Metoda nejmenších čtverců.

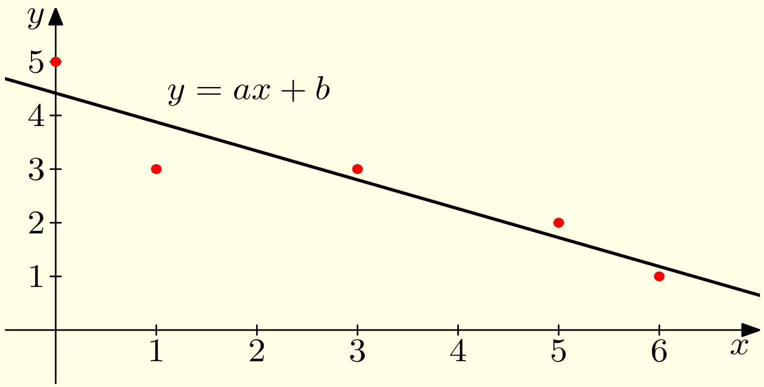
Robert Mařík

22. ledna 2006

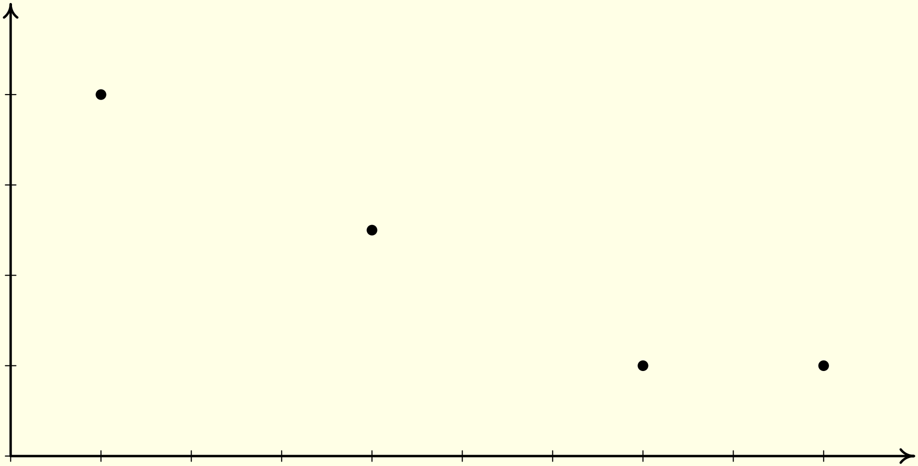
Obsah

1 Motivace a geometrický význam

Předpokládejme, že teorie ukazuje na skutečnost, že mezi veličinami x a y je lineární vztah ve tvaru $y = ax + b$. Měřením byly pro konkrétní hodnoty veličiny x naměřeny odpovídající hodnoty veličiny y a výsledek byl zanesen do grafu. Body které jsme obdrželi však neleží na jedné přímce, protože měření je vždy zatíženo nějakou chybou a naše teorie navíc vždy nemusí odpovídat praxi stoprocentně. Máme tedy body v rovině, které leží přibližně v jedné přímce a chceme najít co nejpřesnější matematický model, tj. stanovit koeficienty a, b tak, aby přímka $y = ax + b$ ležela co nejblíže bodům z měření.



