Licenčné štúdium Archimedes

Aproximácia a vyhladzovanie kriviek

PriFUK, katedra jadrovej chémie
Bratislava 07.02.2012 RNDr. Silvia Dulanská, PhD

Obsah

Úloha **1.** Napište matici pro případ lineárního regresního spline vyjádřeného přes useknuté polynomy pro případ dvou uzlových bodů ksí 1 = 1 a ksí 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].

Použitý program: AdStat

Zdrojová dáta , ADSTAT

```
 F1=nápověda   F2=uložení   F3=čtení    F4=numerický mód
Počet řádků:  10   Počet sloupců:    2    Vodorovně
                                  Numerický editor

       2.000000 E-01       1.000000
       4.000000 E-01       1.000000
       6.000000 E-01       2.000000
       1.500000            3.000000
       2.000000            4.000000
       3.000000            6.000000
       3.500000            6.000000
       6.000000            4.000000
       8.000000            2.000000
       1.000000 E+01       2.000000
```

V programu Adstat boli v module Kalibrace postupne zadávané jednolivé spliny spolu so zvolenými dvomi uzlovými bodmi ξ_1 (1) a ξ_2 (4).

```
                     MODUL KALIBRACE : Lineární spline
  Data      Metoda     Zadání       Výpočet      Výsledky     Graf        Konec
                       Podmínky
                      P O D M Í N K Y   V Ý P O Č T U

 Název : _

 Hladina významnosti: 0.050                Výstupní soubor   : RESULTS .TXT
 Uzly   : Zadávání uzlů                     Počet uzlů        : 2

 Hodnoty uzlů  :
  a      :  2.000000E-01    k[ 1] :  1.000000E+00    k[ 2] :  4.000000E+00
  b      :  1.000000E+01
```

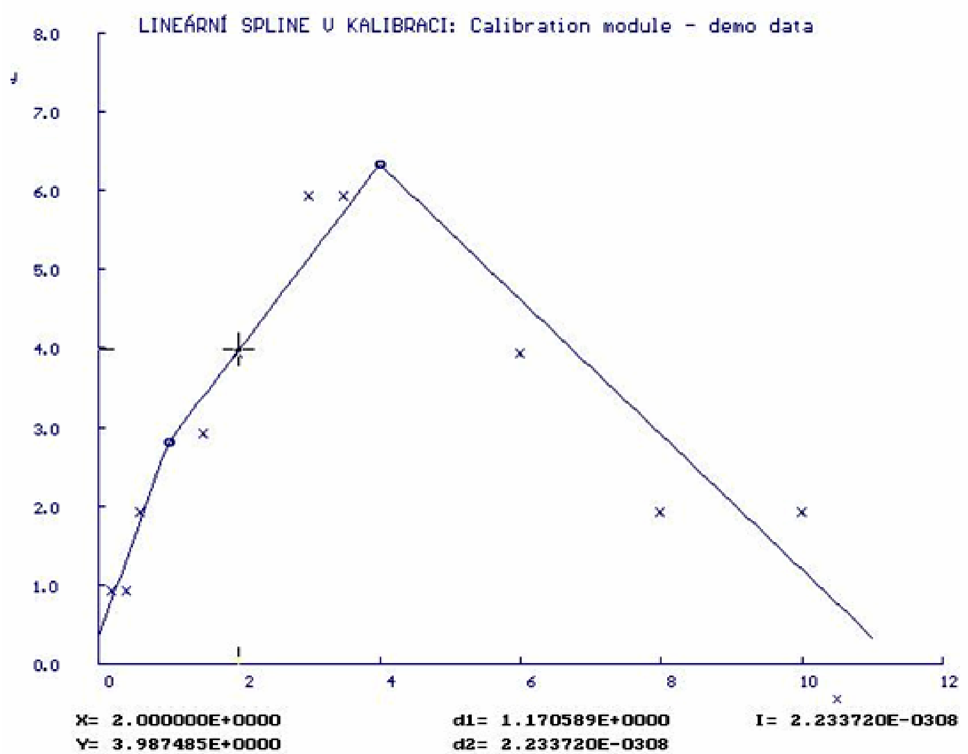
Výsledky

Lineárny spline

```
PARAMETRY KALIBRACE:
Koeficienty rovnice : g[i]*x+h[i] pro k[i-1] < x <= 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
```

