

# **SEMESTRÁLNÍ PRÁCE**

Licenční studium **ARCHIMEDES**  
**STATISTICKÉ ZPRACOVÁNÍ DAT**

Předmět  
**4.4 APROXIMACE KŘIVEK A VYHLAZOVÁNÍ KŘIVEK**

Katedra ochrany lesa a myslivosti, ČZU v Praze

**Ing. Jan Lubojacký**

**Otázka 1.:** Napište matici pro případ lineárního regresního spline vyjádřeného přes useknuté polynomy pro případ dvou uzlových bodů  $k_1 = 1$  a  $k_2 = 4$ . Experimentální body jsou  $x = [0.2 \ 0.4 \ 0.6 \ 1.5 \ 2.0 \ 3.0 \ 3.5 \ 6.0 \ 8.0 \ 10.0]$ ,  $y = [1 \ 1 \ 2 \ 3 \ 4 \ 6 \ 6 \ 4 \ 2 \ 2]$ .

**Program:**

ADSTAT 1.25 – modul Kalibrace

**Vstupní data:**

Tab. 1 Vstupní data

x	y
0.2	1
0.4	1
0.6	2
1.5	3
2.0	4
3.0	5
3.5	5
6.0	4
8.0	2
10.0	2

**Řešení:**

Pomocí programu ADSTAT – modulu Kalibrace byla data zpracována metodou Lineární spline, přičemž byly zvoleny dva uzlové body:  $K [1] = 1$  a  $K [2] = 4$ . Následně byl získán matematický zápis tří nezávislých přímek.

**Výsledky:**

Výstup použité metody je následující:

(1) PARAMETRY KALIBRACE:

Koeficienty rovnice :  $g[i]*x+h[i]$  pro  $k[i-1] < x \leq k[i]$

k[i]	g[i]	h[i]
1.0000E+00	2.4745E+00	3.4240E-01
4.0000E+00	1.1706E+00	1.6463E+00
1.0000E+01	-8.5614E-01	9.7532E+00

(2) ANALÝZA REZIDUÍ:

Bod	Měřená hodnota	Predikovaná hodnota	Absolutní reziduum	Relativní reziduum
i	yexp[i]	yvyp[i]	e[i]	er[i]
1	1.0000E+00	8.3730E-01	-1.6270E-01	-1.9431E+01
2	1.0000E+00	1.3322E+00	3.3220E-01	2.4936E+01
3	2.0000E+00	1.8271E+00	-1.7290E-01	-9.4632E+00
4	3.0000E+00	3.4022E+00	4.0219E-01	1.1822E+01
5	4.0000E+00	3.9875E+00	-1.2515E-02	-3.1386E-01
6	6.0000E+00	5.1581E+00	-8.4193E-01	-1.6322E+01
7	6.0000E+00	5.7434E+00	-2.5663E-01	-4.4683E+00
8	4.0000E+00	4.6164E+00	6.1638E-01	1.3352E+01
9	2.0000E+00	2.9041E+00	9.0409E-01	3.1132E+01
10	2.0000E+00	1.1918E+00	-8.0819E-01	-6.7812E+01

Reziduální součet čtverců, RSC : 2.9538E+00  
 Průměr absolutních hodnot reziduí, Me : 4.5097E-01  
 Průměr relativních reziduí, Mer[%] : 19.905  
 Odhad reziduálního rozptylu,  $s^2(e)$  : 4.9230E-01  
 Odhad směrodatné odchylky reziduí,  $s(e)$  : 7.0164E-01

(3) ANALÝZA DERIVACÍ A INTEGRÁLŮ:

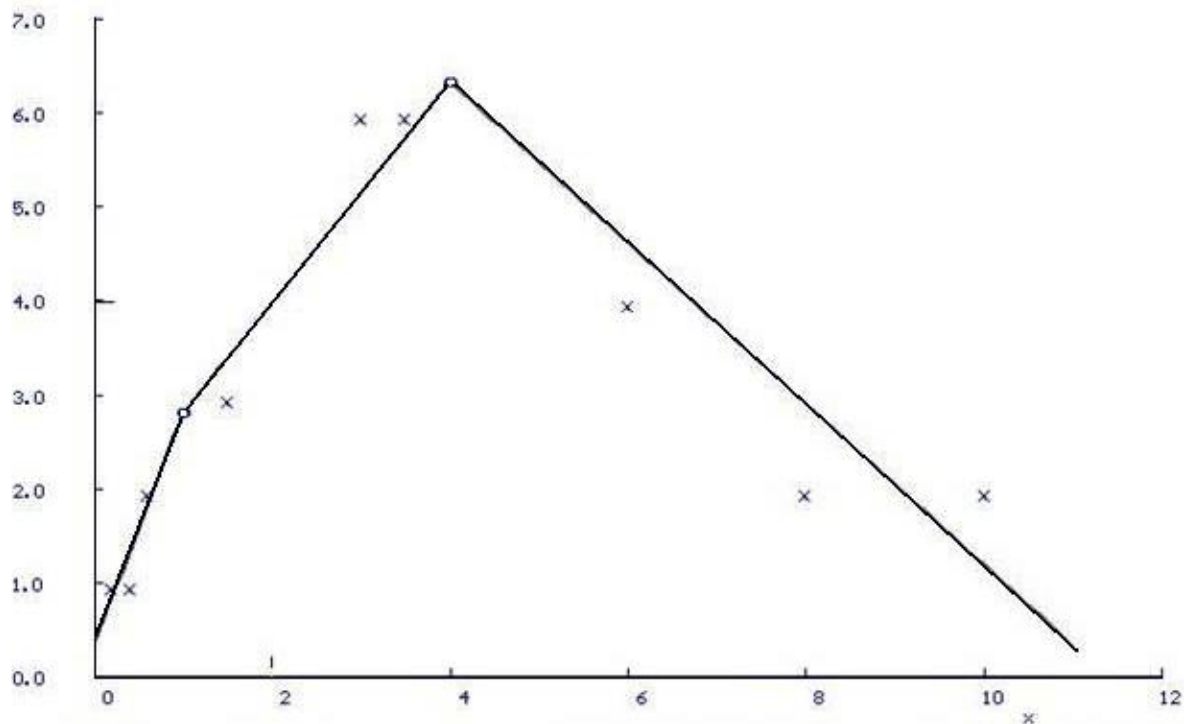
Bod	Predikovaná hodnota	První derivace	Druhá derivace	Integrál
i	Yvyp[i]	der1[i]	der2[i]	int[i]
1	8.3730E-01	2.4745E+00	0.0000E+00	0.0000E+00
2	1.3322E+00	2.4745E+00	0.0000E+00	2.1695E-01
3	1.8271E+00	2.4745E+00	0.0000E+00	5.3288E-01
4	3.4022E+00	1.1706E+00	0.0000E+00	3.0165E+00
5	3.9875E+00	1.1706E+00	0.0000E+00	4.8639E+00
6	5.1581E+00	1.1706E+00	0.0000E+00	9.4366E+00
7	5.7434E+00	1.1706E+00	0.0000E+00	1.2162E+01
8	4.6164E+00	-8.5614E-01	0.0000E+00	2.6125E+01
9	2.9041E+00	-8.5614E-01	0.0000E+00	3.3646E+01
10	1.1918E+00	-8.5614E-01	0.0000E+00	3.7741E+01

(4) KALIBRAČNÍ MEZE:

Kritická úroveň, yc: 2.159562E+00 xc: 8.869637E+00  
 Limita detekce, yd: 3.391457E+00 xd: 1.490831E+00

(4) KALIBRAČNÍ TABULKA:

