil razpon njihovih vrednosti med -1 in 1. Ta interval ima veliko prednost, da ima tudi vsaka potenca spremenljivke vrednosti na tem intervalu. Ker je simetričen, se tudi izognemo korelaciji med x^i in x^{i+1}, saj je le ena izmed teh dveh funkcij na izbranem intervalu simetrična. V obeh primerih sem uporabljal le potenčne funkcije.

1.2 Reševanje

Na matriki $\mathbf{S}$ sem najprej uporabil razcep `SVD` na dve ortogonalni kvadratni matriki $\mathbf{U}$ in $\mathbf{V}$ ter pravokotno $\mathbf{W}$.

$$\mathbf{S} = \mathbf{UWV^T} \tag{3}$$

$$\mathbf{a} = \mathbf{S}^{-1}\mathbf{y} = \mathbf{VW^{-1}U^T y} \tag{4}$$

Matriki $\mathbf{S}$ in $\mathbf{W}$ nista kvadratni, zato sem uporabil pravilo za izračun psevdoinverza. $\mathbf{W}$ je po definiciji razcepa diagonalna, zato jo najprej transponiramo, nato pa popravimo diagonalne elemente

$$[\mathbf{W}^{-1}]_{ii} = \begin{cases} 1/\mathbf{W_{ii}}, & \mathbf{W_{ii}} > \varepsilon \\ 0, & \text{sicer} \end{cases} \tag{5}$$

Med računanjem sem preizkušal različne vrednosti za ε in na ta način dobil različno število uporabljenih parametrov k, da bo χ^2_{red} čim bližje 1.

To "bližino" sem definiral kot $b(x) = x + 1/x$, tako da velja $b(x) > 1$ za vsak x, ki ni enak 1. Sedaj sem lahko iskal takšne vrednosti parametrov, da bo $b(\chi^2_{red})$ čim manjši. Hkrati pa sem ši želel čim manjše število parametrov, zato sem v algoritmu minimiziral $b(x) + \alpha k$. Konstanta α odloča, ali nas zanimajo bolj točni ali bolj varčni modeli.

1.3 Napake parametrov

Napake oz. kovariančno matriko koeficientov sem izračunal po predpisu in Numerical Recipes

$$\mathbf{Cov} = \mathbf{V^T W^{-1} V} \tag{6}$$

Še posebej so me zanimali diagonalni elementi kovariančne matrike, saj sem med seboj primerjal napake in s tem pomembnost parametrov.

1.4 Iskanje optimalnega števila parametrov

Optimizacijo sem začel z velikim številom testnih funkcij, ki so bile kar prvih nekaj potenc neodvisnih spremenljivk. Po zgornjih formulah sem izračunal koeficiente pred temi funkcijami in njihove napake.

Preizkusil sem vse možne kombinacije teh testnih funkcij, nato pa obdržal tisto, ki je dala najmanjši χ^2_{red}. Zaradi velikega števila kombinacij (2^n, kjer je n število testnih funkcij) je bil račun dolgotrajen, na srečo pa sta razcep matrike in iskanje optimalne vrednosti parametrov hitra. V obeh nalogah sem uporabil okrog 10 testnih funkcij, v tem primeru se je račun končal v manj kot sekundi.

2 Toplotna prevodnosti

Za testne funkcije sem uporabil nastavke oblike

$$\varphi_{ij} = T^i P^j \tag{7}$$

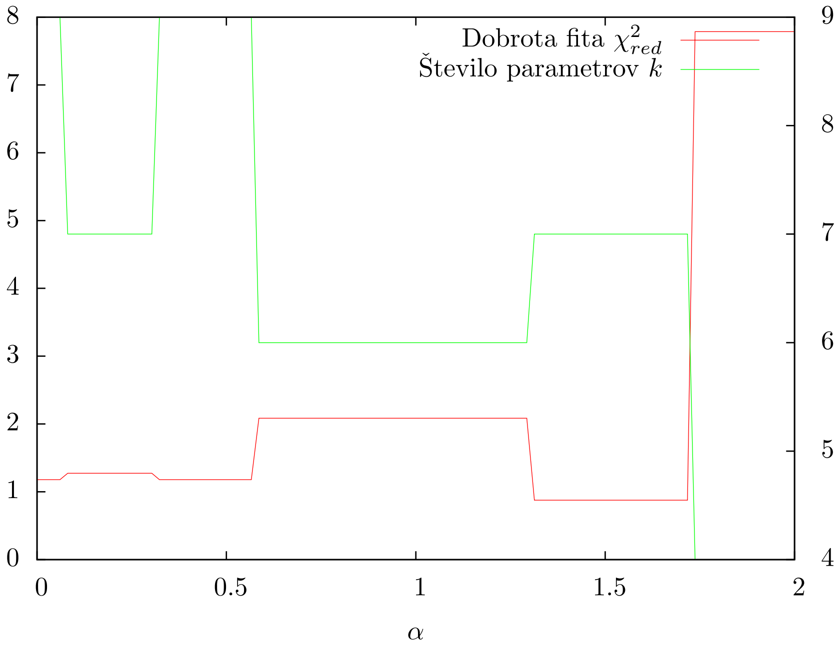
kjer sta T in P temperatura in moč grelca, skalirana na interval $[-1, 1]$. Ker imamo deset meritev, sem uporabil 9 testnih funkcij, tako da sta bila tako i kot j enaka 0, 1 ali 2.

