Modelska analiza I
1999/2000
11. naloga - Luščenje modelskih parametrov: nelinearni modeli

Andrej Studen, MA-FI

3. januar 2000

1 Levenberg-Marquardt

Podobno kot ponavadi pomeni več računanja računalnika manjšo obremenitev za človeka. Kajti metoda sama opravi vse. Sama določi ukrivljenost, sama se odloči za pravilno dolžino koraka,sama obrača matrike in sploh vse. Tako da nam ni treba vedeti skoraj ničesar.
Zdaj imamo našo poskusno funkcijo definirano kot:

$$y(x, \vec{\mathbf{a}}),$$

radi pa bi spremenili vektor $\vec{a}$ tako, da bi bil:

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

najmanjši. Torej imamo za vsak $\vec{\mathbf{a}}$ definirano funkcijo $\chi^2(\vec{\mathbf{a}})$, ki pa jo lahko okrog minimuma razvijemo v kvadratno formo:

$$\chi^2 = \chi^2_{\min} + \vec{\mathbf{a}}\vec{\mathbf{d}} + \frac{1}{2}\vec{\mathbf{a}}\mathbf{D}\vec{\mathbf{a}}.$$

V minimumu bo $\nabla\chi^2 = 0$, kjer razumem ∇ kot odvode po komponentah $\vec{\mathbf{a}}$, kar nam da za iskanje minimalnih parametrov $\vec{\mathbf{a}}$ pogoj:

$$\vec{\mathbf{a}}_{\min} = \vec{\mathbf{a}} - \mathbf{D}^{-1}\left[\nabla\chi^2(\vec{\mathbf{a}})\right],$$

kjer je matrika D ravno krivinska matrika:

$$D_{ij} = \frac{\partial^2\chi^2}{\partial a_i \partial a_j},$$

definirana s še eno polovičko, da se znabimo faktorja pri dovodu linearnih členov. Zgornja enačba se tako preiše v:

$$\mathbf{D}\ \delta\vec{\mathbf{a}} = \nabla\chi^2,$$

odkoder lahko s preprostimi rutinami za reševanje seta linearnih enačb preberem spremembo parametrov.
Seveda pa zgornji način ni edini. Včasih se je bolje premakniti v smeri največjega padca za nek konstanten korak. Takrat bomo preimk izračunali kot:

$$\delta\vec{\mathbf{a}} = \lambda\ \nabla\chi^2.$$

Oba načina lahko združimo v enega, ko rečemo:

$$(\mathbf{D} + \lambda\mathbf{I})\ \delta\vec{\mathbf{a}} = \nabla\chi^2.$$

Tabela 1: Primerjava poskušanja in nelinearne Levenberg-Marquardtove metode. Že s kartiranjem smo iz modela izvlekli zelo veliko, saj so spremembe v χ^2 minimalne. Nelinearni (boljši) rezultati so natisnjeni krepko.

način	A	λ_1	B	λ_2	K	χ^2
2	5263.337130	-0.000420	8184.932985	-0.00483		72.914960
2	5232.913828	0.000417	8176.526862	0.004755		**72.586306**
2K	5728.299361	-0.001244	6078.258205	-0.00698022	1847.850782	39.943186
2K	5716.738583	0.001239	6094.275317	0.006964	1843.237586	**39.940170**

Določiti moramo le še matriko **D** in parameter $|lambda$. Vidimo, da moramo servirati rutini odvode naše poskusne funkcije po parametrih; seveda je to povezano z gradientom χ^2, skratka za člene gradienta in krivinske matirke dobimo:

$$\nabla\chi^2 = \sum_{i=1}^{N} \frac{2(y_i - y(x, \vec{\mathbf{a}}))}{\sigma_i^2} \nabla y,$$

$$\frac{\partial^2\chi^2}{\partial a_j \partial a_l} = \sum_{i=1}^{N} \frac{2(y_i - y(x, \vec{\mathbf{a}}))}{\sigma_i^2} \frac{\partial^2 y}{\partial a_j \partial a_l} + \sum_{i=1}^{N} \frac{2}{\sigma_i^2} \frac{\partial y}{\partial a_j} \frac{\partial y}{\partial a_l}.$$

