Kapitola 10

Fitování aneb metoda nejmenších čtverců

Metoda nejmenších čtverců (least square method) je založena na minimalizaci sumy druhých mocnin odchylek zadaných a upřesněných hodnot,

$$S^2(\vec{p}) = \sum_{i=1}^{N} (y_i - f(x_i, \vec{p}))^2 w_i \tag{10.1}$$

kde $\vec{p}$ je vektor hledaných parametrů funkce $f(x_i, \vec{p})$, x_i a y_i jsou vektory souřadnic bodů, kterými je prokládána závislost daná předpisem $f(x_i, \vec{p})$. Bodů $[x_i, y_i]$ je celkem N, počet parametrů budeme označovat písmenem M. w_i zde značí váhy jednotlivých bodů $[x_i, y_i]$. Minimum hodnoty sumy čtverců $S^2(\vec{p})$ je dáno nulovou hodnotou gradientu $S^2(\vec{p})$ dle $\vec{p}$ (analogie s první derivací funkce jedné proměnné v minimu). Tuto podmínku můžeme zapsat pomocí soustavy rovnic

$$\frac{\partial S^2}{\partial p_j} = -2 \sum_{i=1}^{N} (y_i - f(x_i, \vec{p})) \frac{\partial f(x_i, \vec{p})}{\partial p_j} w_i. \tag{10.2}$$

Soustavu rovnic 10.2 můžeme snadno řešit např. v případě lineárního nebo kvadratického fitu, kdy obdržíme námé vztahy (v případě lineárního fitu $f(x) = p_1 x + p_0$ - viz strana 72)

$$\sum_{i=1}^{N} y_i = p_0 N + p_1 \sum_{i=1}^{N} x_i \tag{10.3}$$

$$\sum_{i=1}^{N} y_i x_i = p_0 \sum_{i=1}^{N} x_i + p_1 \sum_{i=1}^{N} x_i^2 \tag{10.4}$$

a pro kvadratický fit ($f(x) = p_2x^2 + p_1x + p_0$ - viz strana 75)

$$\sum_{i=1}^{N} y_i = p_0N + p_1\sum_{i=1}^{N} x_i + p_2\sum_{i=1}^{N} x_i^2 \quad (10.5)$$

$$\sum_{i=1}^{N} y_ix_i = p_0\sum_{i=1}^{N} x_i + p_1\sum_{i=1}^{N} x_i^2 + p_2\sum_{i=1}^{N} x_i^3 \quad (10.6)$$

$$\sum_{i=1}^{N} y_ix_i^2 = p_0\sum_{i=1}^{N} x_i^2 + p_1\sum_{i=1}^{N} x_i^3 + p_2\sum_{i=1}^{N} x_i^4. \quad (10.7)$$

Obecně se však jedná o soustavu rovnic, kterou lze řešit pouze iteračně. To znamená, že z výchozího odhadu řešení se snažíme najít skutečné řešení minimalizující 10.1. Postup je následující: nahradíme v předpisu minima 10.2 hodnotu funkce $f(x_i, \vec{p})$ a její derivace pomocí Taylorova rozvoje v okolí odhadu řešení $\vec{p_0}$ (omezíme se na první členy rozvoje):

$$f(x_i, \vec{p}) = f(x_i, \vec{p_0}) + \sum_{m=1}^{M} \frac{\partial f(x_i, \vec{p_0})}{\partial p_m}\Delta p_m$$

$$\frac{\partial f(x_i, \vec{p})}{\partial p_k} = \frac{\partial f(x_i, \vec{p_0})}{\partial p_k} + \sum_{m=1}^{M} \frac{\partial^2 f(x_i, \vec{p_0})}{\partial p_k \partial p_m}\Delta p_m. \quad (10.8)$$

