Modelska analiza I
1999/2000
12. naloga - Varčni modeli

Andrej Studen

3. januar 2000

1 Singular Value Decomposition

To je strokovno ime za razcep matrike, ki je lahko tudi singularna. Pri singularnih matrikah nas moti, da relativno majhne lastne vrednosti pomenijo skoraj neskončne vrednosti inverza, ki ga iščemo. Navadno namrečrešujemo enačbo:

$$\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}},$$

kjer je **A** singularna matrika. Za vsako matriko lahk naredimo razcep na tri matrike, katerih prva je po obliki enaka začetni matriki, druga je diagonalna, kvadratna s stranico moči števila stolpcev A-ja, zadnja pa je prav tako kvadratna. Tako dobimo razcep:

$$\mathrm{A} = \mathrm{UWV}^{\mathrm{T}},$$

za V in U pa veljaše, da sta ortonormirani matriki po krajši stranici matrike A. Za nas bo imela navadno matrika A več vrstic kot stoplcev in bo $A_{ij} = f_j(x_i)$, kjer je f_j j -ta poskusna funkcija, izračunana v i-ti točki. Imamo predoločen sistem. Takrat se izkaže, da lahko rešitev zgornje matrične enačbe zapišem kot:

$$\vec{\mathbf{x}} = \sum_{j=1, w_j \neq 0}^{M} \frac{(U^{(j)}\vec{\mathbf{b}})}{w_j} V^{(j)},$$

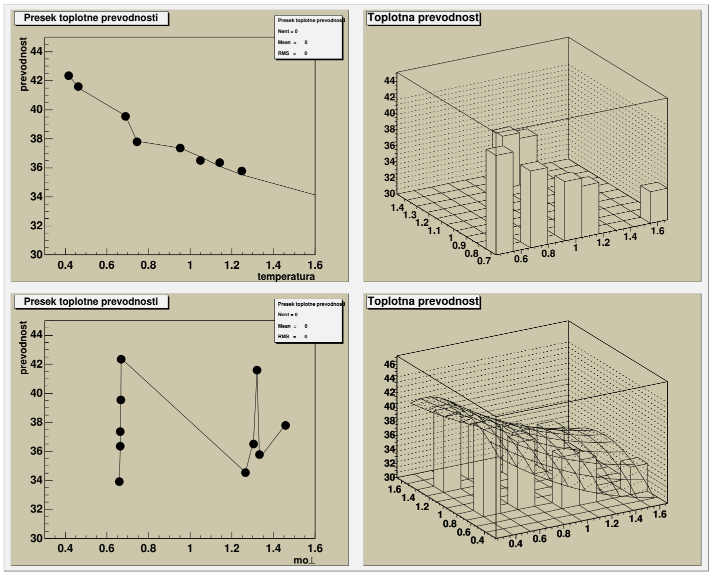
kjer sem z zgornjim oklepajem označil stolpce matrik U in V, zoprnih segmentov in neskončnosti pa se zneibm s seštevanjem samo po neničelnih diagonalcih. Izkaže se, da tako res dobimo minimalno rešitev. Matrika V nosi samo korelacije med posameznimi koeficienti, medtem ko so v U skrite ortogonalne funkcije, ki jih sestavim iz začetnega seta prostih funkcij. Tako je očitno, da bodo v vsakem stolpcu V-ja ravno skrite odvisonsti napram enemu parametru in bomo njegovo razpršenost σ_j dobili kot:

$$\sigma_{a_j}^2 = \sum_{i=1}^{M} \left(\frac{V_{ji}}{w_i}\right)^2.$$

Tako nam metoda sama razbije naše funkcije na ortogonalen set in vsaki poišče primerno diagonalno vrednost, ki pomeni hkrati tudi zastopanost v re'šitvi. Majhnih diagonalcev se tako znebimo in s tem tudi funkcij, ki nimajo smisla. Ta ortogonalen set zapišemo kot:

$$F_k(x_i) = \sum_{j=1}^{M} V_{jk} f_j(x_i) = w_k U_{ik},$$

seveda so malo pokmembne funkcije zaradi predčlena w_k res malo važne. Da pa so funkcije ortogonalne sledi iz pogoja $\mathrm{U}^{\mathrm{T}}\mathrm{U} = \mathrm{I}$.



Slika 1: Tako je videti 2D krivulja, ki jo modeliramo nad podatki, ki so zbrani v dveh premicah. Prileganje po teh presekih kažejo še izseki. Vendar pa je dodatna dimanzija vedno razlog za preplah, saj tisti hribček vmes težko upravičimo.