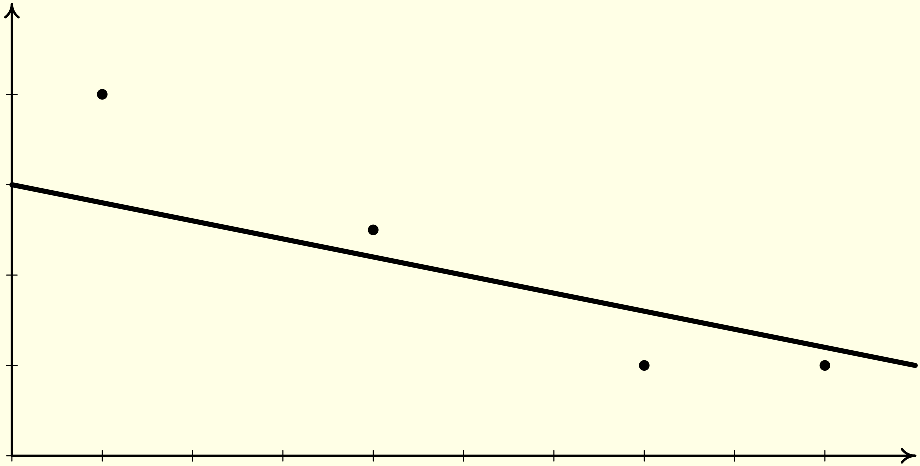
Metoda nejmenších čtverců



Soubor bodů

x	1	4	7	9
y	4	2.5	1	1

Metoda nejmenších čtverců

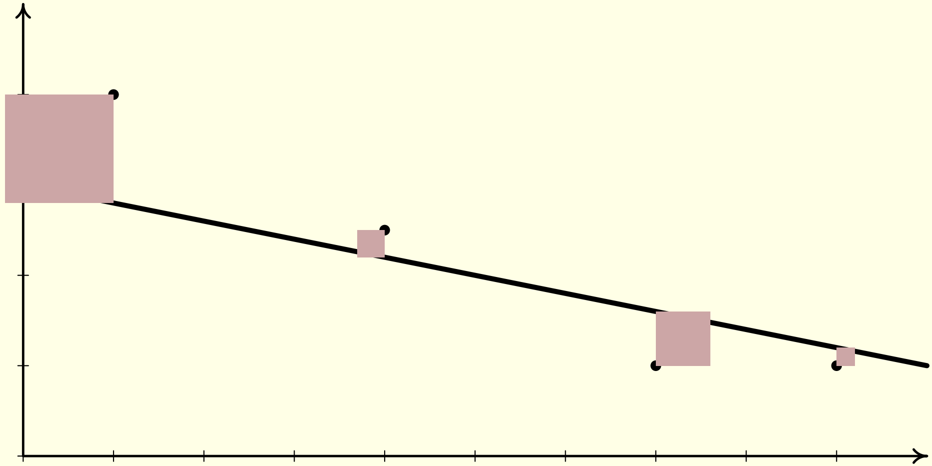


Soubor bodů

x	1	4	7	9
y	4	2.5	1	1

Snažíme se vystihnout chování bodů pomocí lineární závislosti. Přímka nebude pochopitelně procházet všemi body, chceme tedy alespoň, aby procházela co nejblíže okolo nich.

Metoda nejmenších čtverců

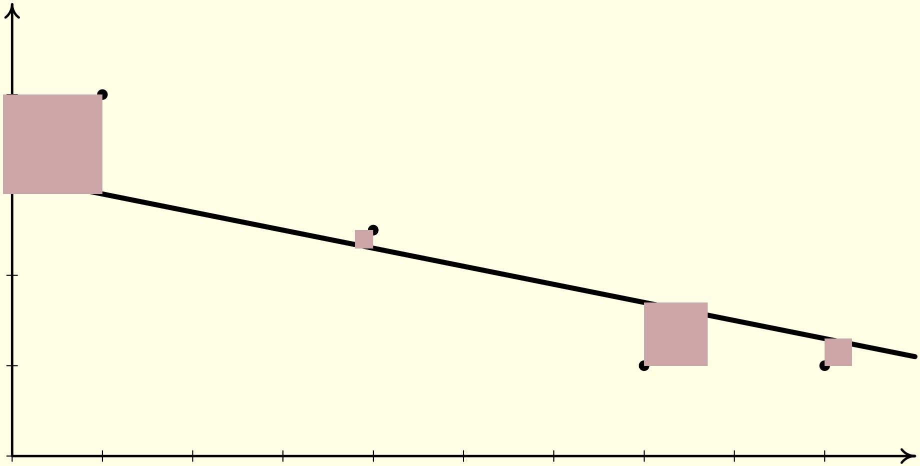


Soubor bodů

x	1	4	7	9
y	4	2.5	1	1

Snažíme se vystihnout chování bodů pomocí lineární závislosti. Přímka nebude pochopitelně procházet všemi body, chceme tedy alespoň, aby procházela co nejblíže okolo nich.
Za optimální přímku považujeme tu, která minimalizuje součet ploch červených čtverců.