Nyní dosadíme rovnice 10.8 do 10.2 a zanedbáme kvadratické členy v Δp_m. To můžeme provést neboť jsou to členy druhého řádu a pokud nejsme s první aproximací řešení $\vec{p_0}$ příliš daleko od minima, je jejich velikost malá. Po dosazení a nezbytných algebraických úpravách dostaneme:

$$\sum_{m=1}^{M}\left(\sum_{i=1}^{N}\left[\frac{\partial f(x_i, \vec{p_0})}{\partial p_k}\frac{\partial f(x_i, \vec{p_0})}{\partial p_m}w_i + (y_i - f(x_i, \vec{p_0}))\frac{\partial^2 f(x_i, \vec{p_0})}{\partial p_k \partial p_m}w_i\right]\right)\Delta p_m = \sum_{i=1}^{N}(y_i - f(x_i, \vec{p_0}))\frac{\partial f(x_i, \vec{p_0})}{\partial p_k}w_i. \quad (10.9)$$

Pokud nejsme příliš daleko od skutečného minima sumy čtverců 10.1, můžeme zanedbat člen s druhou derivací funkce $f(x_i, \vec{p_0})$. Tento člen se při numerickém hledání minima 10.1 běžně zanedbává nejen z tohoto důvodu – výpočet druhé derivace není u složitějších úloh bez komplikací. Takže nakonec máme soustavu M rovnic pro složky vektoru změn parametrů $\Delta\vec{p}$ vektoru odhadu řešení $\vec{p_0}$. Složky vektoru $\Delta\vec{p}$ nám udávají, jak máme změnit původní parametry $\vec{p_0}$, abychom se více přiblížili minimu 10.1.

$$\sum_{m=1}^{M}\left(\sum_{i=1}^{N}\left[\frac{\partial f(x_i, \vec{p_0})}{\partial p_k}\frac{\partial f(x_i, \vec{p_0})}{\partial p_m}w_i\right]\right)\Delta p_m = \sum_{i=1}^{N}(y_i - f(x_i, \vec{p_0}))\frac{\partial f(x_i, \vec{p_0})}{\partial p_k}w_i. \quad (10.10)$$

Nový vektor řešení je dán jednoduchým vztahem

$$\vec{p} = \vec{p}_0 + \Delta\vec{p}. \tag{10.11}$$

10.1 Gaussova metoda

Soustava rovnic 10.10 je známa jako Gaussova metoda minimalizace sumy čtverců 10.1. Pokud se podíváme na rovnice 10.10 pozorněji, můžeme je snadno přepsat do maticového tvaru:

$$\hat{J} * (\hat{J} * \hat{w})' \Delta\vec{p} = \hat{J} * \hat{w} * \hat{y}_{res}, \tag{10.12}$$

kde $\hat{J}$ je matice Jacobiánu (tj. parciálních derivací modelové funkce podle hledaných parametrů) definovaná předpisem

$$\hat{J} = \begin{pmatrix} \frac{\partial f(x_1,\vec{p})}{\partial p_1} & \frac{\partial f(x_2,\vec{p})}{\partial p_1} & \dots & \frac{\partial f(x_N,\vec{p})}{\partial p_1} \\ \frac{\partial f(x_1,\vec{p})}{\partial p_2} & \frac{\partial f(x_2,\vec{p})}{\partial p_2} & \dots & \frac{\partial f(x_N,\vec{p})}{\partial p_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f(x_1,\vec{p})}{\partial p_M} & \frac{\partial f(x_2,\vec{p})}{\partial p_M} & \dots & \frac{\partial f(x_N,\vec{p})}{\partial p_M} \end{pmatrix}. \tag{10.13}$$

$\hat{w}$ je diagonální matice váhových koeficientů jednotlivých bodů $[x_i, y_i]$,

$$\hat{w} = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & w_N \end{pmatrix}, \tag{10.14}$$

$\hat{y}_{res}$ je vektor odchylek zadaných a zpřesněných hodnot y,

$$\hat{y}_{res} = \begin{pmatrix} y_1 - f(x_1, \vec{p}) \\ y_2 - f(x_2, \vec{p}) \\ \vdots \\ y_N - f(x_N, \vec{p}) \end{pmatrix} \tag{10.15}$$

a $\Delta\vec{p}$ je vektor hledaných změn parametrů $\vec{p}$. Apostrof ' značí transpozici matice. Soustavě 10.10 se často říká soustava *normálních* rovnic. Do soustavy 10.10 nyní dosadíme nový vektor řešení 10.11 a opakujeme postup, dokud rozdíl součtu čtverců nového a předchozího řešení nedosáhne zadané hodnoty.

