Úlohy lineárnej a nelineárnej aproximácie riešené MS Excel-om

Jana Boržíková

ABSTRACT: In this paper are described using MS Excel to solve approximation problems in numerical mathematics at technical universities.

1. ÚVOD

Výpočtová technika je osobitnou kategóriou didaktickej techniky. Pri jej využívaní vo vyučovaní numerickej matematiky na vysokých školách technického zamerania sledujeme dva ciele: niekedy je nutné, aby študenti sami vytvárali algoritmy riešenia i samotné programy, inokedy zase, aby boli len používateľmi hotových programov.

Prvý cieľ je popísaný v [1]. V tomto článku sa zameriam na druhý cieľ: využitie niektorých špeciálnych funkcií, ktoré má k dispozícii MS Excel. Takéto riešenie, ktoré je síce veľmi rýchle a neobmedzené vzhľadom na vstupné podmienky, prezentuje len samotný výsledok a nie je jasný použitý algoritmus. No pre ďalšiu prax a pre výpočty na iných vyučovacích predmetoch je vhodnejšie využívať práve tieto funkcie.

Viacnásobné opakovanie jednej úlohy v rôznych súvislostiach a rôznymi metódami umožňuje nielen aplikovať, ale aj upevňovať a prehlbovať vedomosti, čiastočne ich aj rozširovať a zovšeobecňovať. Je zrejmé, že rôzne úrovne zvládnutia Excel-u (príp. iného softwaru) zabezpečia nie len trvalé osvojenie si metódy študentmi, ale aj to, že budú pristupovať k riešeniu rôzne a tým rozvíjajú svoju tvorivosť a schopnosti riešiť nové úlohy. [2]

2. Formulácia úlohy

V technickej praxi sa stretávame s úlohou nájsť funkciu, ktorá by čo najlepšie nahradzovala namerané hodnoty.

Takouto úlohou môže byť napríklad sledovanie vplyvu jednotlivých parametrov pri Elektroerozívnom obrábaní (EDM) na výslednú drsnosť povrchu. Alebo inverzná úloha: sledovanie vplyvu hrúbky a drsnosti povrchu na ostatné parametre. Úloha bola riešená v dizertačnej práci [3]. Sledované parametre ovplyvňujúce EDM sú:

- výkon P (W),
- prúd I (A),
- rezná rýchlosť v_s (mm/min),
- hrúbka l (mm),
- drsnosť R_a (μm) pri konštantnej alebo rôznej hrúbke.

Z matematického hľadiska ide o úlohu aproximovať namerané hodnoty funkciou viac premenných:

$$R_a = f(l, I, v_s, P), \qquad (1)$$

prípadne

$$P = f(l, R_a),\ I = f(l, R_a),\ v_s = f(l, R_a). \qquad (2), (3), (4)$$

3. Aproximácia funkciou viac premenných

Keďže chceme hodnoty následne aj graficky znázorniť, obmedzíme sa na funkcie (2), (3), (4). Pri metóde najmenších štvorcov aproximujeme zadaných n bodov [l, R_a, P] funkčnou závislosťou

$$P = f(l, R_a, A) = f(l, R_a, a_1, \ldots, a_k),$$

kde neznáme parametre a_j, $j = 1, \ldots, k$ musíme určiť tak, aby krivka čo najlepšie priliehala k hodnotám. Aby bola splnená požiadavka, za kritérium budeme uvažovať minimum funkcie:

$$S(A) = \sum_{i=1}^{n} \left[y_i - f\left(l_i, R_{ai}, A\right) \right]^2 \Rightarrow \min$$

Odtiaľ vyplývajú podmienky pre koeficienty a_j. Pri lineárnej aproximácii

$$P = f\left(l_i, R_{ai}, A\right) = a_1 + a_2.l + a_3.R_a \qquad (5)$$

ide o funkciu troch premenných, pričom extrém nadobudne vtedy, ak prvé parciálne derivácie položíme rovné nule.

$$\frac{\partial S}{\partial a_1} = 2\sum_{i=1}^{n} y_i - a_1 - a_2 l_i - a_3 R_{ai} = 0$$

$$\frac{\partial S}{\partial a_2} = 2\sum_{i=1}^{n} \left(y_i - a_1 - a_2 l_i - a_3 R_{ai} \right).l_i = 0$$

$$\frac{\partial S}{\partial a_3} = 2\sum_{i=1}^{n} \left(y_i - a_1 - a_2 l_i - a_3 R_{ai} \right).R_{ai} = 0$$