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



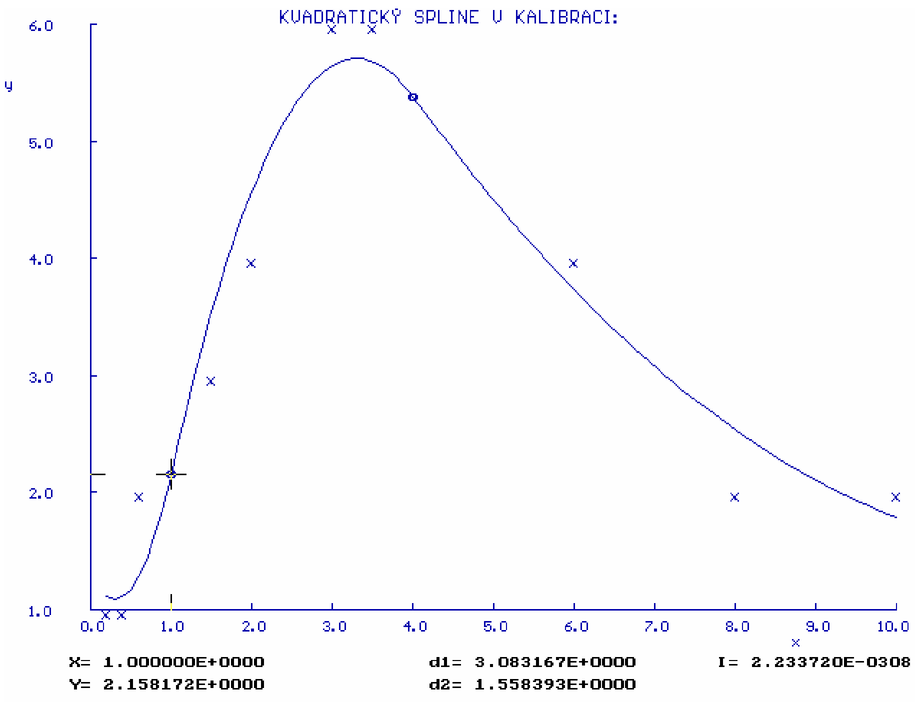
Graf 1 Lineárny spline, Astat.

Kvadratický spline

PARAMETRY KALIBRACE:
Koeficienty rovnice : f[i]*x^2+g[i]*x+h[i] pro k[i-1] < x <= k[i]

k[i]	f[i]	g[i]	h[i]
1.0000E+00	2.2282E+00	-1.3731E+00	1.3031E+00
4.0000E+00	-6.6984E-01	4.4230E+00	-1.5950E+00
1.0000E+01	5.6266E-02	-1.3858E+00	1.0023E+01

Reziduální součet čtverců, RSC : 1.7840E+00
Průměr absolutních hodnot reziduí, Me : 3.7380E-01
Průměr relativních reziduí, Mer[%] : 15.602
Odhad reziduálního rozptylu, s^2(e) : 3.5680E-01
Odhad směrodatné odchylky reziduí, s(e) : 5.9733E-01