Směrodatné odchylky nalezených parametrů určíme pomocí matice na levé straně 10.12,

$$\hat{\sigma} = \sqrt{(\hat{J} * (\hat{J} * \hat{w})')^{-1}} s, \tag{10.16}$$

kde s je směrodatná odchylka hodnot y_i,

$$s = \sqrt{\frac{S^2(\vec{p})}{\nu}}, \tag{10.17}$$

kde $S^2(\vec{p})$ je suma čtverců definovaná vztahem 10.1 a $\nu = N - M$ je počet stupňů volnosti. Matici

$$\hat{C} = (\hat{J} * (\hat{J} * \hat{w})')^{-1}, \tag{10.18}$$

se říká *kovarianční.*

Ukažme si použití výše odvozených vztahů na jednoduchém příkladě lineárního fitu, tj. fitováním polynomem prvního stupně. Mějme 5 dvojic bodů $[x_i, y_i]$, které chceme proložit („fitovat“) lineární funkcí $f(x) = ax + b$ s počátečním odhadem řešení a=2 a b=1. Body mají následující hodnoty, váhové faktory pokládáme rovné 1:

1	3.131
2	5.001
3	7.149
4	9.171
5	11.028

Výpočet zastavíme, až bude rozdíl sum čtverců dvou po sobě následujících iterací menší než 10^{-6}.

Spočtěme sumu čtverců 10.1 pro odhad řešení $\vec{p}$=[2,1]:

$$S^2(\vec{p}) = \sum_{i=1}^{N} (y_i - f(x_i, \vec{p}))^2 w_i = 0.069388. \tag{10.19}$$

Matice Jacobiánu $\hat{J}$ 10.13 má jednoduchý tvar ($\vec{p} = [a, b]$ je vektor parametrů)

$$\hat{J} = \begin{pmatrix} \frac{\partial f(x_1,\vec{p})}{p_1} & \frac{\partial f(x_2,\vec{p})}{p_1} & \frac{\partial f(x_3,\vec{p})}{p_1} & \frac{\partial f(x_4,\vec{p})}{p_1} & \frac{\partial f(x_5,\vec{p})}{p_1} \\ \frac{\partial f(x_1,\vec{p})}{p_0} & \frac{\partial f(x_2,\vec{p})}{p_0} & \frac{\partial f(x_3,\vec{p})}{p_0} & \frac{\partial f(x_4,\vec{p})}{p_0} & \frac{\partial f(x_5,\vec{p})}{p_0} \end{pmatrix} = \\ = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{10.20}$$

Vektor $\hat{y}_{res}$ 10.15 má pro odhad řešení tyto prvky:

$$\hat{y}_{res} = \begin{pmatrix} 0.1310000 \\ 0.0010000 \\ 0.1490000 \\ 0.1710000 \\ 0.0280000 \end{pmatrix}. \tag{10.21}$$

Levá strana 10.12 je matice 2×2 (máme dvě rovnice pro dvě změny $\Delta\vec{p}$ vektoru řešení),

$$\hat{J} * (\hat{J} * \hat{w})\Delta\vec{p} = \begin{pmatrix} 55 & 15 \\ 15 & 5 \end{pmatrix} \Delta\vec{p} \tag{10.22}$$

Podobně, pro pravou stranu 10.12 po dosazení dostaneme

$$\hat{J} * \hat{w} * \hat{y}_{res} = \begin{pmatrix} 1.40400 \\ 0.48000 \end{pmatrix}. \tag{10.23}$$

