Metoda nejmenších čtverců

- **Zadání:**

  Chceme aproximovat nějakou funkci, přičemž známe její diskrétní hodnoty $y_1, \ldots, y_n$ v bodech $x_1, \ldots, x_n$. Nepožadujeme ale, aby aproximace vycházela zadanými body, například proto, že hodnoty $y_1, \ldots, y_n$ nejsou zadány zcela přesně (jsou to třeba výsledky měření s určitou chybou). Místo toho máme představu o tom, jakou funkci k aproximaci použít. (Víme to například z teorie nebo z pohledu na graf zadaných bodů.) Hledáme tedy aproximaci funkce, která nemusí být její interpolací.

  Například máme čtyři body, o kterých víme, že by měly ležet na parabole, ale ve skutečnosti na ní úplně přesně neleží. Hledáme takovou parabolu, která data co nejlépe aproximuje.

- **Postup:**

  V mnoha případech nejlepší aproximaci $f(x)$ najdeme tak, že hledáme minimum funkce

  $$\tilde{S} = \sqrt{\sum_i \Big(y_i - f(x_i)\Big)^2}$$

  přes všechny podobné aproximace (např. polynomy stejného řádu).

  Co je to vlastně funkce $\tilde{S}$? Máme-li například pouze tři body x_1, x_2, x_3, pak (y_1, y_2, y_3) je bod Eukleidova prostoru (bod zadaných hodnot), $(f(x_1), f(x_2), f(x_3))$ je další bod Eukleidova prostoru (bod aproximace hodnot) a funkce $\tilde{S}$ je vzdálenost mezi těmito body (eukleidovská norma).

- **Příklad:**

  Chceme aproximovat data polynomem druhého stupně $\quad f(x) = a_0 + a_1\, x + a_2\, x^2$.

  - Hledáme tedy koeficienty a_0, a_1, a_2 takové, abychom minimalizovali

    $$\tilde{S} = \sqrt{\sum_i \big(y_i - a_0 - a_1\, x_i - a_2\, x_i^2\big)^2}.$$

    Odmocninu můžeme vynechat, protože odmocnina je monotónní funkce a tedy

    $$S = \tilde{S}^2 = \sum_i \big(y_i - a_0 - a_1\, x_i - a_2\, x_i^2\big)^2$$

    má minimum ve stejném bodu jako $\tilde{S}$.

  - Minimum budeme hledat pomocí derivace. V minimu platí

    $$\mathrm{grad}_{\vec{a}}\, S(a_0, a_1, a_2) = \vec{0}$$

    neboli

    $$\frac{\partial S}{\partial a_0} = 0, \qquad \frac{\partial S}{\partial a_1} = 0, \qquad \frac{\partial S}{\partial a_2} = 0.$$

  - Budeme tyto derivace postupně počítat. Nejprve derivujeme podle a_0

    $$\frac{\partial S}{\partial a_0} = \frac{\partial}{\partial a_0} \sum_i \big(\underbrace{y_i - a_1\, x_i - a_2\, x_i^2}_{\text{ozn. } k_i} - a_0\big)^2 = \frac{\partial}{\partial a_0} \sum_i (k_i - a_0)^2 = \sum_i \frac{\partial}{\partial a_0} (k_i - a_0)^2$$
    $$= \sum_i \Big(-2\,(k_i - a_0)\Big)\ ,$$

kde jsme označili

$$k_i = y_i - a_1\, x_i - a_2\, x_i^2.$$

Derivace podle a_1 je

$$\begin{aligned}
\frac{\partial S}{\partial a_1} &= \frac{\partial}{\partial a_1} \sum_i \big(\underbrace{y_i - a_0 - a_2\, x_i^2}_{\text{ozn. } l_i} - a_1\, x_i\big)^2 = \frac{\partial}{\partial a_1} \sum_i \left(l_i - a_1\, x_i\right)^2 = \sum_i \frac{\partial}{\partial a_1} \left(l_i - a_1\, x_i\right)^2 \\
&= \sum_i \Big(-2\, x_i \left(l_i - a_1\, x_i\right)\Big)\ ,
\end{aligned}$$

kde jsme označili

$$l_i = y_i - a_0 - a_2\, x_i^2.$$

Konečně derivace podle a_2 je

$$\begin{aligned}
\frac{\partial S}{\partial a_2} &= \frac{\partial}{\partial a_2} \sum_i \big(\underbrace{y_i - a_0 - a_1\, x_i}_{\text{ozn. } m_i} - a_2\, x_i^2\big)^2 = \frac{\partial}{\partial a_2} \sum_i \left(m_i - a_2\, x_i^2\right)^2 = \sum_i \frac{\partial}{\partial a_2} \left(m_i - a_2\, x_i^2\right)^2 \\
&= \sum_i \Big(-2\, x_i^2 \left(m_i - a_2\, x_i^2\right)\Big)\ ,
\end{aligned}$$

kde jsme označili

$$m_i = y_i - a_0 - a_1\, x_i.$$

– Derivace (podělené -2) položíme rovny nule

$$\begin{aligned}
0 &= \sum_i \left(k_i - a_0\right) = \sum_i \left(y_i - a_0 - a_1\, x_i - a_2\, x_i^2\right), \\
0 &= \sum_i x_i \left(l_i - a_1\, x_i\right) = \sum_i x_i \left(y_i - a_0 - a_1\, x_i - a_2\, x_i^2\right), \\
0 &= \sum_i x_i^2 \left(m_i - a_2\, x_i^2\right) = \sum_i x_i^2 \left(y_i - a_0 - a_1\, x_i - a_2\, x_i^2\right)
\end{aligned}$$

a rovnice přepíšeme jako

$$\begin{aligned}
\sum_i y_i &= a_0 \sum_i 1 + a_1 \sum_i x_i + a_2 \sum_i x_i^2, \\
\sum_i x_i\, y_i &= a_0 \sum_i x_i + a_1 \sum_i x_i^2 + a_2 \sum_i x_i^3, \\
\sum_i x_i^2\, y_i &= a_0 \sum_i x_i^2 + a_1 \sum_i x_i^3 + a_2 \sum_i x_i^4,
\end{aligned}$$

což je soustava lineárních rovnic pro koeficienty a_0, a_1, a_2 hledané aproximace. Maticově ji můžeme zapsat jako

$$\begin{pmatrix} \sum 1 & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i\, y_i \\ \sum x_i^2\, y_i \end{pmatrix},$$

kde pro N aproximovaných bodů jdou sumy od 1 do N a $\sum 1 = N$.