T. j. musia vyhovovať sústave:

$$a_1 n + a_2 \sum_{i=1}^{n} l_i + a_3 \sum_{i=1}^{n} R_{ai} = \sum_{i=1}^{n} y_i$$

$$a_1 \sum_{i=1}^{n} l_i + a_2 \sum_{i=1}^{n} l_i^2 + a_3 \sum_{i=1}^{n} R_{ai} l_i = \sum_{i=1}^{n} y_i .l_i$$

$$a_1 \sum_{i=1}^{n} R_{ai} + a_2 \sum_{i=1}^{n} l_i + a_3 \sum_{i=1}^{n} R_{ai}^2 = \sum_{i=1}^{n} y_i R_{ai}$$

Pri exponenciálnej aproximácii $P = a_1.a_2^{l}.a_3^{R_a}$ musíme použiť metódu linearizácie:

$$\ln P = \ln a_1 + l.\ln a_2 + R_a \ln a_3$$

a substitúciou $u = \ln P, A_1 = \ln a_1, A_2 = \ln a_2, A_3 = \ln a_3$ upravíme na tvar: $u = A_1 + lA_2 + R_{ai} A_3$, čo je opäť lineárna závislosť (5). Po výpočte koeficientov vrátime späť hodnoty $a_1 = e^{A_1}, a_2 = e^{A_2}, a_3 = e^{A_3}$.

Tesnosť nahradenia je charakterizovaná indexom korelácie:

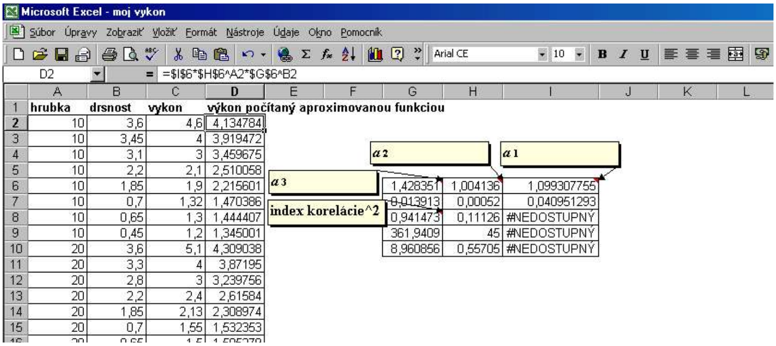
$$IK = \sqrt{1 - \frac{\sum_{i=1}^{n} \left(P_i - y_i \right)^2}{\sum_{i=1}^{n} \left(\bar{y} - y_i \right)^2}},$$

kde P_i sú vypočítané hodnoty a $\bar{y}$ je aritmetický priemer nameraných.

4. Riešenie pomocou MS Excel

Ak sa chceme vyhnúť zdĺhavosti výpočtu vzhľadom na počet nameraných hodnôt a rozmer vektora nezávislej premennej, využívame pre riešenie funkcie viac premenných funkcie MS Excel: LINREGRESE a LOGLINREGRESE. Tie pomocou metódy najmenších štvorcov vrátia maticu hodnôt, ktorá popisuje priamku (exponenciálnu krivku) najlepšie odpovedajúcu zadaným hodnotám. Predpokladajú vstupné hodnoty: vektor závislých hodnôt y_i a vektor (pole) nezávislých hodnôt X_i.

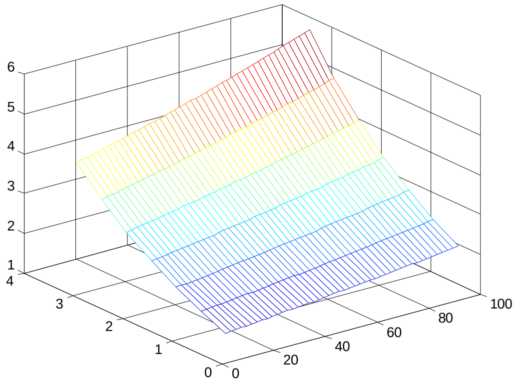
Výsledná matica hodnôt je na obrázku 1. Bola použitá funkcia LOGLINREGRESE. Vidíme, že hodnoty sú aproximované funkciou $P = 1{,}09930755.1{,}004136^{l}.1{,}428351^{R_a}$, index korelácie je 0,9703. Jej graf je na obrázku 2a). Použili sme programovací systém Matlab.