Měřená hodnota	Inverzní odhad	Konfidenční interval	
		dolní mez	horní mez
yexp[i]	xvyp[i]	Llxvyp[i]	Luxvyp[i]
1.000000E+00	1.022404E+01	1.320657E+01	8.906301E+00
1.500000E+00	9.640025E+00	1.219206E+01	8.445860E+00
2.000000E+00	9.056010E+00	1.120022E+01	1.404270E+00
2.500000E+00	8.471995E+00	1.024345E+01	1.679348E+00
3.000000E+00	1.156421E+00	9.340587E+00	1.970675E+00
3.500000E+00	1.583556E+00	8.515223E+00	6.222915E+00
4.000000E+00	2.010691E+00	7.785205E+00	5.477987E+00
4.500000E+00	2.437827E+00	1.233587E+00	4.639412E+00
5.000000E+00	5.551919E+00	2.098578E+00	-----
5.500000E+00	4.967903E+00	2.610284E+00	-----



Obr. 1 Graf lineárního spline vytvořený v modulu kalibrace programu ADSTAT (upraveno)

Z Obr. 1 je patrné, že vstupní body byly proloženy třemi přímkami, což bylo způsobeno zadáním 2 uzlových bodů. Metodou nejmenších čtverců bylo potřeba nalézt takové přímky, které budou nejlépe proloženy danými body (přitom musí být spojitě). Parametry popisující aproximaci současně dávají koeficienty tří nezávislých přímek ve tvaru  $y = g[i] * x + h[i]$  pro každý úsek křivky samostatně. Koeficientem  $g$  je popisován počáteční bod přímky a koeficientem  $h$  pak sklon přímky. Koeficientem  $k$  je ve výsledné matici určován definiční obor funkce.

**Závěr:**

Zadanými dvěma uzlovými body byla výsledná křivka, prokládající vstupní body, rozdělena na tři spojitě přímky. Vypočítána byla matice koeficientů těchto tří přímek (viz Výsledky – bod (1))

**Otázka 2.:** Spočítejte integrál funkce  $1/(1+x^2)$  pro interval  $[0, 0.8]$  pomocí programu Späth.

**Program:**

ADSTAT – modul Vyhlazování

**Vstupní data:**

Tab. 2 Vstupní data

x	y
0	1.0000
0.1	0.9901
0.2	0.9615
0.3	0.9174
0.4	0.8621
0.5	0.8000
0.6	0.7353
0.7	0.6711
0.8	0.6098