Hledané změny parametrů $\Delta\vec{p}$ jsou dány řešením soustavy rovnic

$$\begin{pmatrix} 55 & 15 \\ 15 & 5 \end{pmatrix} \Delta\vec{p} = \begin{pmatrix} 1.40400 \\ 0.48000 \end{pmatrix} \rightarrow \Delta\vec{p} = \begin{pmatrix} -0.0036000 \\ 0.1068000 \end{pmatrix}. \tag{10.24}$$

Nový vektor řešení má tedy tvar (původní měl hodnoty [2 1])

$$\vec{p} = \vec{p} + \Delta\vec{p} = \begin{pmatrix} 1.9964 \\ 1.1068 \end{pmatrix}. \tag{10.25}$$

Suma čtverců pro nové řešení je $S^2 = 0.023178$, rozdíl sum čtverců nového a starého řešení, $\Delta S^2 = 0.046210$, což je více než zadaná přesnost. Výše uvedeným způsobem spočteme nový vektor změn $\Delta\vec{p}$. Pro kontrolu uvádíme hodnoty vektorů a matic (mimo Jacobiánu $\hat{J}$, který v našem případě zůstává stejný neboť nezávisí na parametrech $\vec{p}$).

$$\hat{y}_{res} = \begin{pmatrix} 0.027800 \\ -0.098600 \\ 0.053000 \\ 0.078600 \\ -0.060800 \end{pmatrix}, \tag{10.26}$$

$$\begin{pmatrix} 55 & 15 \\ 15 & 5 \end{pmatrix} \Delta\vec{p} = \begin{pmatrix} 7.1054 \cdot 10^{-15} \\ 1.7764 \cdot 10^{-15} \end{pmatrix} \rightarrow \Delta\vec{p} = \begin{pmatrix} 1.7764 \cdot 10^{-16} \\ -1.7764 \cdot 10^{-16} \end{pmatrix}. \tag{10.27}$$

Vidíme, že změny parametrů v této iteraci jsou velice malé, či–li nejspíše jsme dosáhli minima 10.1. Suma čtverců pro nové řešení je $S^2 = 0.023178$, rozdíl sum čtverců nového a starého řešení, $\Delta S^2 = 2.1511 \cdot 10^{-16}$, což je méně než zadaná přesnost a řešení bylo nalezeno. Zbývá spočítat směrodatné odchylky nalezených parametrů. Použijeme vztah 10.16,

$$\hat{\sigma} = \sqrt{(\hat{J} * (\hat{J} * \hat{w})')^{-1}} s = \begin{pmatrix} 0.027796 \\ 0.092189 \end{pmatrix}, \tag{10.28}$$

tj. hledané parametry funkce a a b jsou následující (počet platných míst je nadhodnocen):

$$a = 1.996400 \pm 0.027796 \tag{10.29}$$

$$b = 1.106800 \pm 0.092189 \tag{10.30}$$

Graficky je lineární fit zachycen na obrázku 10.1.

Podobně můžeme řešit, „fitovat“, složitější funkce, např. $f(x) = ae^{-bx}$. V tomto případě je Jacobián o něco složitější, postup však zůstává stejný.

$$\begin{aligned} \hat{J} &= \begin{pmatrix} \frac{\partial f(x_1,\vec{p})}{\partial p_1} & \frac{\partial f(x_2,\vec{p})}{\partial p_1} & \dots & \frac{\partial f(x_N,\vec{p})}{\partial p_1} \\ \frac{\partial f(x_1,\vec{p})}{\partial p_2} & \frac{\partial f(x_2,\vec{p})}{\partial p_2} & \dots & \frac{\partial f(x_N,\vec{p})}{\partial p_2} \end{pmatrix} \\ &= \begin{pmatrix} e^{-bx_1} & e^{-bx_2} & \dots & e^{-bx_N} \\ -axe^{-bx_1} & -axe^{-bx_2} & \dots & -axe^{-bx_N} \end{pmatrix}. \end{aligned} \tag{10.31}$$