Izračun sem ponovil pri različnih vrednosti konstante α. Pripadajoča χ^2_{red} in optimalno število parametrov sta prikazana na grafu.

Seveda je od zahtevane natančnosti odvisno, katera izbira za α je najboljša. Meni osebno se je za demonstracijske namene zdela najboljša izbira $\alpha \approx 1,5$, s sedmimi koeficienti in $\chi^2_{red} \approx 0,9$. S to izbiro sem izračunal koeficiente. Odvisnost toplotne prevodnosti od temperature in moči gretja lahko sedaj zapišemo kot

$$\lambda(T, P) = \sum_{i,j} a_{ij} T^i P^j \tag{8}$$

s koeficienti, podanimi v tabeli 1.



Slika 1: Odvisnost števila parameter in dobrote fita od konstante α

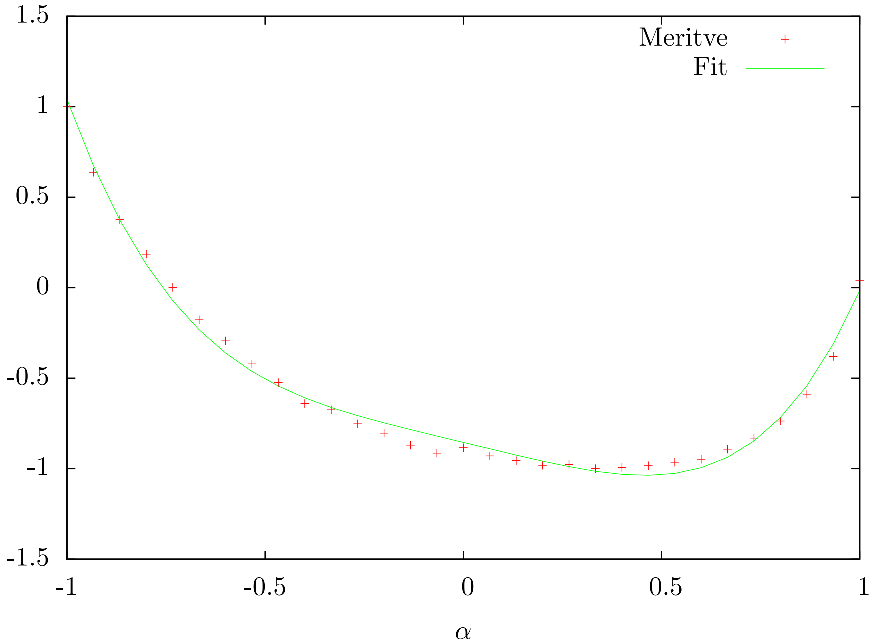
i	j	a_{ij}	$\sigma_{a_{ij}}$
0	0	-0.42493	0.14529
1	0	-1.39982	0.17735
2	0	0.43696	0.13286
0	1	0	0
1	1	0.40031	0.15236
2	1	0.16002	0.15608
0	2	0.11553	0.18965
1	2	0.80708	0.26507
2	2	0	0

