Univerzita
Pardubice
Fakulta
chemicko-technologická

# LICENČNÍ STUDIUM ARCHIMEDES

Semestrální práce z předmětu 4.4:

Aproximace křivek a vyhlazování křivek

**Ing. Tomáš Vítek, Ph. D.**
Oddělení rybářství a hydrobiologie
Mendelova univerzita v Brně

**Otázka 1:** Napište matici pro případ lineárního regresního spline vyjádřeného přes useknuté polynomy pro případ dvou uzlových bodů $\xi_1 = 1$ a $\xi_2 = 4$. 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].

**Řešení:**

Jako aproximační model bude použita složená polynomická funkce. Každý polynom bude definován v intervalu $I_j$ mezi uzlovými body $\xi_{j-1}$, $\xi_j$ a ve zbylých intervalech nulový. V případě lineárního spline $C^0$ (tři úsekové polynomy prvního stupně, tedy přímky) tak získáme průběh funkce spojitý, ale nikoliv vyhlazený. Pro hladký průběh by bylo nutné použít kvadratický spline $C^1$ (polynomy druhého stupně), který je spojitý ve své první derivaci. Pro případ ještě lepšího průběhu lze využít kubický spline $C^2$, funkce pak bude tvořena třemi polynomy třetího stupně a bude vyhlazená i ve své první derivaci a spojitá ve druhé derivaci.

Do software ADSTAT v modulu Kalibrace zadáme hodnoty signálu (*x*), a odezvy (*y*), zvolíme lineární spline, zadáme ručně uzlové body a provedeme výpočet včetně konstrukce grafu.

Pro případ použití polynomů $C^0$ vypočte software následující parametry (soubor *uloha1A.txt*):

PARAMETRY KALIBRACE:

Koeficienty rovnice: $g_i*x+h_i$ pro $k_{(i-1)} < x \leq k_i$

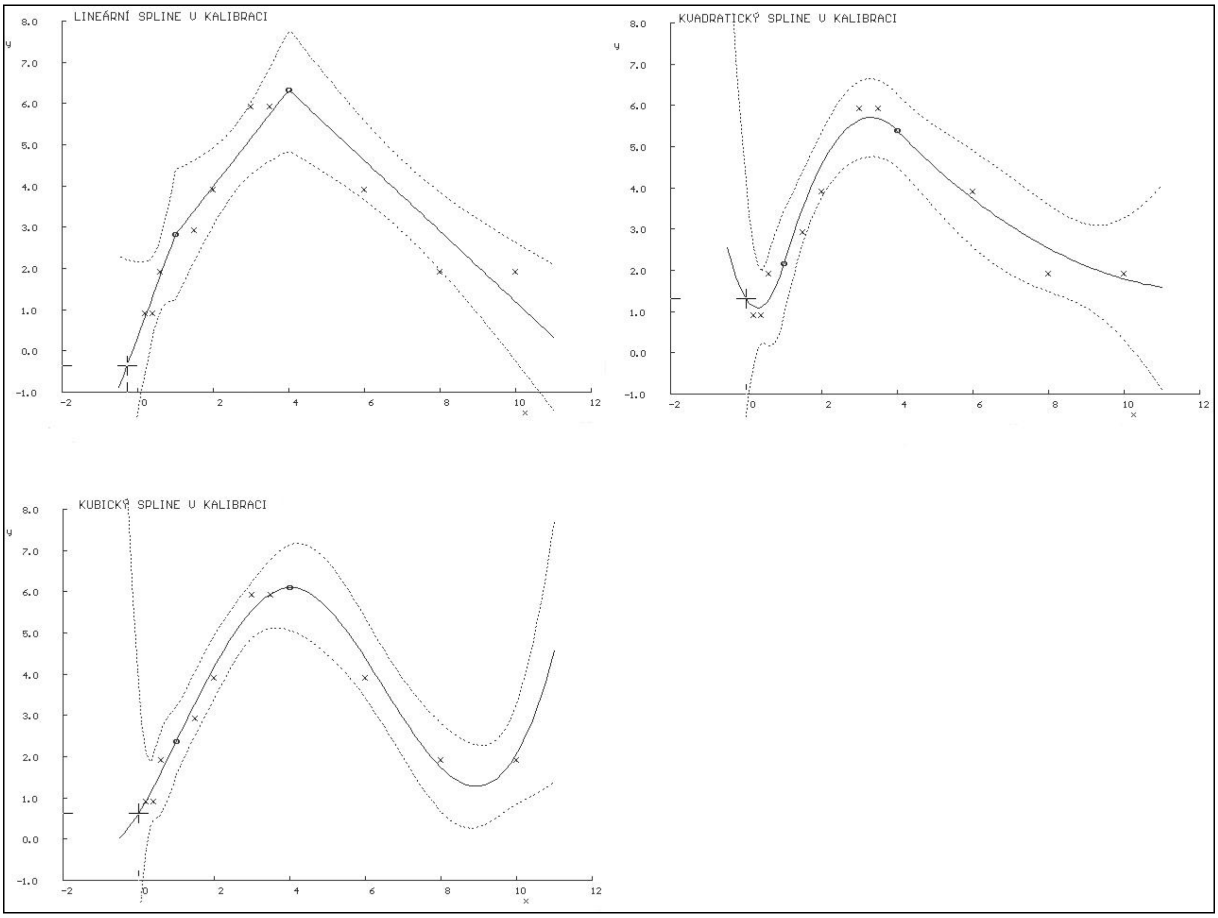
| $k_i$ | $g_i$ | $h_i$ |
|---|---|---|
| 1.0000E+00 | 2.4745E+00 | 3.4240E-01 |
| 4.0000E+00 | 1.1706E+00 | 1.6463E+00 |
| 1.0000E+01 | -8.5614E-01 | 9.7532E+00 |