Tabela 1: Koeficienti in napake, skupaj s χ^2 za SVD rešitev s šestimi parametri, kot opisano v testu. Napake nekaterih parametrov so ogromne (npr. a_1), vendar pa je χ^2 v mejah normale.

	a_0	a_1	a_2	a_3	a_4	a_5	χ^2
koeficienti	40.55	16.74	-14.99	-8.93	0.88	3.68	6.910726
napake	4.71	16.98	1.26	3.66	0.34	0.17	

2 ARMCO

Tu je težko izbirati funkcije. POskusil sem kar s polnim razvojem do drugega reda:

$$\lambda(P,T) = a_0 + a_1 P + a_2 T + a_3 P^2 + a_4 PT + a_5 T^2,$$

pa še nisem dobil singularnih vrednosti, tudi če sem zviševal prag. Malo sem tudi spreminjal funkcije (vstavljal sem tretji red namesto drugega), pa sem dobil singularno točko, kar pomeni funkcijo preveč) ali pa precej večji χ^2. Tako sem ocenil ta rezultat kot maksimum, ki ga dosežem s 6 funkcijami. Ker nisem dobil nobene singularne vrednosti, prispeva vsaka od rešitev znatno h končnemu rezultatu in jih ne smem zavreči. Razen, če najdem boljše nekje v neskončnem vektorskem prostoru polinomov. Seveda bi to vzelo neskončno časa, pa sem obupal. Zato sem poskusil še z več funkcijami, pa se je izkazalo, da celo preveč zmanjšujejo χ^2, kljub temu, da niso singularne. Ko pa sem jih vstavil namesto že uporabljenih, niso opravile nič bolje. Zato bo končna rešitev kar tista zapisana zgoraj, njeni parametri pa sledijo, zapisani v tabeli.

3 Rja

Že prejšnijč smo poskušali prilagoditi sicer eksponentno krivuljo na naše podatke kot polinom. Uspeli smo najti lepo re'istev s štirimi potencami:

$$I_{(4)} = I_0 + AU + BU^2 + CU^3.$$

Tabela 2: Primerjava izračunanih χ^2 za različne načine računanje. Zanimivo da je petparametrična rešitev z U^6 bolj singularna (diagonalci so manjši) kot tista z U^4, čeprav prva bolj zmanjša χ^2.

način	χ^2
nelin.	934.438879
(4)	1526.932416
(4) $+U^4$	790.255619
(4) $+U^6$	605.884156

Tabela 3: Koeficienti po modelu, izračunani iz prvih štirih koeficientov. Ne razlikujejo se bistveno, kar nam še enkrat pove, da so prvi štirje bolj važni kot vsak naslednji, ki jim sledi, in zato delujejo le kot korekcije.

način	I_0	U_a	U_b	I_{konst}
(4)	0.002556	160.110533	65.514579	0.000315
(4) $+U^4$	0.002460	129.431640	68.384807	0.000255
(4) $+U^6$	0.002481	137.031456	67.281324	0.000260
nelin.	0.003056	196.471663	77.362471	0.000282

Seveda najprej preizkusim, če je to upravičeno. SVD nam res dopusti štiri parametre v modelu. Ko pa poskusimo s petim, že začnemo dobivati zmešnjavo. Recimo peta potenca ne pomaga, nam da takoj singularno vrednost. Zanimivo pa je, da nam dasta tako U^6 kot U^4 pomemben prispevek k χ^2, prva celo večjega:

$$I = I_{(4)} + DU^6 \qquad \text{in} \qquad I = I_{(4)} + DU^4.$$

Vendar pa lahko obe okvalificiramo kot nepotrebni, če le dovolj zvišamo prag. Vendar prve štiri takrat še vztrajajo in imajo zato večjo težo. Vendar pa U^6 odpade prej kot U^4, čeprav prispeva več k χ^2.

Če vzamemo več kot štiri člene v razvoju, eksponentnega modela in potenčnega razvoja nem oremo več primerjati, saj ima nelinearni eksponentni model le štiri paramtere. A poskusimo vseeno izračunati parametre sicer boljšega fita, a le iz prvih štirih koeficientov. Tako bomo videli, kakšno težo ima preostali člen. Ker so koeficienti podobni, je tudi dodaten člen majhen popravek. Torej je smiselno opisati I-U diagram s štirimi koeficienti in so vsi ostali le priveski.