Prvi člen v drugem izrazu lahko zanemarimo, saj je izraz $(y_i - y(x, \vec{\mathbf{a}}))$ blizu 0, še posebej če upoštevamo vsoto, ki izniči statistične napake. Zato definiramo matriko α in vektor β kot:

$$\beta_j = \sum_{i=1}^{N} \frac{2(y_i - y(x, \vec{\mathbf{a}}))}{\sigma_i^2} \frac{\partial y}{\partial a_j},$$

$$\alpha_{jl} = \sum_{i=1}^{N} \frac{2}{\sigma_i^2} \frac{\partial y}{\partial a_j} \frac{\partial y}{\partial a_l} + \lambda\delta_{jl},$$

in rešujemo matrično enačbo:

$$\alpha_{jl}\ \delta a_l = \beta_j.$$

Prost je ostal edinole še parameter λ, ki pa ga uravnava rutina sama. Mi moramo torej priskrbeti funkcijo, ki bo za dan problem izračunala vrednost funkcije $y(x, \vec{\mathbf{a}})$ in njenih odvodov po parametrih.
Ostane še vprašanje stavka: *Metoda je našla rešitev.* Nimamo še tega kriterija in moramo ga nastaviti na roke. Seveda imamo radi, da se rešitve ne ustavijo, dokler se χ^2 zmanjšuje, vseeno pa nam je, če se zvišuje. Ko se le malo spreminja, je velika verjetnost, da hodimo v krogu in zato je smiselno takrat računanje prekiniti. Sam sem to storil, ko se je χ^2 zmanjšal za manj kot eno stotino.
Metoda nam da tudi oceno za napake parametrov. Diagonala inverza krivinske matrike, tj. matrike α za λ=0 je ravno merilo za točnost posameznih parametrov. Tako lahko ocenimo tudi točnost v okviru modela. Še vedno pa nam meri χ^2 odstopanja od modela. Inverz oziroma kovariantno matriko pa nam vrne tudi rutina, le parameter λ postavimo na 0.

2 Klirens

Primerjamo dvo-kompartmentska modela z ozadjem in brez njega. Dodal sem še minimume, ki sem jih dobil prejšnjič, ko sem dva parametra vstavil na pamet. Iste podatke sem vstavil tudi za začetni približek. Metoda je v dveh korakih našla malo izboljšano verzijo, Amerike pa ravno ni odkrila.

3 Farmakologija

Tu dodamo modelu še en parameter. Prejšnjič smo model linearizirali in našli edini minimum, ki ga z nelinearno metodo ne moremo izboljšati. Novi parameter pa bo upravičen, če bo opazno spremenil naš χ^2.

Tabela 2: Primerjava modelov in s tem tudi načina računanja za merjenje odziva tkiva na reagente. Za določanje p uporabimo nelinearne metode. Rezultati so res bistveno boljši.

način	y_0	a	p	χ^2
lin	104.675526	21.175714	1.000000	26.109741
nelin	99.609671	19.765679	1.362990	9.356928

Novi model je:

$$y = \frac{y_0 x^p}{a^p + x^p},$$

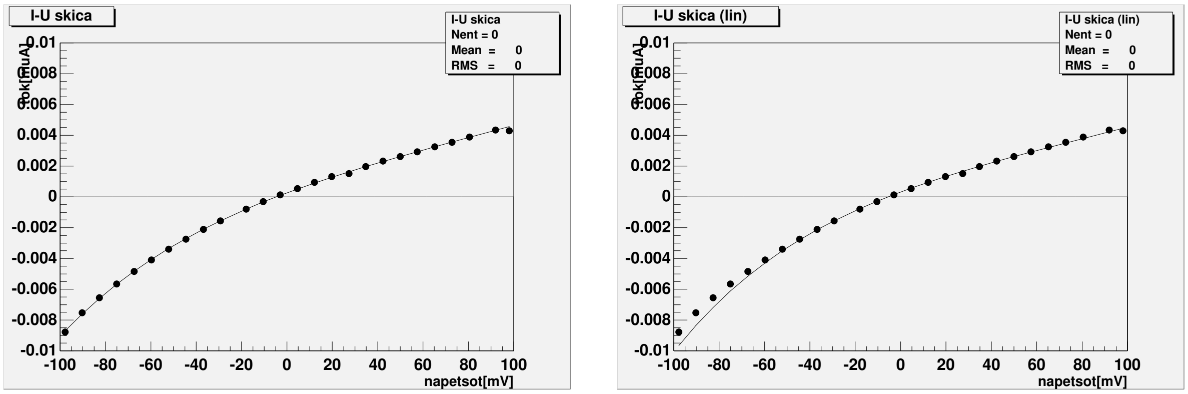
ki preide za $p = 1$ ravno v naš stari model in lahko rezultate celo primerjamo. Če začnemo kar s tem približkom, se metoda ustali po dveh korakih, rešitev pa upravičuje novi parameter.

