České Vysoké Učení Technické v Praze, Fakulta Elektrotechnická

Semestrální práce z předmětu Optimální rozhodování a řízení

Nejmenší čtverce a jiné normy

Michal Dvořák

24 - 05 - 2007

1 Úvod

Cílem této práce je porovnat metody aproximace funkcí založené na minimalizaci určitého kritéria. Nejrozšířenější je metoda nejmenších čtverců (LS - Least Squares), která je založena na minimalizaci normy H_2. Dále je popsána metoda nejmenších absolutních odchylek (LAD - Least Absolute Deviations) založená na minimalizaci H_1 normy a nakonec metoda minimalizující normu H_∞ známá pod jménem Čebyševova aproximace či minimax aproximace.

Základní úlohou aproximace je proložení naměřených dat vhodnou funkcí, která je bude dobře reprezentovat ve smyslu zvoleného kritéria. Na rozdíl od úlohy interpolace, zde není požadováno, aby hledaná funkce procházela přesně zadanými body. To odpovídá realitě, kdy naměřené hodnoty jsou zatíženy chybou měření a není důvod domnívat se, že tyto hodnoty jsou skutečně správné hodnoty hledané funkce.

1.1 Formulace úlohy aproximace

Je dán vektor n různých uzlových bodů $\mathbf{x} = (x_1, ..., x_n)^T \in \mathbb{R}^n$ a vektor hodnot $\mathbf{y} = (y_1, ..., y_n)^T \in \mathbb{R}^n$. Dále je dána třída funkcí $\mathcal{F}$, definovaných alespoň v bodech $x_0, ..., x_n$ ze které máme vybrat takovou funkci $\varphi(x) \in \mathcal{F}$, jejíž hodnoty se v bodech x_i "co nejméně liší" od y_i.

Uvažujeme případ, kdy $\mathcal{F}$ je lineární obal k vybraných (elementárních) funkcí. Pak hledaná funkce $\varphi(x)$ je lineární kombinací

$$\varphi(x) = \sum_{j=1}^{k} c_j \varphi_j(x), \tag{1}$$

kde $\mathbf{c} = (c_1, ..., c_k)^T \in \mathbb{R}^k$ jsou koeficienty, které je třeba určit.

2 Metoda nejmenších čtverců

Tato metoda aproximace je založena na minimalizaci součtu druhých mocnin odchylek hledané funkce od zadaných hodnot. Jinými slovy, hledáme koeficienty $c_1, ..., c_n$ lineární kombinace (1) takové, abychom minimalizovali druhou mocninu euklidovské normy

$$H_2^2 = \|\mathbf{r}\|_2^2 = \sum_{i=1}^{n} (\varphi(x_i) - y_i)^2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{k} c_j \varphi_j(x_i) - y_i \right)^2 . \tag{2}$$

Zavedením matice

$$\mathbf{A} = \begin{bmatrix} \varphi_1(x_1) & \dots & \varphi_k(x_1) \\ \vdots & \ddots & \vdots \\ \varphi_1(x_n) & \dots & \varphi_k(x_n) \end{bmatrix} \tag{3}$$

můžeme vektor odchylek (reziduí) zapsat také ve tvaru

$$\mathbf{r} = \mathbf{Ac} - \mathbf{y} \tag{4}$$

a celý problém hledání optimálních koeficientů vyjádřit jako

$$\mathbf{c}^* = \arg \min_{\mathbf{c} \in \mathbb{R}^k} \|\mathbf{r}\|_2^2 = \arg \min_{\mathbf{c} \in \mathbb{R}^k} \|\mathbf{Ac} - \mathbf{y}\|_2^2 . \tag{5}$$

2.1 Řešení přímou minimalizací normy

Normu, kterou minimalizujeme můžeme vyjádřit jako

$$\|\mathbf{r}\|_2^2 = \mathbf{r}^T \mathbf{r} = (\mathbf{Ac} - \mathbf{y})^T (\mathbf{Ac} - \mathbf{y}) . \tag{6}$$

Roznásobením tohoto výrazu získáme kvadratickou formu

$$\|\mathbf{r}\|_2^2 = \mathbf{c}^T \mathbf{A}^T \mathbf{Ac} - \mathbf{c}^T \mathbf{A}^T \mathbf{y} - \mathbf{y}^T \mathbf{Ac} + \mathbf{y}^T \mathbf{y}, \tag{7}$$

jejíž minimum leží v bodě, ve kterém je derivace kvadratické formy nulová.