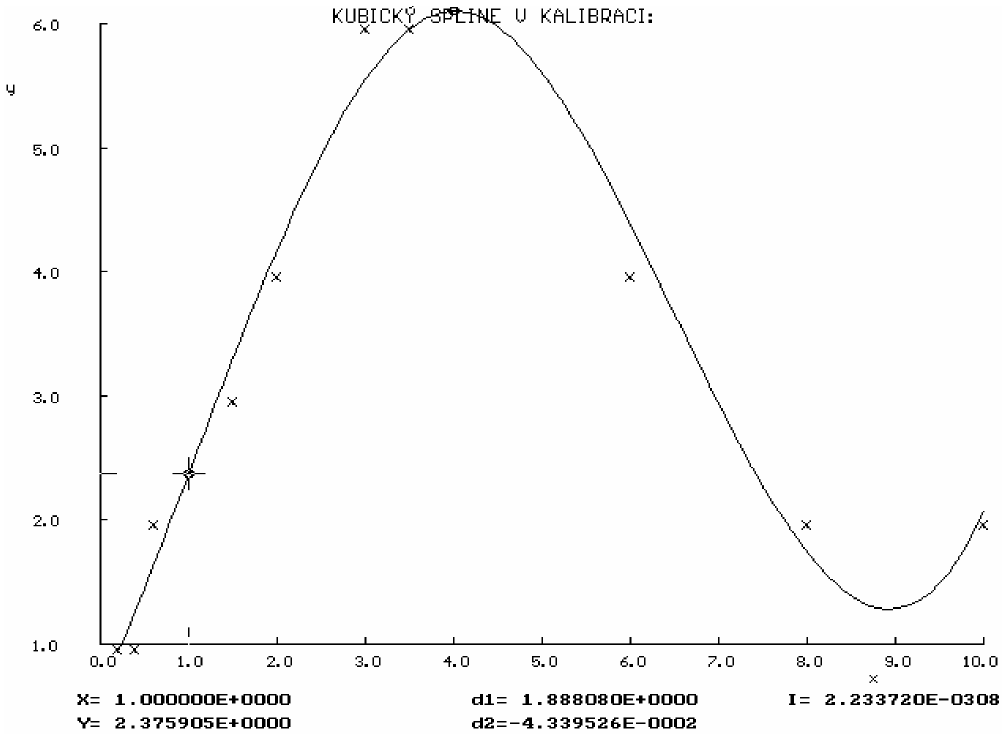


Graf 2 Kvadratický spline, AdStat.

PARAMETRY KALIBRACE:
Koeficienty rovnice : e[i]*x^3+f[i]*x^2+g[i]*x+h[i] pro k[i-1] < x <= k[i]

k[i]	e[i]	f[i]	g[i]	h[i]
1.0000E+00	-1.5890E-01	4.5494E-01	1.4549E+00	6.2496E-01
4.0000E+00	-6.4434E-02	1.7153E-01	1.7383E+00	5.3049E-01
1.0000E+01	8.1078E-02	-1.5746E+00	8.7229E+00	-8.7823E+00

Reziduální součet čtverců, RSC	: 7.6537E-01
Průměr absolutních hodnot reziduí, Me	: 2.3986E-01
Průměr relativních reziduí, Mer[%]	: 10.064
Odhad reziduálního rozptylu, s^2(e)	: 1.9134E-01
Odhad směrodatné odchylky reziduí, s(e)	: 4.3743E-01



Graf 3 Kubický spline, AdStat.

Záver

Pomocou programu AdStat boli porovnané jednotlivé grafy vybraných aproximácii – lineárna, kubická a kvadratická. Bola nájdená najlepšia aproximácia pre expermintálne dáta a to aproximácia triedy C^2 (kubický spline). Priebeh u tejto aproximácii je hladký, aproximujúca krivka najlepšie prekladá experimentálne body a má najužší interval spoľahlivosti. Výhodou kubických spline, je že sú spojité v prvých dvoch deriváciách, čo umožňuje zostrojenie hladkej krivky v prvej derivácii.

Vo výstupe programu Adstat boli uvedené aj jednotlivé parametry pre popis daných aproximácie u každém aproximácii .

Veľkosť matice ako je uvedené v zadaní otázky je možné vypočítať podľa postupu Meloun, Militký[1]

$$
\begin{pmatrix}
n & \sum x_i & \sum (x_i - \xi_1)_+ & \sum (x_i - \xi_2)_+ \\
\sum x_i & \sum x_i^2 & \sum x_i (x_i - \xi_1)_+ & \sum x_i (x_i - \xi_2)_+ \\
\sum (x_i - \xi_1)_+ & \sum x_i (x_i - \xi_1)_+ & \sum (x_i - \xi_1)_+ (x_i - \xi_2)_+ & 0 \\
\sum (x_i - \xi_2)_+ & \sum x_i (x_i - \xi_2)_+ & 0 & 0
\end{pmatrix} =
$$

$$
= \begin{pmatrix}
10{,}00 & 35{,}20 & 25{,}20 & -4{,}80 \\
35{,}20 & 228{,}02 & 192{,}86 & 87{,}26 \\
25{,}20 & 192{,}86 & 92{,}06 & 0 \\
-4{,}80 & 87{,}26 & 0 & 0
\end{pmatrix}
$$

[1] Meloun, Miltky, 857 str. Kompendium statistického spracovaní dat. Academia. 2004.