ANALÝZA REZIDUÍ:

| Bod | Měřená Hodnota | Predikovaná hodnota | Absolutní reziduum | Relativní reziduum |
|---|---|---|---|---|
| $i$ | $y_{(i)exp}$ | $y_{(i)exp}$ | $e_{(i)}$ | $e_{(i)r}$ |
| 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, $M_e$: | 4.5097E-01 |
| Průměr relativních reziduí, $M_{er}$ (%): | 19.905 |
| Odhad reziduálního rozptylu, $s^2(e)$: | 4.9230E-01 |
| Odhad směrodatné odchylky reziduí, $s(e)$: | 7.0164E-01 |

Obr. 1: Aproximace funkce metodou interpolace přes useknuté polynomy v případě použití lineárního, kvadratického a kubického spline (soubory *uloha1A.txt, uloha1B.txt, uloha1C.txt, graph01.jpg*):




Z obrázku je patrné, že nejlepšího výsledku je dosaženo aproximací třídy $C^2$ (kubický spline). Průběh funkce je dokonale hladký, výborně prokládá experimentální data a intervaly spolehlivosti jsou nejužší. Uvedený způsob aproximace není vhodný k predikci mimo interval vložených hodnot $x$.

**Otázka 2:** Spočítejte integrál funkce $1/(1 + x^2)$ pro interval [0; 0,8] pomocí programu Späth.

**Řešení:**
Vyhlazující funkce aproximuje experimentální body ve smyslu nejmenších čtverců odchylek. Späthův postup aproximace využívá vyhlazující kubický spline, který je řešením pětidiagonální soustavy lineárních rovnic pomocí Choleského metody. Uživatel zadává vyhlazující parametr $\beta$.

Pro $\beta \to 0$ je potlačena podmínka hladkosti a výsledkem je regresní přímka.
A pro $\beta \to \infty$ je výsledkem kubický spline $S_3(x)$.

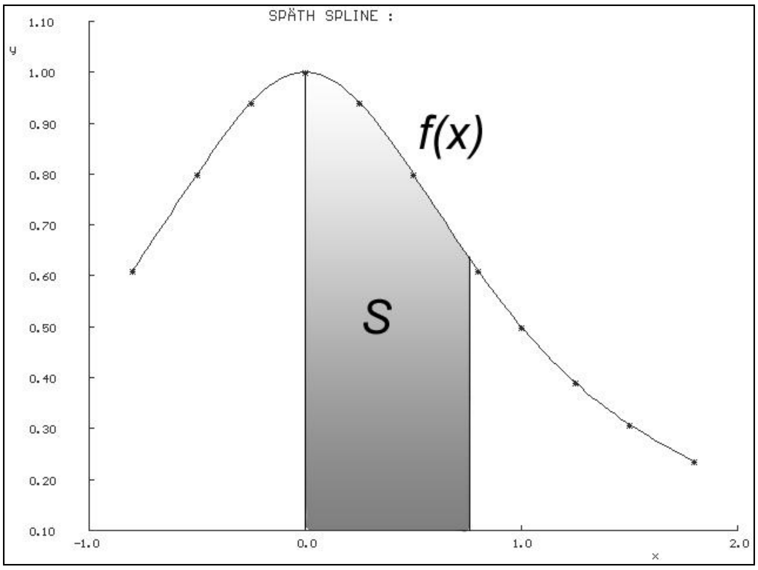
Na zadanou funkci aplikujeme metodu vyhlazení dle Spätha v software ADSTAT, modul vyhlazování, zvolíme velký parametr vyhlazení (např. $10^5$, soubor *uloha2.txt*).
Pokud aplikujeme uvedený postup na následující data, výsledný spline je uveden na obr. 2 (soubor *graph02.jpg*). Integrál v intervalu [0; 0,8] je plochou pod křivkou k ose $x$. Rovněž jej

můžeme určit z outputu software, který v bloku 2 uvádí hodnoty integrálu ve zvolených funkčních hodnotách.