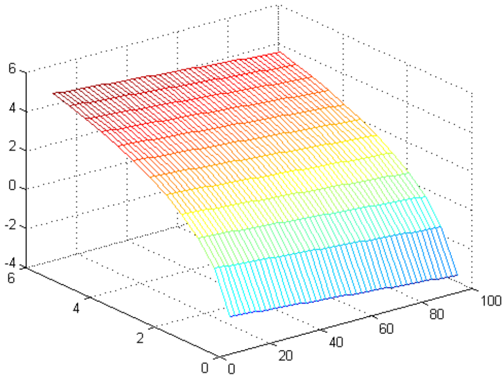
	A	B	C	D
1	hrubka	drsnost	vykon	výkon počítaný aproximovanou funkciou
2	10	3,6	4,6	4,134784
3	10	3,45	4	3,919472
4	10	3,1	3	3,459675
5	10	2,2	2,1	2,510058
6	10	1,85	1,9	2,215601
7	10	0,7	1,32	1,470386
8	10	0,65	1,3	1,444407
9	10	0,45	1,2	1,345001
10	20	3,6	5,1	4,309038
11	20	3,3	4	3,87195
12	20	2,8	3	3,239756
13	20	2,2	2,4	2,61584
14	20	1,85	2,13	2,308974
15	20	0,7	1,55	1,532353

	G	H	I
6	1,428351	1,004136	1,099307755
7	0,013913	0,00052	0,040951293
8	0,941473	0,11126	#NEDOSTUPNÝ
9	361,9409	45	#NEDOSTUPNÝ
10	8,960856	0,55705	#NEDOSTUPNÝ

Obr. 1



Obr. 2a



Obr. 2b

Týmto spôsobom môžeme vypočítať a porovnávať rôzne funkčné závislosti. V prípade špeciálnych tvarov funkcie, či nevyhovujúceho indexu korelácie využívame substitúcie.

Napríklad, pri inverznej závislosti $R_a = f(l, P)$ sme týmto spôsobom zistili že index korelácie je 0,833, čo je nízka hodnota, tak sme požadovanú závislosť vyjadrili zo vzťahu $P = (R_a, l)$, čím sme zachovali index korelácie. Hľadaná funkcia P má tvar:

$$R_a = \frac{\left(\ln P - \ln a_1 - l.\ln a_2\right)}{\ln a_3}$$

$$R_a = \frac{\left(\ln P - 0{,}09468 - l.0{,}00413\right)}{0{,}35652} =$$

$$= 2{,}80489 \ln P - 0{,}0115842.l - 0{,}26557$$

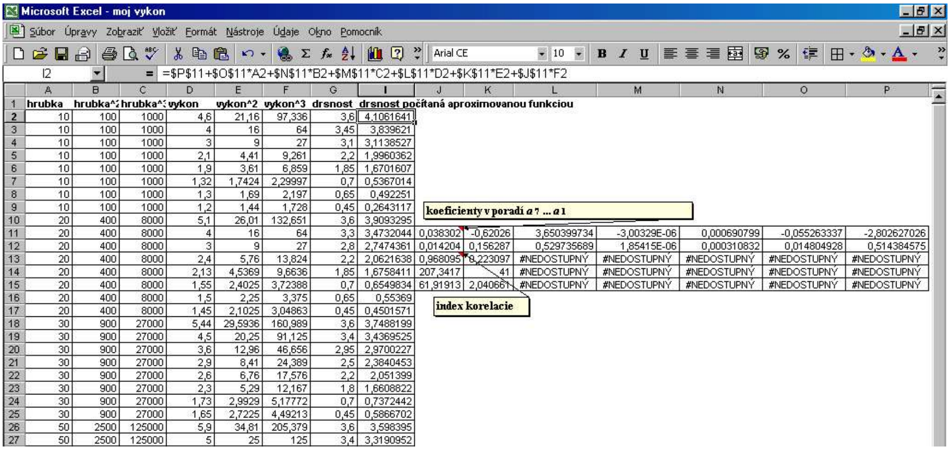
Funkciu znázorníme graficky na obrázku 2b.

Lineárna závislosť $R_a = f(l, P)$ tiež nedosahovala požadované hodnoty a navyše pri zobrazení $R_a = f(P)$ sme jasne videli nelineárnu závislosť, tak substitúciou môžeme vytvoriť pole $\left(l_i, l_i^2, l_i^3, P, P_i^2, P_i^3\right)$ a pomocou funkcie LINREGRESE (obr. 3) nájdeme funkciu

$$R_a = -2{,}802627 - 0{,}055263.l + 0{,}00069.l^2 - 3.10^{-6}.l^3$$

$$+ 3{,}6504.P - 0{,}62026.P \quad + 0{,}0383.P^3$$

s indexom korelácie 0,9839, ktorú znázorníme graficky na obrázku 4.