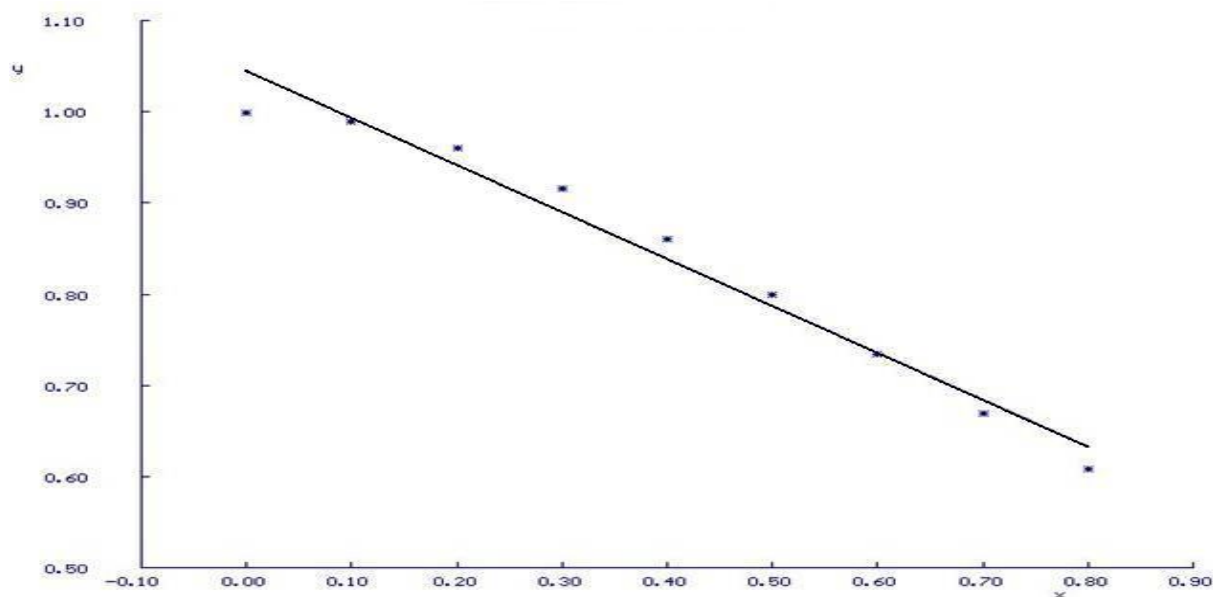
**Řešení:**

Nejprve bylo vypočítáno 9 vstupních hodnot a to tak, že ze zadaného intervalu bylo vybráno 9 hodnot  $x$ , které byly dosazeny do vstupní funkce a dopočítány hodnoty  $y$ .

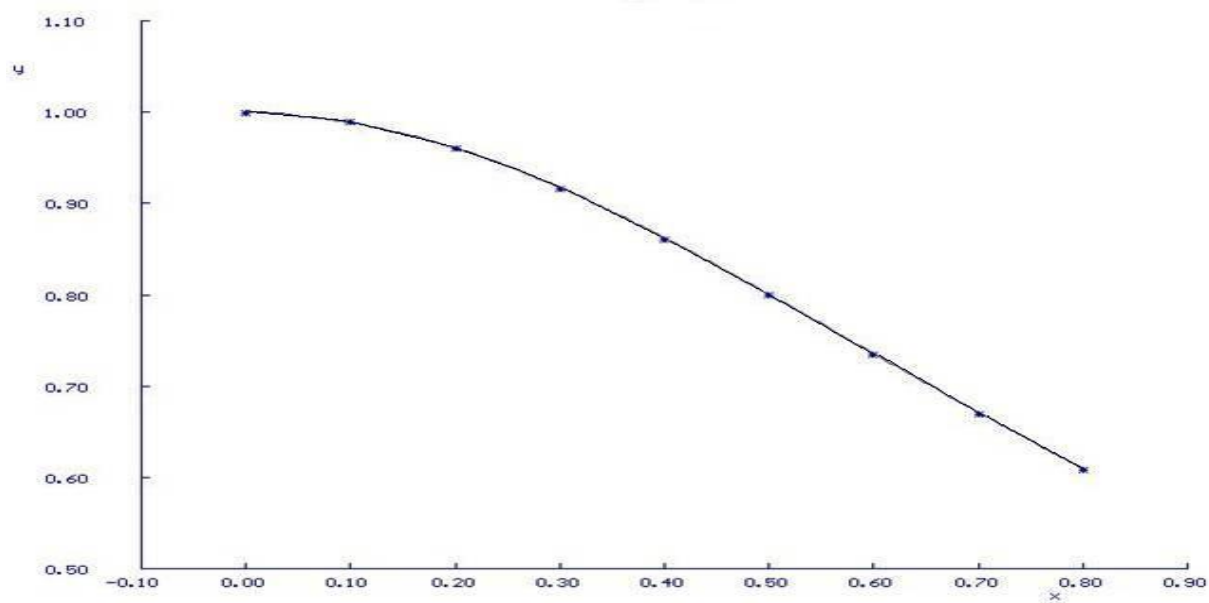
Tato vstupní data byla následně zadána do programu ADSTAT – modul Vyhlazování a zvolena metoda Späth. Bylo nutné zadat hodnotu  $P$ , která představuje lokální parametr vyhlazení. Je-li hodnota  $P$  blízká nule, výsledkem je regresní přímka, neboť dochází k potlačení podmínky hladkosti. Na druhou stranu, blíží-li se hodnota  $P$  nekonečnu, prochází křivka všemi body. Obě krajní možnosti jsou dobře patrné z Obr. 2 a 3.

**Výsledky:**

Z Obr. 2 a 3 je dobře patrné, že mnohem lepšího proložení bylo dosaženo při zadání vyšší hodnoty  $P$  ( $P = 10,000$ ).