Úloha **2.** *Spočítejte integrál funkce 1/(1+x^2) pro interval [0,0.8] programem Spaeth*

Riešenie

Spaeth použil pri konštrukcii algoritmu pre vyhladenie kubický spline postup[2], ktorý vychádza z rovnice pre *g*. Vyhladzujúci spline vyjadril v tvare lokálních kubických polynómov a k riešeniu piatich diagonálních sústav lineárních rovníc využil kompaktné varianty Choleského metódy. V programu Spaeth sú použité lokálne parametry vyhladenia pre

a) $\beta \rightarrow 0$, i= 1,…..,n je potlačená podmienka hladkosti a získáme funkciu *g(x)* ako regresná priamka

b) $\beta \rightarrow \infty$, i= 1,…..,n, vychádza funkcia $g(x)$ ako kubické interpolačne spline $S_3(x)$.

Pre zvolený interval [0,0.8] sa vypočítali príslušné *y* hodnoty a dosadili do vstupných parametrov v programe „Vyhladzovanie“ ADSTAT. Pre názornú ukážku sú uvedené výsledky a grafy splinov s rozdielnymi parametrami P=0,01 a P=10000.

Použitý program

Adstat, zdrojové dáta

```
F1=nápověda  F2=uložení  F3=čtení  F4=numerický mód  F5=zoom  ESC=konec
Počet řádků:  9  Počet sloupců:  2  Vodorovně  demo.dat
                         Numerický editor

      0                    1.000000
      1.000000 E-01        9.901000 E-01
      2.000000 E-01        9.615000 E-01
      3.000000 E-01        9.174000 E-01
      4.000000 E-01        8.621000 E-01
      5.000000 E-01        8.000000 E-01
      6.000000 E-01        7.353000 E-01
      7.000000 E-01        6.711000 E-01
      8.000000 E-01        6.098000 E-01
```

Zvolenie metódy (ADSTAT)

[2] Meloun, Militký, Kompendium statistického spracovaní dat. Academia , 871 s.

```
VYHLAZOVÁNÍ : Späth Spline
Data    Metoda    Zadání    Výpočet    Výsledky    Grafy    Konec

        Reinsch spline
        Späth spline
        Savitzki-Golay spline
```

a) Podmienky vyhladzovania, parameter vyhladenia P =0,01

```
VYHLAZOVÁNÍ : Späth Spline
Data    Metoda    Zadání    Výpočet    Výsledky    Grafy    Konec

                  Podmínky
                  Volby

P O D M Í N K Y   V Ý P O Č T U

Název: Demo data

Parametr vyhlazení       P  :                   1.0000E-02
                      Podmínky dobře ?  [A] _
```

Výstup (P=0,01)

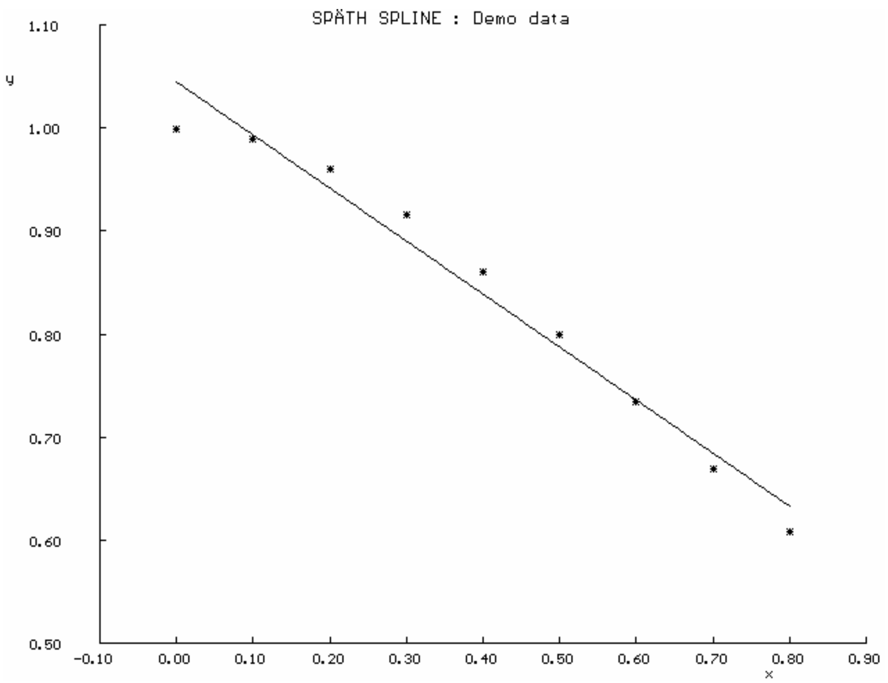
V tabuľke sú uvedené hodnoty vyhladenej funkcie, prvej derivácie, druhej derivácie a integrálu v jednotlivých bodoch pre daný prípad , podmienky vyhladenia P=0,01.