Data použitá pro spline vyhlazovaní:

| *i* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| *x* | -0,8 | -0,5 | -0,25 | 0 | 0,25 | 0,5 | 0,8 | 1 | 1,25 | 1,5 |
| *y* | 0,61 | 0,80 | 0,94 | 1,00 | 0,94 | 0,80 | 0,61 | 0,50 | 0,39 | 0,31 |

Obr. 2: Výsledný spline (soubor *graph02.jpg*) s označením plochy vyjadřující integrál v intervalu [0; 0,8].




Tabulka derivací a integrálů vyhlazené funkce (soubor *uloha2.txt*):

TABULKA DERIVACÍ A INTEGRÁLŮ:

| $i$ | $x_i$ | $y_i$ | $g^{(1)}(x_i)$ | $g^{(2)}(x_i)$ | $I(x_i)$ |
|---|---|---|---|---|---|
| 1 | -8.0000E-01 | 6.0980E-01 | 6.3687E-01 | 0.0000E+00 | 0.0000E+00 |
| 2 | -5.0000E-01 | 8.0000E-01 | 6.2877E-01 | -5.4056E-02 | 2.1154E-01 |
| 3 | -2.5000E-01 | 9.4120E-01 | 4.4264E-01 | -1.4349E+00 | 4.3016E-01 |
| 4 | **0.0000E+00** | **1.0000E+00** | **-6.5166E-04** | **-2.1115E+00** | **6.7511E-01** |
| 5 | 2.5000E-01 | 9.4120E-01 | -4.3994E-01 | -1.4029E+00 | 9.2004E-01 |
| 6 | 5.0000E-01 | 8.0000E-01 | -6.3858E-01 | -1.8623E-01 | 1.1387E+00 |
| 7 | **8.0000E-01** | **6.0980E-01** | **-5.9697E-01** | **4.6363E-01** | **1.3499E+00** |
| 8 | 1.0000E+00 | 5.0000E-01 | -4.9962E-01 | 5.0989E-01 | 1.4605E+00 |

$$\int_0^{0,8} g(x) = 1,3499 - 0,6751 = 0,6748$$

**Otázka 3:** Uveďte základní rozdíly mezi spline vyhlazováním a spline regresí.

**Řešení:**

Při spline vyhlazování nejde v pravém slova smyslu o nalezení konkrétního funkčního tvaru aproximující funkce, ale o nalezení rekonstruované bezšumné závislosti. Uzlové body $\xi_j$ jsou totožné se souřadnicemi $x$ zadaných experimentálních dat. Podmínkou aproximující funkce je dostatečná hladkost při současně co nejmenších čtvercových odchylkách od experimentálních bodů. Podmínka hladkosti je vyjádřena pomocí integrálu a souvisí s normou druhé derivace, minimalizace odchylek je podmíněna kriteriem vážených nejmenších čtverců. Při stanovení optimální vyhlazující funkce jsou tedy řešeny dvě minimalizační úlohy. Nejčastěji jsou v jednotlivých intervalech mezi uzlovými body použity kubické polynomy $S_3(x)$. Výsledná funkce je hladká ve funkčních hodnotách a v hodnotách první derivace, v hodnotách druhé derivace je ještě spojitá. Spline vyhlazování používáme obvykle v případě velkého počtu uzlových bodů, výsledná funkce má potom velmi složitý předpis, proto nejsou podstatné samotné koeficienty, ale křivka jako taková.

Spline regrese je vlastně specifickým případem vyhlazování při malém počtu uzlových bodů. Uzlové body $\xi_j$ jsou definovány mimo experimentální body. Výsledná funkce je potom lineární kombinací úsekových polynomů vymezených uzlovými body. Jedná se o regresní úlohu, kdy parametry lokálně definovaných funkcí jsou hledány metodou nejmenších čtverců. Hladkost aproximující funkce závisí na třídě $C^m$. Podmínku spojitosti ve funkčních hodnotách splňují polynomy třídy $m$+1. Pro třídu $C^0$ se jedná o lineární spline, pro třídu $C^1$ o kvadratický spline a v případě třídy $C^2$ kubický spline. Výsledkem spline regrese je tedy funkční předpis včetně koeficientů.

**Otázka 4:** Co můžeme říci o funkci 53H($x$)?

**Rešení:**

Filtrace obecně umožňuje vyhlazení v podobě eliminace šumové složky ve zpracovávaných signálech. V případě, kdy se v datech vyskytují i hrubé nenáhodné chyby (outliers) je třeba aplikovat robustní nelineární filtry. Jedním z nejúčinnějších je filtr 53H, kombinující tři mediány pátého stupně. Je dán výrazem

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

Pro zajištění lepšího vyhlazení lze mediány aplikovat opakovaně.