$$\frac{\partial}{\partial \mathbf{c}} \|\mathbf{r}\|_2^2 = 2\mathbf{A}^T \mathbf{Ac} - 2\mathbf{A}^T \mathbf{y} = 0. \tag{8}$$

Za předpokladu, že sloupce matice $\mathbf{A}$ jsou lineárně nezávislé, existuje analytické řešením ve tvaru

$$\mathbf{c}^* = \left(\mathbf{A}^T \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{y}. \tag{9}$$

Výraz (9) se však pro numerický výpočet příliš nehodí a proto se používá rozkladu matice $\mathbf{A}$ metodou Choleskyho faktorizace, QR rozkladu či SVD rozkladu.

3 Metoda nejmenších absolutních odchylek

Na rozdíl od metody LS je tato metoda založena na minimalizaci normy

$$H_1 = \|\mathbf{r}\|_1 = \|\mathbf{Ac} - \mathbf{y}\|_1 = \sum_{i=1}^{n} |\varphi(x_i) - y_i|\,, \tag{10}$$

kdy hledáme optimální koeficinety

$$\mathbf{c}^* = \arg\min_{\mathbf{c}\in\mathbb{R}^k} \|\mathbf{Ac} - \mathbf{y}\|_1\,. \tag{11}$$

Minimalizaci normy H_1 lze řešit metodou lineárního programování následujícím způsobem

$$\min\left\{\mathbf{e}^T\mathbf{r} : \mathbf{Ac} - \mathbf{y} \le \mathbf{r}, \mathbf{Ac} - \mathbf{y} \ge -\mathbf{r}\right\}. \tag{12}$$

Zavedením vektoru $\mathbf{z} = \begin{bmatrix} \mathbf{c} \\ \mathbf{r} \end{bmatrix}$ a sloučením obou podmínek můžeme výraz (12) převést na standardní úlohu lineárního programování

$$\min\left\{\mathbf{f}^T\mathbf{z} : \widetilde{\mathbf{A}}\mathbf{z} \le \widetilde{\mathbf{y}}\right\}, \tag{13}$$

kde

$$\widetilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & -\mathbf{E} \\ -\mathbf{A} & -\mathbf{E} \end{bmatrix}, \quad \widetilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ -\mathbf{y} \end{bmatrix}, \quad \mathbf{f} = [0\ldots0, 1\ldots1]^T, \tag{14}$$

$\mathbf{E}$ je jednotková matice a $\mathbf{e}$ je jednotkový vektor.

4 Čebyševova aproximace

Tato metoda, také nazývaná minimax aproximace, je založena na minimalizaci normy

$$H_\infty = \|\mathbf{r}\|_\infty = \max\{|r_1|, \ldots, |r_n|\}\,. \tag{15}$$

Minimalizaci normy H_∞ lze opět řešit metodou lineárního programování následujícím způsobem

$$\min\{t : \mathbf{Ac} - \mathbf{y} \le t\mathbf{e}, \mathbf{Ac} - \mathbf{y} \ge -t\mathbf{e}\}\,, \tag{16}$$

kde $t = H_\infty$ je maximální chyba, kterou se snažíme minimalizovat. Zavedením vektoru $\mathbf{z} = \begin{bmatrix} \mathbf{c} \\ t \end{bmatrix}$ a sloučením obou podmínek můžeme výraz (16) převést na standardní úlohu lineárního programování

$$\min\left\{\mathbf{f}^T\mathbf{z} : \widetilde{\mathbf{A}}\mathbf{z} \le \widetilde{\mathbf{y}}\right\}, \tag{17}$$

kde

$$\widetilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & -\mathbf{e} \\ -\mathbf{A} & -\mathbf{e} \end{bmatrix}, \quad \widetilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ -\mathbf{y} \end{bmatrix}, \quad \mathbf{f} = [0\ldots0, 1]^T. \tag{18}$$

5 Porovnání metod na příkladu

Všechny uvedené metody byly porovnány na příkladu aproximace dat přímkou. Data byla vygenerována podle rovnice

$$y = 2x + 3 + e, \tag{19}$$

kde $e \sim \mathcal{N}(\mu, \sigma^2)$ je náhodná chyba s normálním rozdělením se zvolenými parametry $\mathcal{N}(0, 0.5)$.