4 Rjavenje

Za rjavnje kovin delajo testne I-U diagrame. Model odvisnosti naj bi podajala enačba:

$$I = I_0 \left[\exp(U/U_a) - \exp(-U/U_b)\right] + I_c,$$

kjer naj bi paramteri povedali nekaj o korozivnih lastnostih kovine. Konstanto sem dodal sam, kajti model je postavljen okrog ničle, tj. I(U=0)=0, kar pa očitno ni res, popravimo pa lahko kar s to konstanto, neznanim tokom, ki teče neodvsino od priključene napetosti. Upravicitve te konstante je tudi sicer kar veliko pa jih sploh ne bi našteval.

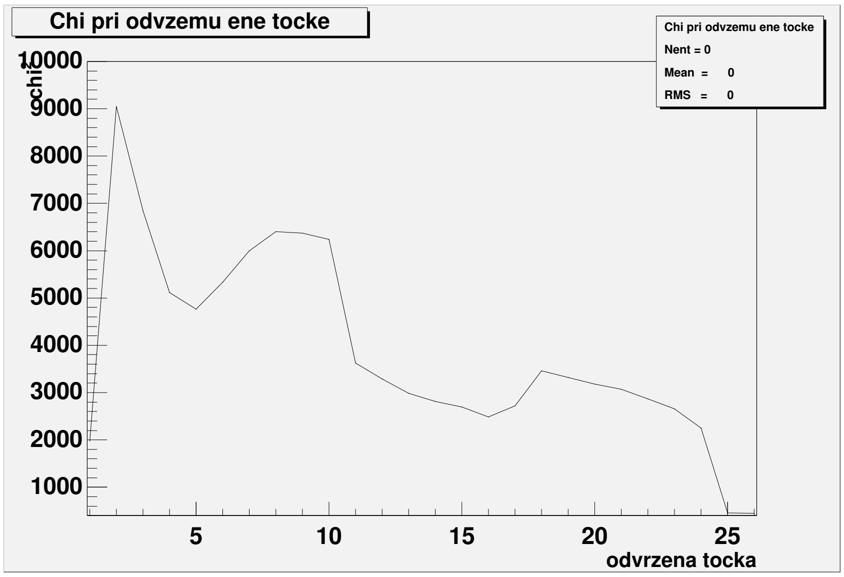


Slika 1: Primerjava ujemanja za najboljši možen scenarij, tj. izpuščeni dve točki in χ^2 na zavidljivih 450 (levo) in modeliranega razvoja ekpsonenta po vseh točkah(desno). Tudi desna rešitev, ki ima χ^2 približno 1500, je dovolj dobra. Na levi sliki se krivulja odpne od zadnjih dveh točk, ki sta moteči.

Rešitev lahko izboljšamo tudi tako, da nekaj točk odvržemo. Korenito v eksperiment nisem posegal, vendar pa sem poskusil naredit kak korak v tej smeri. Tako sem spremljal odvisnost χ^2, če se odpovem eni od točk. Izkaže se, da najbolj prispevata k napaki zadnji točki, zato sem obe izključil iz obravnanve in dobil najbolj prilegajočo se krivuljo.

Poleg nelinearnega postopka lahko tudi drugače uženemo naš model. Zaradi relativno majhne krivine lahko eksponent razvijemo v potenčno vrsto in dobimo linearnost v odvisnosti parametrov - to pomeni, da $\nabla\chi^2$ ni odvisen od $\vec{\mathbf{a}}$. Ker imamo štiri parametre, razvijamo toliko časa, da imamo štiri konstante. Dobimo polinom tretje stopnje:

$$I = I_c + I_0\left(\frac{1}{U_a} + \frac{1}{U_b}\right)U + I_0\left(\frac{1}{U_a^2} - \frac{1}{U_b^2}\right)U^2 + I_0\left(\frac{1}{U_a^3} + \frac{1}{U_b^3}\right)U^3 = I_c + AU + BU^2 + CU^3.$$



Slika 2: Tu se vidi, da sta največji odpadnici zadnji dve točki. Krivuljo dobim tako, da odvržem i-to točko in z nelinearno metodo določim parametre in s tem χ^2, na koncu pa točko vstavim nazaj.