(2) TABULKA DERIVACÍ A INTEGRÁLŮ:

Bod i	Nezávisle proměnná x[i]	Závisle proměnná yexp[i]	První derivace der1[i]	Druhá derivace der2[i]	Integrál int[i]
1	0.0000E+00	1.0000E+00	-5.1457E-01	0.0000E+00	0.0000E+00
2	1.0000E-01	9.9010E-01	-5.1457E-01	-4.4425E-05	1.0187E-01
3	2.0000E-01	9.6150E-01	-5.1458E-01	-9.1717E-05	1.9859E-01
4	3.0000E-01	9.1740E-01	-5.1459E-01	-1.1902E-04	2.9017E-01
5	4.0000E-01	8.6210E-01	-5.1460E-01	-1.1897E-04	3.7660E-01
6	5.0000E-01	8.0000E-01	-5.1461E-01	-9.5420E-05	4.5789E-01
7	6.0000E-01	7.3530E-01	-5.1462E-01	-5.8998E-05	5.3403E-01
8	7.0000E-01	6.7110E-01	-5.1462E-01	-2.2945E-05	6.0502E-01
9	8.0000E-01	6.0980E-01	-5.1463E-01	-3.3881E-21	6.7087E-01

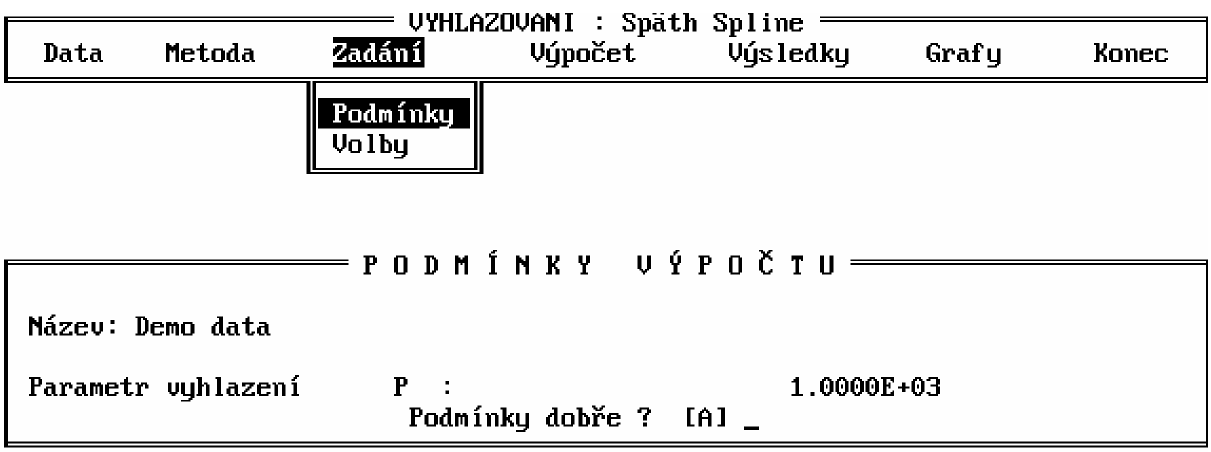
ANALÝZA REZIDUÍ:

Průměr residuí	:	1.2336E-17
Průměr absolutních hodnot residuí	:	1.8603E-02
Reziduální směrodatná odchylka	:	2.2475E-02
Reziduální rozptyl	:	5.0511E-04
Šikmost reziduí	:	5.7225E-01
Špičatost reziduí	:	2.2880E+00



Graf 4 Vyhladzovanie spline, parameter vyhladzovania P=0,01.

b) Voľba parametru P= 0,001

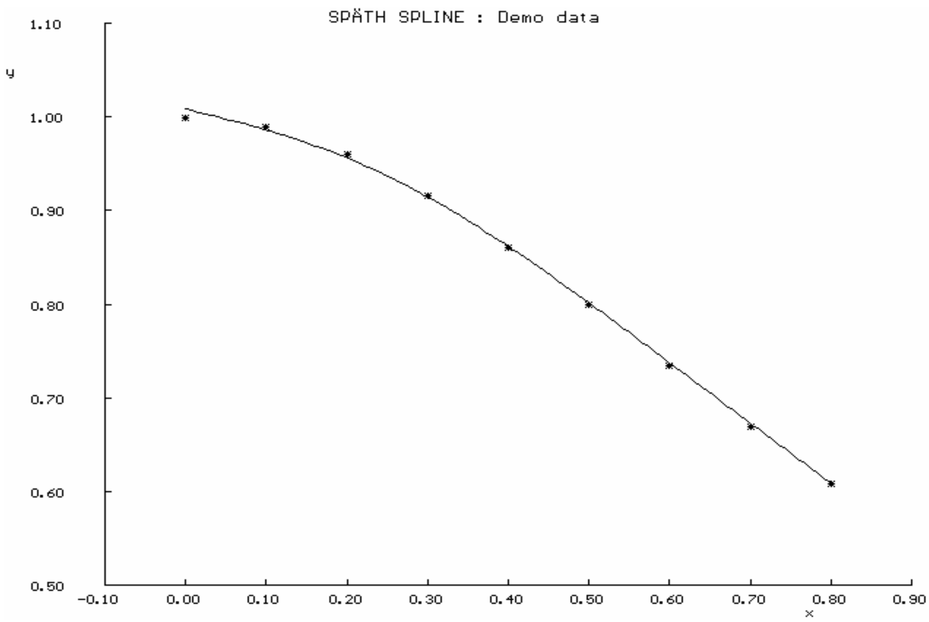


(2) TABULKA DERIVACÍ A INTEGRÁLŮ:

Bod i	Nezávisle proměnná x[i]	Závisle proměnná yexp[i]	První derivace der1[i]	Druhá derivace der2[i]	Integrál int[i]
1	0.0000E+00	1.0000E+00	-2.0868E-01	0.0000E+00	0.0000E+00
2	1.0000E-01	9.9010E-01	-2.5035E-01	-8.3341E-01	9.9755E-02
3	2.0000E-01	9.6150E-01	-3.5525E-01	-1.2646E+00	1.9695E-01
4	3.0000E-01	9.1740E-01	-4.7655E-01	-1.1614E+00	2.9059E-01
5	4.0000E-01	8.6210E-01	-5.7294E-01	-7.6643E-01	3.7948E-01
6	5.0000E-01	8.0000E-01	-6.2772E-01	-3.2922E-01	4.6267E-01
7	6.0000E-01	7.3530E-01	-6.4519E-01	-2.0064E-02	5.3962E-01
8	7.0000E-01	6.7110E-01	-6.4212E-01	8.1346E-02	6.1013E-01
9	8.0000E-01	6.0980E-01	-6.3805E-01	-1.3878E-17	6.7422E-01

ANALÝZA REZIDUÍ:

Průměr residuí	: -1.2336E-17
Průměr absolutních hodnot residuí	: 3.0044E-03
Reziduální směrodatná odchylka	: 3.8419E-03
Reziduální rozptyl	: 1.4760E-05
Šikmost reziduí	: 6.9158E-01
Špičatost reziduí	: 3.0635E+00



Graf 5 Vyhladzovanie spline, parameter vyhladzovania P= 1000.