Obr. 2 Späth spline vstupních hodnot při  $P = 0.02$



Obr. 2 Späth spline vstupních hodnot při P = 10,000

#### TABULKA DERIVACÍ A INTEGRÁLŮ:

Bod	Nezávisle proměnná	Závisle proměnná	První derivace	Druhá derivace	Integrál
i	x[i]	yexp[i]	der1[i]	der2[i]	int[i]
1	0.0000E+00	1.0000E+00	-1.0464E-01	0.0000E+00	0.0000E+00
2	1.0000E-01	9.9010E-01	-1.8949E-01	-1.6970E+00	9.9576E-02
3	2.0000E-01	9.6150E-01	-3.5927E-01	-1.6985E+00	1.9719E-01
4	3.0000E-01	9.1740E-01	-5.0257E-01	-1.1676E+00	2.9123E-01
5	4.0000E-01	8.6210E-01	-5.9443E-01	-6.6962E-01	3.8028E-01
6	5.0000E-01	8.0000E-01	-6.3989E-01	-2.3952E-01	4.6343E-01
7	6.0000E-01	7.3530E-01	-6.4701E-01	9.7170E-02	5.4022E-01
8	7.0000E-01	6.7110E-01	-6.2886E-01	2.6576E-01	6.1055E-01
9	8.0000E-01	6.0980E-01	-6.1557E-01	0.0000E+00	6.7459E-01

Výsledná hodnota integrálu byla vypočtena jako:

$$\int_0^{0.8} (1 / (1 + x^2)) = 0.675 - 0 = 0.675$$

#### Závěr:

Hodnota integrálu  $1/(1 + x^2)$  pro interval  $[0, 0.8]$  je rovna 0.675. Tato hodnota představuje obsah plochy pod křivkou integrálu, která je ohraničena intervalem  $[0, 0.8]$ .

**Otázka 3.:** Uved'te základní rozdíly mezi spline vyhlazováním a spline regresí.

K řešení úloh založených na aproximaci závislostí existuje řada možností:

- (a) aproximace polynomy pro zvolenou normu  $S_p$ ,
- (b) spline regrese s volbou stupně spline a uzlových bodů,
- (c) úseková regrese,
- (d) spline vyhlazování s volbou parametrů vyhlazení,
- (e) neparametrická regrese s volbou parametrů vyhlazení,
- (f) číslicová filtrace pro ekvidistantní dělení souřadnic  $x$  s volbou stupně a délky filtru.

Výběr vhodného postupu závisí na cíli zpracování dat. Odstranění náhodných šumů umožňují prakticky všechny postupy a liší se zejména definicí aproximující funkce a její složitostí.

Pro účely tvorby empirických modelů vyhovuje obvykle nejlépe **spline regrese**. Uzlové body tvoří hranice intervalů, pro něž jsou definovány jednotlivé funkce.

**Spline vyhlazování** je účelné pro případy, u kterých budou data následně numericky derivována nebo integrována. Pro spline vyhlazování je charakteristické, že uzlové body jsou totožné s  $x$ -ovými souřadnicemi zadaných experimentálních dat. Požadavkem je, aby vyhlazující funkce byla dostatečně hladká a spojitá ve zvoleném počtu derivací.

**Otázka 4.:** Co můžeme říci o filtru 53H?

**Řešení a závěr:**

Jedná se o typ robustních nelineárních filtrů, které jsou používány v případech, kdy lze v datech očekávat i hrubé nenáhodné chyby (tzv. outliers). Tyto filtry byly vytvořeny jako varianty robustních vychlazovacích metod.

Filtr 53H je jednoduchý filtr, jež je dán výrazem:

$$Z_i = S(5, i-2)/4 + S(5, i-1)/2 + S(5, i)/4$$

Jde o kombinaci mediánu a tří hodnot (bodů). Filtr 53H je považován za jeden z nevhodnějších filtrů, neboť je dostatečně robustní a přitom neposkytuje „nadměrně“ vyhlazené úseky.