Celkem bylo vygenerováno 21 vzorků na intervalu $\langle 0, 10 \rangle$ a několik hodnot bylo dodatečně vychýleno, aby se více projevil rozdíl mezi jednotlivými metodami, především odolnost aproximace oproti výrazně vychýleným hodnotám.

5.1 Řešení

Protože data byla vygenerována podle lineární závislosti, hledáme aproximaci ve tvaru

$$\varphi(x) = c_2 x + c_1, \tag{20}$$

která minimalizuje příslušné kritérium.

Nejprve najdeme aproximaci pomocí metody LS. Podle rovnice (3) vygenerujeme matici dat ve tvaru

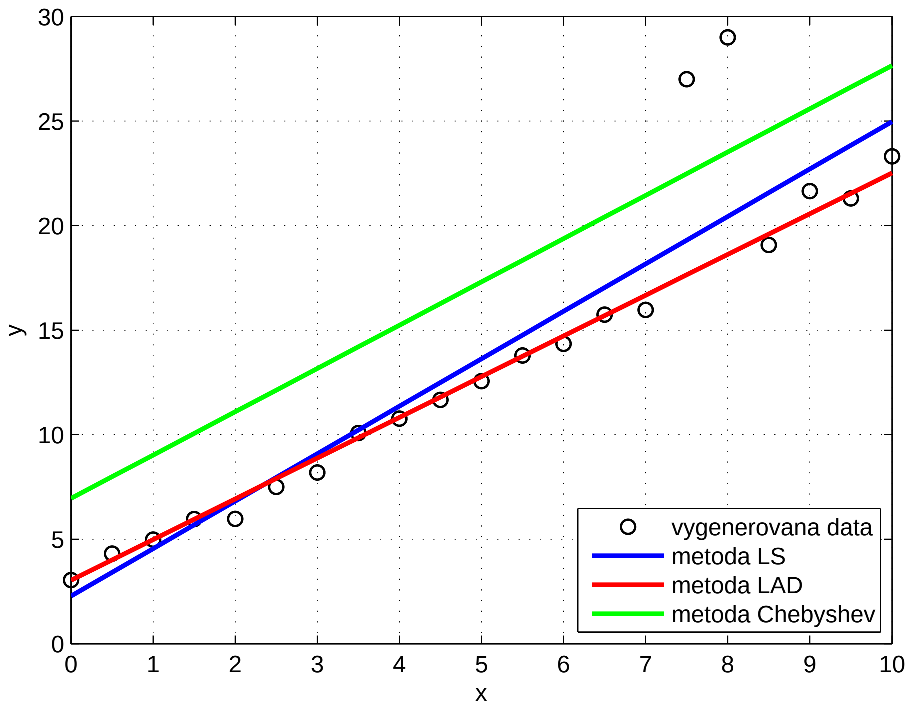
$$\mathbf{A} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \tag{21}$$

a dosazením do výrazu (9) získáme přímo koeficienty c_1, c_2.

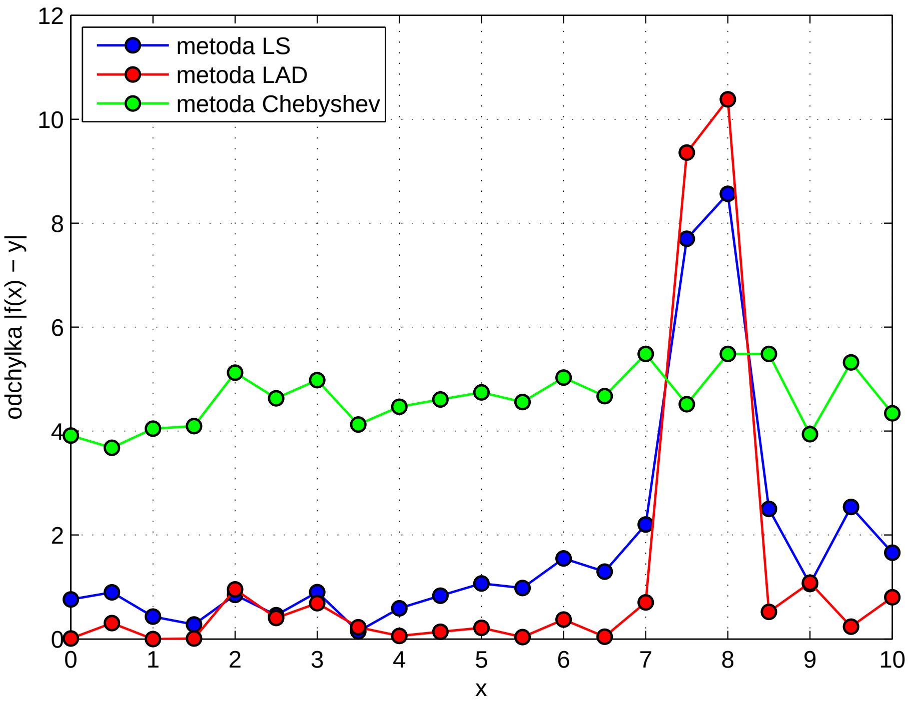
Při řešení metodou LAD vezmeme opět matici (21) a navíc sestavíme rozšířené matice podle předpisu (14). Úlohu lineárního programování (13) vyřešíme v programu MATLAB voláním funkce