Tabela 3: Korozijska krivulja in poskusi modeliranja eksperimentalnih podatkov. Najprej poiščemo parametre z nelinearno metodo, vključujoč vse točke. Nato nekatere točke odvržem, seveda tiste, pri katerih se χ^2 najbolj zmanjša. Nazadnje poskusim še z razvojem eksponentov; dobim model, katerega parametri so povezani linearno s poskusnimi funkcijami. Rezultati hvalijo nelinearne metode.

Ndrop	I_0	U_a	U_b	I_{const}	χ^2
0	0.003056	196.471663	77.362471	0.000282	934.438879
25	0.002835	170.450409	74.355829	0.000268	453.497237
26	0.002783	164.666286	73.645362	0.000264	445.471221
25,26	0.002791	165.516855	73.751830	0.000265	445.204413
lin	0.002556	160.110533	65.514579	0.000315	1526.932416

Zvezo med parametri obrnemo, da povežemo rezultate modeliranja z modelom:

$$\frac{1}{U_a} = \frac{C}{B} - \sqrt{\frac{6D}{B} - \left(\frac{C}{B}\right)^2},$$

$$\frac{1}{U_b} = -\frac{C}{B} - \sqrt{\frac{6D}{B} - \left(\frac{C}{B}\right)^2},$$

$$I_0 = \frac{B}{\sqrt{24D/B - (2C/B)^2}}.$$

Primerjavo med parametri in njihovo uspešnostjo podaja tabela. Vseh točk je ravno 26. Vidimo, da tudi linarni model ni slab, parametri so kar podobni.

Tabela 4: Višina, širina in lega prve spektralne črte, opremljena z napako. Kot pri naslednjih dveh slikah so primerjani različni načini iskanja parametrov. Tako se izkaže, da obstajata dva lokalna minimuma, eden z velikim razmerjem med širino prvega in drugega vrha, drugi z majhnim. Širino označim kar s s in o tem govorita tudi začetni označbi. K obema možnostima nato dodam še deseti parameter, ozadje, s konstanto, in primerjam še to. Parametri niso tako različni, kot namigujejo relativno močno razmaknjeni χ^2.

način	amplituda	Δ	širina	Δ	lega	Δ
$s_1 > s_2$	681.99	22.70	33.17	0.07	117.17	0.05
+K	712.92	28.17	28.55	0.05	116.69	0.03
$s_1 < s_2$	543.67	62.94	-15.09	0.06	115.67	0.03
+K	526.60	64.66	14.72	0.06	115.12	0.03

Tabela 5: Enako kot na prejšnji sliki, le za drugi vrh. Širina se sicer poveča, če jo silimo v tisto smer, ostali parametri pa so zelo podobni.

način	amplituda	Δ	širina	Δ	lega	Δ
$s_1 > s_2$	424.69	81.41	9.60	0.07	158.76	0.03
+K	462.57	69.70	10.26	0.07	158.37	0.03
$s_1 < s_2$	388.53	15.74	81.02	0.11	151.16	0.09
+K	406.14	17.33	71.04	0.13	150.39	0.12

5 Spekter

Spekter sestavimo kot vsoto treh nenormiranih Gaussovk:

$$y(x, \vec{\mathbf{a}}) = \sum_{i=1}^{3} a_1^i \exp(-((x - a_3^i)/a_2^i)^2).$$

Prilagajamo torej kar devet parametrov in pogumno sem dodal še desetega - konstantno ozadje sevanja. Res je dodatek upravičen, kar kažejo preglednice.
Treba je pazljivo začeti, izkaže se, da ima metoda dva lokalna minima. Parametre razen širine vrhov lahko dokaj dobro ocenimo iz slike meritve. Zaradi relativne bližine in razpotegnjenosti prvih dveh vrhov v spektru lahko najdemo dve podobno verjetni rešitvi. Verjetno je to en ozek (strm) hrib na vrhu drugega, širšega (položnejšega), pri čemer ne vemo, kateri od hribov, prvi ali drugi, je strmejši (ožji). Metoda se temu premisleku ne upira in daje dve podobno verjetni rešitvi. Na koncu se sicer izkaže, da je ena boljša, a obe sta vredni premisleka.

Tabela 6: Enako kot na prejšnjih dveh slikah, le za tretjo črto spektra. Tu se parametri skorajda ne spreminjajo, saj je vrh izrazit. Temu primerni sta tudi majhni napaki pri določanju lege in širine, medtem ko je vrh bolj vprašljiv.