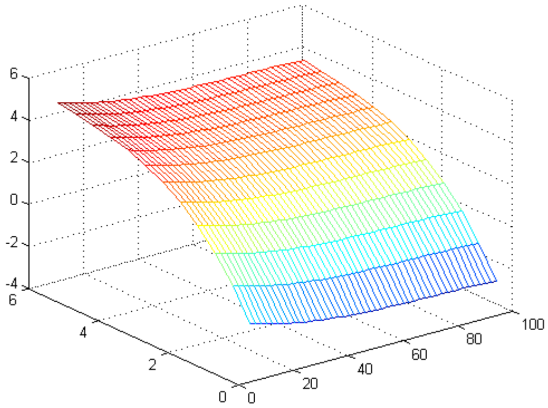

Microsoft Excel - moj vykon

Súbor Úpravy Zobraziť Vložiť Formát Nástroje Údaje Okno Pomocník

I2 = =$P$11+$O$11*A2+$N$11*B2+$M$11*C2+$L$11*D2+$K$11*E2+$J$11*F2

	A	B	C	D	E	F	G	I	J	K	L	M	N	O	P
1	hrubka	hrubka^2	hrubka^3	vykon	vykon^2	vykon^3	drsnost	drsnost počítaná aproximovanou funkciou							
2	10	100	1000	4,6	21,16	97,336	3,6	4,1061641							
3	10	100	1000	4	16	64	3,45	3,839621							
4	10	100	1000	3	9	27	3,1	3,1138527							
5	10	100	1000	2,1	4,41	9,261	2,2	1,9960362							
6	10	100	1000	1,9	3,61	6,859	1,85	1,6701607							
7	10	100	1000	1,32	1,7424	2,29997	0,7	0,5367014							
8	10	100	1000	1,3	1,69	2,197	0,65	0,492251							
9	10	100	1000	1,2	1,44	1,728	0,45	0,2643117							
10	20	400	8000	5,1	26,01	132,651	3,6	3,9093295							
11	20	400	8000	4	16	64	3,3	3,4732044	0,038302	-0,62026	3,650399734	-3,00329E-06	0,000690799	-0,055263337	-2,802627026
12	20	400	8000	3	9	27	2,8	2,7474361	0,014204	0,156287	0,529735689	1,85415E-06	0,000310832	0,014804928	0,514384575
13	20	400	8000	2,4	5,76	13,824	2,2	2,0621638	0,968095	0,223097	#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ
14	20	400	8000	2,13	4,5369	9,6636	1,85	1,6758411	207,3417	41	#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ
15	20	400	8000	1,55	2,4025	3,72388	0,7	0,6549834	61,91913	2,040661	#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ
16	20	400	8000	1,5	2,25	3,375	0,65	0,55369							
17	20	400	8000	1,45	2,1025	3,04863	0,45	0,4501571							
18	30	900	27000	5,44	29,5936	160,989	3,6	3,7488199							
19	30	900	27000	4,5	20,25	91,125	3,4	3,4369525							
20	30	900	27000	3,6	12,96	46,656	2,95	2,9700227							
21	30	900	27000	2,9	8,41	24,389	2,5	2,3840453							
22	30	900	27000	2,6	6,76	17,576	2,2	2,051399							
23	30	900	27000	2,3	5,29	12,167	1,8	1,6608822							
24	30	900	27000	1,73	2,9929	5,17772	0,7	0,7372442							
25	30	900	27000	1,65	2,7225	4,49213	0,45	0,5866702							
26	50	2500	125000	5,9	34,81	205,379	3,6	3,598395							
27	50	2500	125000	5	25	125	3,4	3,3190952							

obr. 3



Obr. 4

Záver

Z uvedeného vyplýva, že MS Excel má bohaté využitie pri riešení úloh nelineárnej regresie v praxi. Umožňuje kreatívne vytvárať rôzne typy funkčných závislostí funkcie viac premenných, prispieva k zvyšovaniu presnosti a rýchlosti riešenia, čo má potom za následok, že inžinier v praxi si môže vybrať z viacerých typov funkcií, podľa vlastných kritérií. Napríklad vysoký index korelácie, definičný obor, diferencovateľnosť, ďalšie vlastnosti a využitia funkcie.

Literatúra:

1. Využitie MS Excel-u vo vyučovaní numerickej matematiky. Zborník zo IV. Vedeckej konferencie doktorandov. UKF Nitra, 2003. ISBN 80-8050-582-9
2. Turek, I.: Kapitoly z didaktiky vysokej školy. Košice TU, 1998. ISBN 80-7099-322-7.
3. Straka, L.: Analýza automatizovaného riadenie výrobných technológií v aplikácii na EDM. Dizertačná práca FVT TU PO, 2003.
4. Orvis, J.: MS Excel pro vědce a inženýry. Praha Computer Press, 1996. ISBN 80-85896-49-4.

Adresa autora: Mgr. Jana Boržíková, KIMaF, Fakulta výrobných technológií TU KE so sídlom v Prešove, 080 01 Prešov. E-mail: borzikova.jana@FVT.sk, t. č.: 051/7723931.