$$\mathbf{z} = \texttt{linprog}\left(\mathbf{f}, \widetilde{\mathbf{A}}, \widetilde{\mathbf{y}}\right).$$

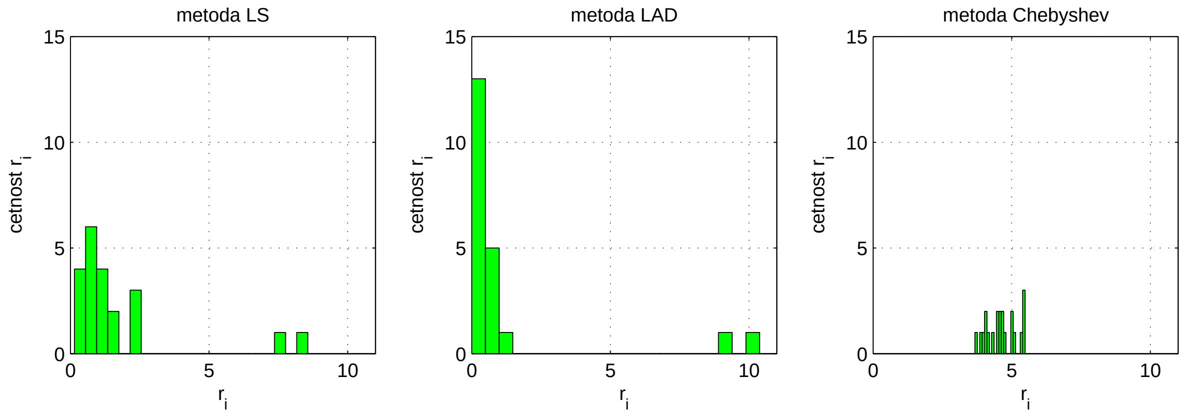
Stejným způsobem řešíme Čebyševovu aproximaci, pouze sestavíme příslušné rozšířené matice podle (18).



Obrázek 1: Porovnání jednotlivých metod při aproximaci lineární závislosti.



Obrázek 2: Absolutní odchylka aproximace v uzlových bodech.



Obrázek 3: Histogram absolutní chyby pro jednotlivé metody.

6 Závěr

Mají-li naměřené hodnoty **y** charakter součtu hodnot aproximované funkce $\varphi(x)$ a šumu s normálním rozložením, pak metoda nejmenších čtverců dává maximálně věrohodné řešení. Nicméně, je-li šum v datech jiného charakteru, jako ve výše uvedeném příkladu, má smysl uvažovat i jiné metody aproximace.

Na obr. 1 jsou průběhy jednotlivých aproximací. Metodou nejmenších absolutních odchylek (LAD) získáme aproximaci, která velmi dobře sleduje většinu dat. Její průběh se neodchýlí vlivem vychýlených hodnot jako u metody nejmenších čtverců. Metodu LAD tedy můžeme považovat za robustní metodu aproximace. Také z histogramu chyb na obr. 3 je patrné, že metoda LAD ma oproti metodě LS největší četnost malých chyb.

Pomocí Čebyševovy aproximace zase získáme průběh, který má nejmenší možnou maximální chybu. To je velmi dobře patrné jak z průběhu chyby na obr. 2 tak z histogramu, který je pro tuto metodu nejužší. Tato metoda se hodí v případě, kdy potřebujeme minimalizovat maximální chybu.

Reference

[1] J. Štecha: *Optimální rozhodování a řízní.* Skriptum ČVUT, FEL 1999.

[2] M. Navara: *Numerické metody.* Skriptum ČVUT, FEL 2003.

[3] S. Boyd, L. Vandenberghe: *Convex Optimization.* Cambridge University Press 2004.