4 Klirens

Tu rešujemo že znano Pronyevo enorazdelčno shemo:

$$\frac{c_{i+1} - c_{i-1}}{t_{i+1} - t_{i-1}} = \lambda c_i^{(n)},$$

kjer sem s $c_i^(n)$ označil n-razsežno povprečenje koncentreacije okrog i-te točke meritve:

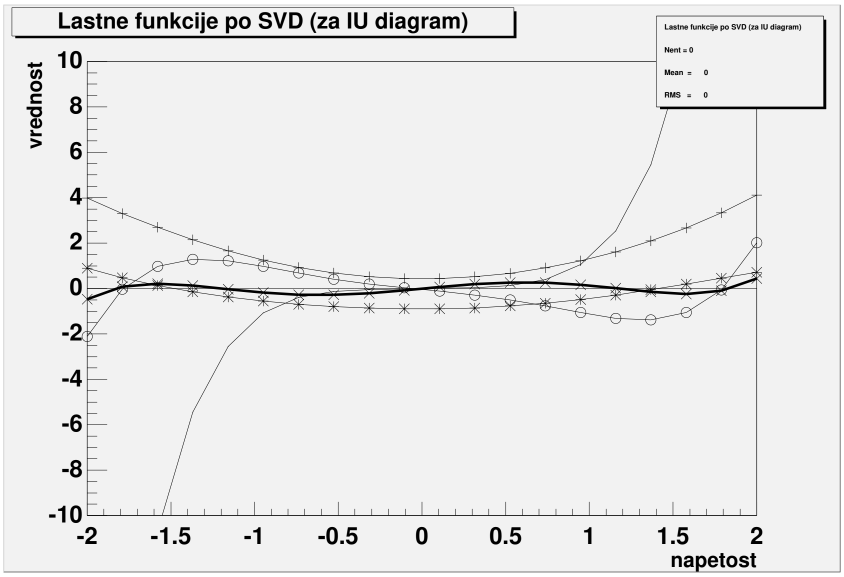
$$c_i^{(1)} = c_i,$$

$$c_i^{(2)} = \frac{1}{2}(c_{i+1} + c_{i-1}),$$

$$c_i^{(3)} = \frac{1}{3}(c_{i+1} + c_i + c_{i-1}).$$

Vmes sem se že odločil, da bom Pronyevo shemo postavil na tri točke, ker sse je to izkazalo za najboljšo rešitev pri linearnem modeliranju. Zato zapišemo zgornjo enačbo za točke $2, \ldots, N-1$ in imamo $N-2$ enačb za en sam parameter, tj. klirens(λ). Navadno smo enačbe pisali kot:

$$\mathbf{A}\vec{\mathbf{a}} = \vec{\mathbf{b}},$$



Slika 2: Slike lastnih funkcij za petparametrično ršitev. Peta funkcije je U^5, zanalašč izbrana tako, da jo SVD prepozna kot singularko. In res se peta funkcija vleče po ničli in jo iz obravnave lahko kar izpustimo (**krepko**).

kjer je bila matrika $\mathbf{A}$ model, odvisen le od časa, vektor $\vec{\mathbf{b}}$ pa le od izmerjenih točk. Tako poskusimo tudi zdaj in zapišemo:

$$(t_{i+1} - t_{i-1})\ \lambda = \frac{c_{i+1} - c_{i-1}}{c_i^{(n)}}.$$

In že imamo zgornji matrični enačbi ustrezno enačbo. Le matrika je zdaj $N-2$ razsežen vektor, prav tako kot je to $\vec{\mathbf{b}}$, vektor koeficeintov pa je skalar. Spet ima vektor $\vec{\mathbf{b}}$ napako, kot je to v navadi, izračunamo pa jo iz napak merjenja koncentracij. Pri deljenju seštevamo relativne napake, pri seštevanju absolutne, končni izraz pa je:

$$\sigma_{\vec{\mathbf{b}}} = \left(\frac{\delta\ c_i^{(n)}}{c_i^{(n)}} + \frac{1}{\sqrt{c_{i+1}} + \sqrt{c_{i-1}}} \right) \frac{c_{i+1} - c_{i-1}}{c_i^{(n)}},$$

in očitno je, da lahko za majhne razlike pridelam kar precejšnje napake.