Záver

Pomocou programu ADSTAT v modulu Spaeth sa porovnaním dvoch parametrov vyhladenia našlo vhodné vyhladenie analyzovanej krivky. Zistilo sa, že pri P= 0,01 bola potlačená podmienka hladkosti a získali sme funkciu *g(x)* ako regresnú priamku. Pre P=1000 vychádza funkcia *g(x)* ako kubické interpolačné spline $S_3(x)$ (Graf 5), čo je pre dané riešenie vhodnejšie.

$$\int_0^{0.8}\left(\frac{1}{1+x^2}\right) = 0{,}6746 - 0 = 0{,}6746$$

Integrál funkce $1/(1+x^2)$ pre interval [0,0.8] sa rovná číslu 0,6746.

Úloha 3. *Uvedte základní rozdíly mezi spline vyhlazováním a spline regresí.*

Spline vyhlazovanie je súčasť numerického vyhlazovania, ktorá sa používa pre odstránenie náhodných šumov ε_i v model $y_i = g(x_i) + \varepsilon_i$. Pre spline vyhlazovanie je charakteristické, že uzlové body ξ_i sú totožné so súradnicami x zadaných experimentálnych hodnôt $\{x_i, y_i\}$ $i = 1, ..., n$. Na rozdiel od spline regresie spline nemusia priamo prechádzať uzlovými bodmi. Vyhladzujúca funkcia by mala byť dostatočne hladká a spojitá vo zvolenom počte derivácií (najčastejším príkladom je keď požadujeme, aby funkcia $g(x)$ bola spojitá v 1. a 2. derivácii. Hlavnou úlohou je rekonštruovanie bezšumovej závislosti a až potom o tvar aproximujúcej funkcie $g(x)$.Spline vyhladzovanie používáme najčastejšie v prípade velkého počtu uzlových bodov.

Spline regresi je špecifickým prípadom vyhladzovania pri malom počte uzlových bodov ξ_j, ktoré sú definované mimo experimentálních bodov. V spline regresii sa snažíme preložiť body po častiach v definovanej funkcii. Uzlové body tvoria teda hranice intervalu, kde sú definované funkcie, pričom ako už bolo povedané uzlovými bodmi priamka neprechádza. Hladkosť aproximujúcej funkcie závisí na triede C^m. Podmienku spojitosti vo funkčných hodnotách splňujú polynómy triedy m+1. Pro triedu C^0 sa jedná o lineárny spline, pre triedu C^1 o kvadratický spline a C^2 kubický spline. Výsledkem spline regresie je funkčný predpis spolu s koeficientami.

Úloha 4. Co můžeme říct o filtru 53H.

Rekurzívne filtre sa používajú k vyhladeniu časových rad a vstupov do číslicových regulátorov. Na základe simulačného štúdia bolo zistené, že medzi najvhodnejšie patrí filter H53, ktorý je dostatočne robusný a pritom neposkytuje nadmerne vyhladené úseky[3]. V literatúre Meloun, Militký[1,4] sa porovnávali vlastnosti lineárnych a nelineárnych filtrov na sínusoidálnych dátach zaťažených ako náhodnými normálne rozdelenými chybami, ale aj hrubými chybami. Z nelineárnych filtrov sa ako najvhodnejší potvrdil filter H53, ktorý možno jednoducho zaradiť do programu pre predspracovanie analytických signálov ak sú očakávané vyskytujúce sa hrubé chyby. Hrubé chyby pri vyhladzovaní pomocou Hippeho filtru a filtru 3T spôsobili na rozdiel od filtru H53 značné prekmitávanie a ani vyhladenie pre gaussovské chyby nebolo dokonalé. Z tohto porovnania sa ako najvhodnejší potvrdil nelineárny filter H53. Filter 53H je teda robustný nelineární filtr, necitlivý k hrubým , je daný výrazom:

$$Zi = \frac{S(5,i-2)}{4} + \frac{S(5,i-1)}{2} + \frac{S(5,i)}{4}$$

K vypočítanému mediánu *i*-tého bodu sa počíta s dvomi predchádzajúcimi.

3 Meloun, Militký, 928s, 2004.