Tabela 1: Optimalne vrednosti parametrov toplotne prevodnosti

3 Tekoči kristali

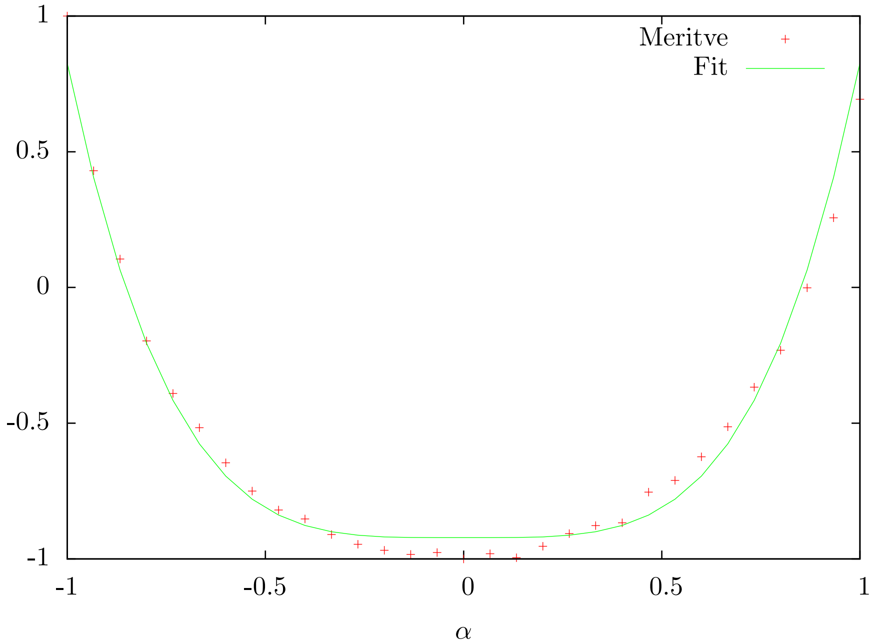
Isti postopek sem uporabil tudi na tabelah s podatki o dolžini disklinacij. Tokrat imamo same eno neodvisno spremenljivko, zato so bile testne funkcije kar njene potence. Za testne funkcije sem uporabil prvih 16 potenc te spremenljivke, omejil pa sem se le na kombinacije z največ 10 paremetri. Konstanto α sem zmanjšal, ker iščemo skupne člene iz vseh treh serijah meritev, zato dopustimo večje število parametrov.

Izračunani parametri so v tabeli 2.

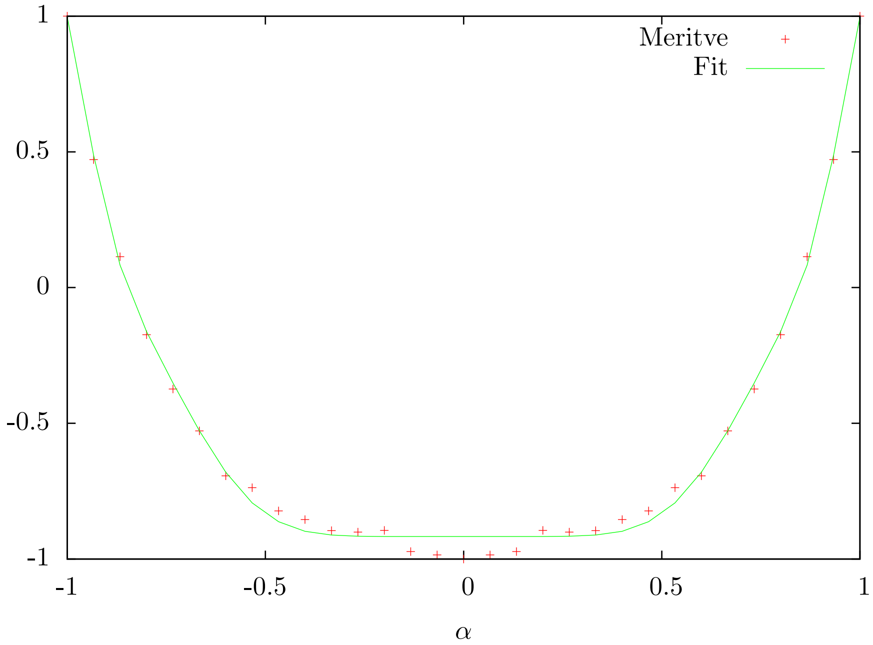


Slika 2: Podatki iz datoteke `wrboost_eight.dat`

Izbral sem takšno število koeficientov, da je bil χ^2_{red} v vseh primerih pod 1. Kljub temu ni nobene testne funkcije, ki bi nastopala v vseh treh konfiguracijah. No podlagi teh rezultatov bi pri podani oceni napake zaključil, da takšnega člena ni.



Slika 3: Podatki iz datoteke `wrboost_omega.dat`



Slika 4: Podatki iz datoteke `wrboost_theta.dat`

i	`eight`	`omega`	`theta`
0	-0.9 $\pm$ 0.1	-0.9 $\pm$ 0.1	-0.9 $\pm$ 0.1
1			
2			
3	-8.1 $\pm$ 0.5		
4	2.4 $\pm$ 0.4	1.7 $\pm$ 0.2	
5	47 $\pm$ 2		
6			
7	-123 $\pm$ 6		
8	-4 $\pm$ 1		49 $\pm$ 2
9	129 $\pm$ 7		
10	3.4 $\pm$ 0.8		-146 $\pm$ 6
11			
12			155 $\pm$ 8
13	-91 $\pm$ 6		
14			-56 $\pm$ 3
15	44 $\pm$ 4		
χ^2_{red}	0,21	0,88	0,20

Tabela 2: Koeficienti a_i za vse tri serije meritev