Metoda nejmenších čtverců

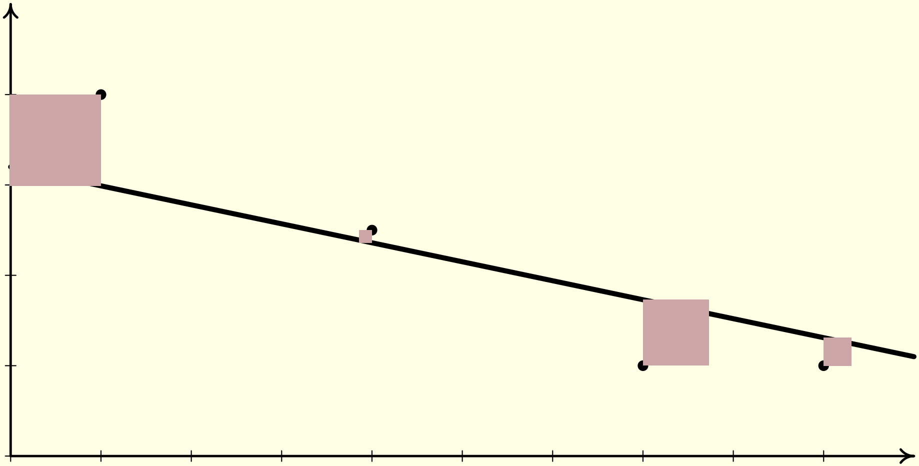


Soubor bodů

x	1	4	7	9
y	4	2.5	1	1

Snažíme se vystihnout chování bodů pomocí lineární závislosti. Přímka nebude pochopitelně procházet všemi body, chceme tedy alespoň, aby procházela co nejblíže okolo nich.
Za optimální přímku považujeme tu, která minimalizuje součet ploch červených čtverců.

Metoda nejmenších čtverců

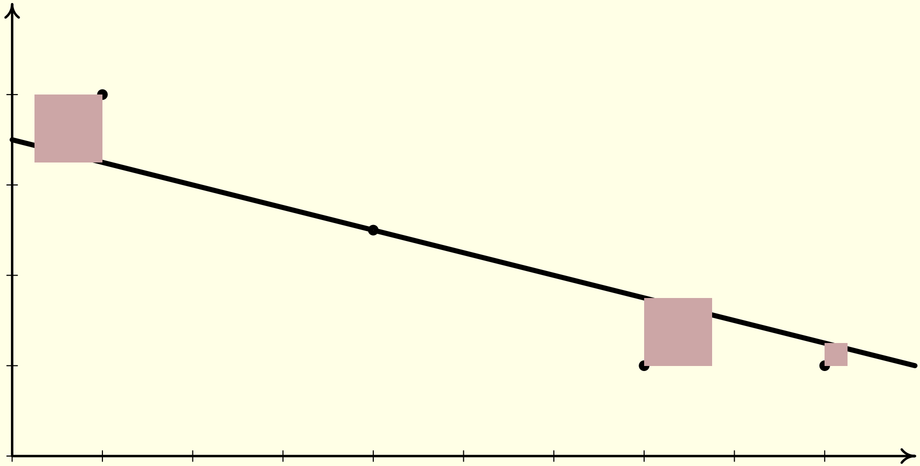


Soubor bodů

x	1	4	7	9
y	4	2.5	1	1

Snažíme se vystihnout chování bodů pomocí lineární závislosti. Přímka nebude pochopitelně procházet všemi body, chceme tedy alespoň, aby procházela co nejblíže okolo nich.
Za optimální přímku považujeme tu, která minimalizuje součet ploch červených čtverců.