No, od tu dalje delo za nas opravi SVD, ki poskuša najti najboljši parameter. Seveda smo s tem dosegli

Tabela 4: Primerjava koeficientov in klirensa za različne metode. Vedno vzamem kar simetričen odvod preko treh točk, srednjo vrednost pa določim kot povprečje števila vrednosti v oklepajih (pri 3 povprečim srednjo, levo in desno vrednost). Izkaže se, da je povprečen koeficient bolj'i kot tisti, dobljen z metodo least-square-fit-a. Razlog je verjetno kar v razglašenosti modela.

Način	Klirens	Koeficient	χ^2
SVD(3)	-0.001039	10364.1	5464.8
AVG(3)	-0.000806	8868.3	4059.0
SVD(2)	-0.001041	10373.6	5483.1
AVG(2)	-0.000801	8833.5	4061.0
SVD(1)	-0.001035	10340.2	5419.4
AVG(1)	-0.000817	8942.3	4060.0

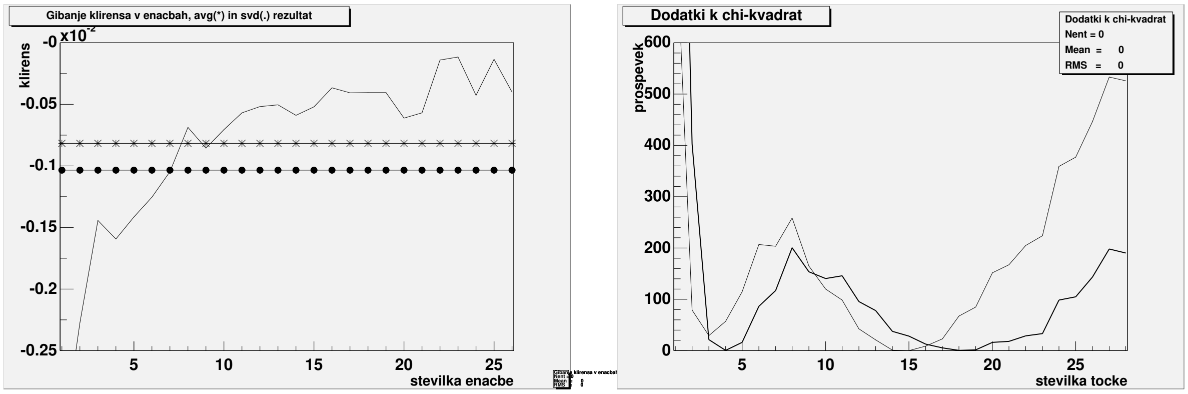
šele polovičko poti. (Mimogrede lahko določimo λ tudi s povprečenjem rešitev za λ vseh 26 enačb. Tej rešitvi bom rekel povprečna rešitev.) Določiti moramo še prosto konstanto pred eksponentom, to spet storimo z minimizacijo χ^2:

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{c_i - A\exp(\lambda t_i)}{\sqrt{c_i}} \right)^2 = \text{min.},$$

odvajamo po konstanti A in zahtevamo ničelni odvod, tako da sledi za A:

$$A = \frac{\sum_{i=1}^{N} c_i \exp(\lambda t_i)/\sigma_i^2}{\sum_{i=1}^{N} \exp(2\lambda t_i)/\sigma_i^2}.$$

To lahko storim tako za SVD kot za AVG rešitev. Primerjam obe in vidim, da zaradi slabosti modela povprečna rešitev celo bolje opiše potek rešitve kot SVD. Prva namreč minimizira ne glede na odmike, medtem ko jih druga upošteva. A odmiki so se močno spremenili zaradi linearizacije in v eksponentni sliki jo zmanjšanje kvadratov bolje odnese. Primerjave kažejo slike in preglednice.



Slika 3: Nastajanje napak pri računanju klirensa po Pronyu. Najprej moramo iz leve krivulje, kjer so nabrane rešiteve vseh 26 enačb, če gledamo vsako zase, izbrati neko srednjo vrednost za klirens. Kaj je to srednja vrednost, se je težko odlo;citi. Odločal sem se med navadnim povprečjem (minimizacija najmanjših kvadratov) in SVD rešitvijo. Slednja upošteva tudi transformirane napake. Vidimo, da se vrednosti razlikujeta, verjetno je na levem grafu SVD boljša rešitev. Nato določimo še konstante z minimizacijo χ^2 in gledamo izraz zanj in prispevke po točkah. Izkaže se, da je v eksponentni skali SVD slabši približek kot AVG. Da je temu tako, je verjetno kriv neustrezen model.