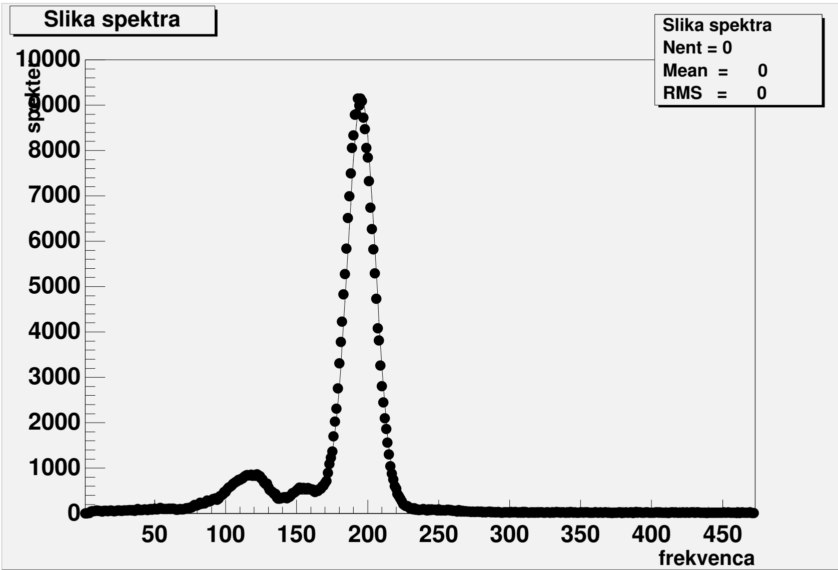
način	amplituda	Δ	širina	Δ	lega	Δ
$s_1 > s_2$	8917.39	562.96	14.72	0.00	195.64	0.00
+K	8930.30	569.72	14.62	0.00	195.61	0.00
$s_1 < s_2$	8868.39	601.38	13.82	0.00	195.66	0.00
+K	8865.52	601.19	13.82	0.00	195.72	0.00

Tabela 7: Uspeh računanja na različne načine. Če smo pripravljeni priznati ozadje, je raz'irjena prva črta podobno verjetna kot raz'irjena druga brez ozadja. Zmagovalec pa je drugi minimum, ki mu dodamo ozadje.

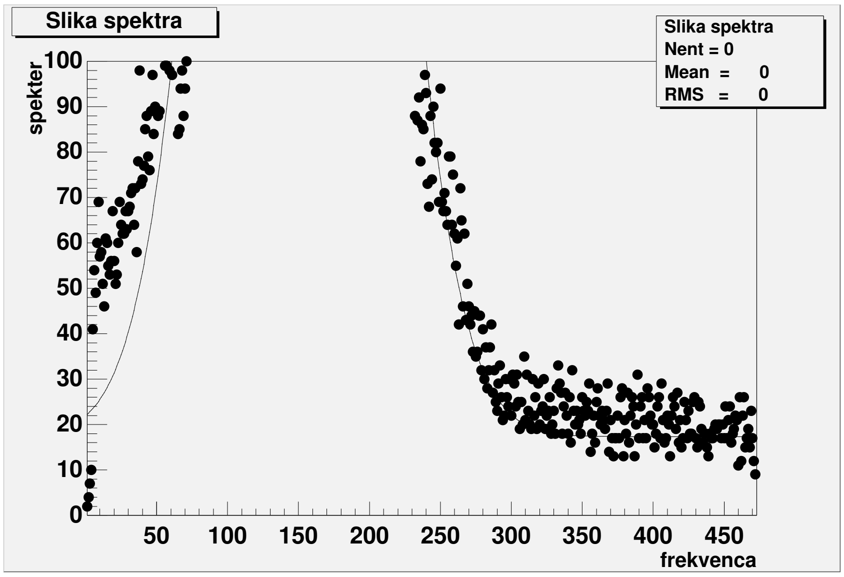
način	χ^2
$s_1 > s_2$	13050.653619
+K	5830.990375
$s_1 < s_2$	5703.965715
+K	2878.376346

Tabela 8: Še primerjava ozadja za različno razmerje širin prvih dveh črt.

način	ozadje	Δ
$s_1 > s_2$	23.82	0.08
$s_1 < s_2$	17.35	0.10



Slika 3: Spekter z analitično modelirano krivuljo. V gosti množici točk se naše krivulje sploh ne vidi, oster vrh pa dobro opiše.



Slika 4: Podrobnejša povečava nam odkrije, da je meritev močno razpršena, tudi močneje kot nam to narekuje σ^2. Upravičuje tudi ozadje v okviru napak. Nekaterih točk bi se tudi lahko znebili (1,2). Sicer pa sem nad prileganjem zadovoljen.