

UNIVERZITA PARDUBICE

Fakulta chemicko - technologická
Katedra analytické chemie

Licenční studium chemometrie:
Počítačové zpracování dat při kontrole a řízení jakosti

SEMESTRÁLNÍ PRÁCE

Předmět: Aproximace křivek, numerické
vyhlazování
Přednášející: Prof. J. Militký

Vyhlazení dat získaných měřením vodivosti

1. Zadání:

Proveďte dle vlastní strategie vyhlazení dat získaných měřením vodivosti oběhové vody v plynové kotelně v období od 24.7. 1996 do 19.12. 1996. Pro $x = 62$ spočítejte derivaci a pro interval $x = 100$ až 149 určete hodnotu integrálu.

2. Data:

Vstupní data jsou v souboru vyhlaz.dat (vodivost je uvedena v $\mu\text{S}/\text{cm}$):

Den	Vodivost
1	435
2	490
3	476
6	429
7	427
8	466
10	422
13	461
14	442
15	495
16	483
17	466
20	420
21	485
22	490
23	524
24	504
28	520
29	454
30	458
31	491
34	634
36	596
37	660
38	636

Den	Vodivost
41	706
42	670
43	650
44	644
45	592
48	512
49	566
50	566
51	556
52	532
55	568
62	528
63	534
64	562
65	576
66	564
69	516
70	514
71	502
72	534
73	558
76	580
77	614
78	604
79	572

Den	Vodivost
80	562
83	552
84	548
85	532
86	496
87	530
90	445
91	456
92	471
93	495
94	538
98	482
99	518
100	522
101	498
104	460
105	489
106	482
107	465
108	459
111	482
112	467
113	462
114	452
115	447

Den	Vodivost
118	474
119	463
120	466
121	449
122	454
125	430
126	410
127	450
128	463
129	450
132	437
133	457
134	433
135	437
136	467
139	475
140	444
141	439
142	465
143	485
146	497
147	510
148	530
149	516

3. Program:

Adstat 1.25: Kalibrace
Vyhlažování.

4. Output z PC:

Soubory: Vyhlaž.out.

5. Obrázky:

V příloze: Graf kubická spline regrese, 8 dnů/interval
Graf kubická spline regrese, 5 bodů/interval
Graf Reinschův spline, $S = 50000$
Graf reziduí vs. x , Reinschův spline, $S = 50000$
Graf reziduí vs. predikce, Reinschův spline, $S = 50000$
Graf Reinschův spline pro $x = 1$ až 100 , $S = 35000$
Graf Reinschův spline pro $x = 101$ až 149 , $S = 6000$
Graf Reinschův spline s $w_i = 1/s_e^2$, $S = 99$
Graf Reinschův spline s optimálními parametry

6. Řešení:

Pro vyhlazení dat jsem nejprve zvolil spline regresi (kubickou), zkusil jsem dvě různé metody pro volbu uzlů, intervaly konstantní délky a konstantní počet bodů v intervalu. Lepších výsledků bylo dosaženo druhou metodou, ale i při maximálním počtu uzlů (5 bodů na interval) byla dosažená reziduální směrodatná odchylka 24,28 příliš vysoká. Proto jsem k vyhlazování použil Reinschovu metodu. Protože nejsou k dispozici údaje o přesnosti jednotlivých měření, mají všechny body stejnou váhu ($w_i=1$). Počáteční hodnotu parametru vyhlazení S vypočteme pomocí vzorce $S = N * s^2$, kde N je počet bodů a s je směrodatná odchylka metody měření. Pro první odhad parametru vyhlazení S jsem použil hodnotu určenou na základě osobních zkušeností s podobnými metodami. Chyba konduktometrických metod by podle zkušeností neměla být větší než 3 až 5 %, proto pro směrodatnou odchylku měření s použijeme hodnotu 20, po zaokrouhlení se S rovná 40000. Křivka odpovídající zvolenému parametru vyhlazení poměrně dobře vyhovuje daným datům. Zvolíme-li vyšší hodnotu S , dojde k většímu vyhlazení, v extrémním případě může být výsledkem vyhlazení přímka určená MNČ. Zvolíme-li nižší hodnotu S , bude se křivka více přibližovat jednotlivým měřením, v extrémním případě bude výsledkem spojnice jednotlivých bodů, vzhledem k charakteru vyhlazovací funkce by došlo k neúměrně velkým oscilacím prokladu. Zkouškami bylo stanoveno, že má smysl pracovat s hodnotami S většími než 25000 a byla nalezena optimální hodnota parametru $S = 50000$, při níž byla směrodatná odchylka reziduí rovna 22,47, což je v souladu s očekávanou chybou měření (4.45 %). Kritériem pro posouzení kvality vyhlazení bylo vizuální posouzení grafů a porovnání hodnot s_e resp. RSC. Při prohlídce grafů reziduí je patrná jejich heteroskedasticita. Z grafu reziduí v závislosti na y je zřejmý trend, s rostoucím y rezidua klesají. Tento trend je způsoben vyhlazováním a jeho existence je snadno pochopitelná. Z grafu reziduí v závislosti na x je patrné, že rozptyl reziduí mírně klesá s rostoucím x . Navíc i z průběhu grafu závislosti y na x je dobře patrné, že soubor dat (měření) můžeme rozdělit podle rozptylu na 2 intervaly: oblast hodnot $x = 0$ až 100 a 101 až 150 . Je patrné, že druhá oblast má menší rozptyl než první. Proto bude vhodné vyhladit data odděleně, nebo upravit spline podle variability dat použitím vah w_i . Optimálního vyhlazení v prvním intervalu bylo dosaženo při $S = 35000$ s

reziduální směrodatnou odchylkou 23,39, ve druhém intervalu pro $S = 6000$ s reziduální směrodatnou odchylkou 13,09. Zjištěné hodnoty reziduálního rozptylu na obou intervalech byly použity pro výpočet vah $w_i = 1/s_e^2$, pro první interval $w_1 = 1/547$ a pro druhý interval $w_2 = 1/171$ pro $S = 99$. Výsledek vyhlazení s těmito parametry není uspokojivý, neboť došlo k silným oscilacím prokladu ve druhé oblasti dat (spline se nechová ideálně) vlivem příliš vysoké hodnoty w_2 , částečného zlepšení lze dosáhnout zvětšením parametru S na hodnotu 250, kdy $s_e = 20,78$. Pro dosažení uspokojivého vyhlazení je nutné především upravit hodnotu w_i . Proto byla optimální hodnota vah w_i určena pokusy. Nejlepšího vyhlazení dat bylo dosaženo při $w_1 = 1/380$, $w_2 = 1/1374$ a $S = 99$, kdy $s_e = 20,64$. Hodnota derivace pro $x = 62$ je 0,541 a hodnota integrálu pro $x = 100$ až 149 je vypočtena jako rozdíl integrálu pro $x = 149$ a integrálu pro $x = 100$ a činí 22800.

7. Závěr:

Optimálního vyhlazení dat bylo dosaženo Reinschovou metodou při $w_1 = 1/380$, $w_2 = 1/1374$ a $S = 99$. Hodnota derivace pro $x = 62$ je 0,541 a integrál pro $x = 100$ až 149 je 22800..

Název souboru: VYHLAZ
Adresář: E:\Pom\approx
Šablona: C:\OFFICE\WORD20\SKOL_UK.DOT
Název: Semestální práce
Předmět: Aproximace křivek, numer. vyhlazování
Autor: Ing. Miloslav Korhoň
Klíčová slova: Militký
Komentáře: hotovo
Datum vytvoření: 30.07.97 22:29
Číslo revize: 11
Poslední uložení: 20.08.97 13:27
Uložil: Ing. Miloslav Korhoň
Celková doba úprav: 35 min.
Poslední tisk: 15.09.00 08:26
Jako poslední úplný tisk
 Počet stránek: 4
 Počet slov: 871 (přibližně)
 Počet znaků: 4 966 (přibližně)