Metoda nejmenších čtverců

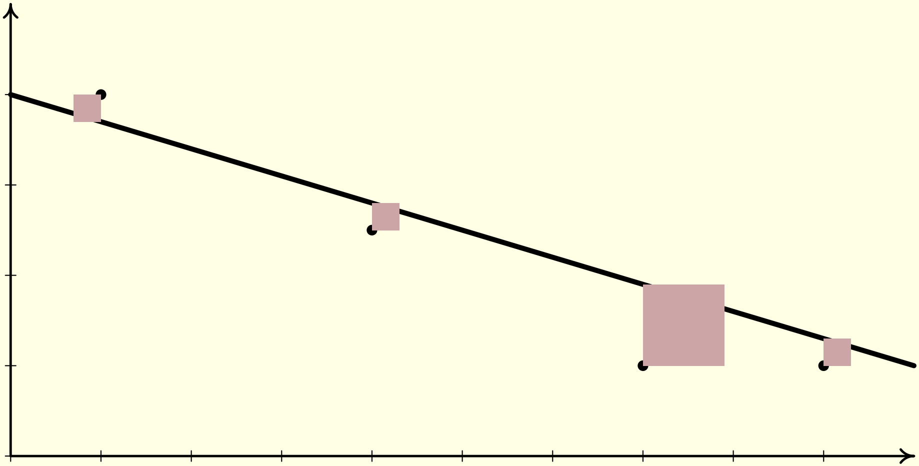


Soubor bodů

x	1	4	7	9
y	4	2.5	1	1

Snažíme se vystihnout chování bodů pomocí lineární závislosti. Přímka nebude pochopitelně procházet všemi body, chceme tedy alespoň, aby procházela co nejblíže okolo nich.
Za optimální přímku považujeme tu, která minimalizuje součet ploch červených čtverců.

Metoda nejmenších čtverců

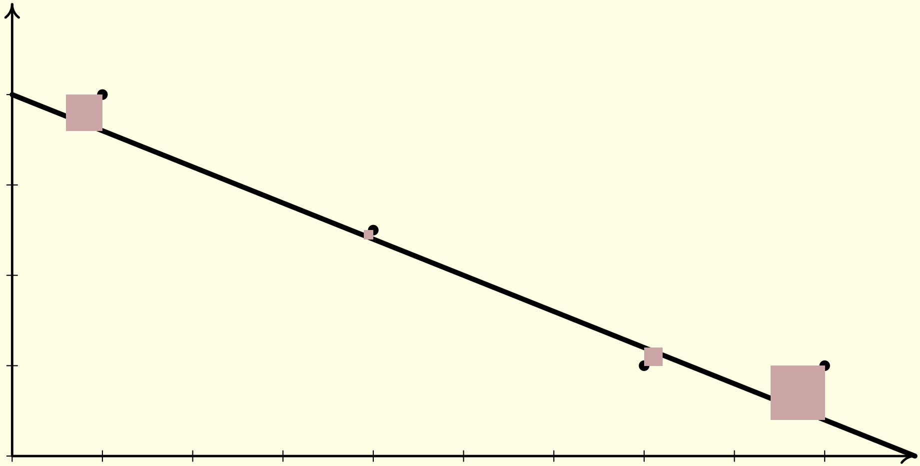


Soubor bodů

x	1	4	7	9
y	4	2.5	1	1

Snažíme se vystihnout chování bodů pomocí lineární závislosti. Přímka nebude pochopitelně procházet všemi body, chceme tedy alespoň, aby procházela co nejblíže okolo nich.
Za optimální přímku považujeme tu, která minimalizuje součet ploch červených čtverců.

Metoda nejmenších čtverců

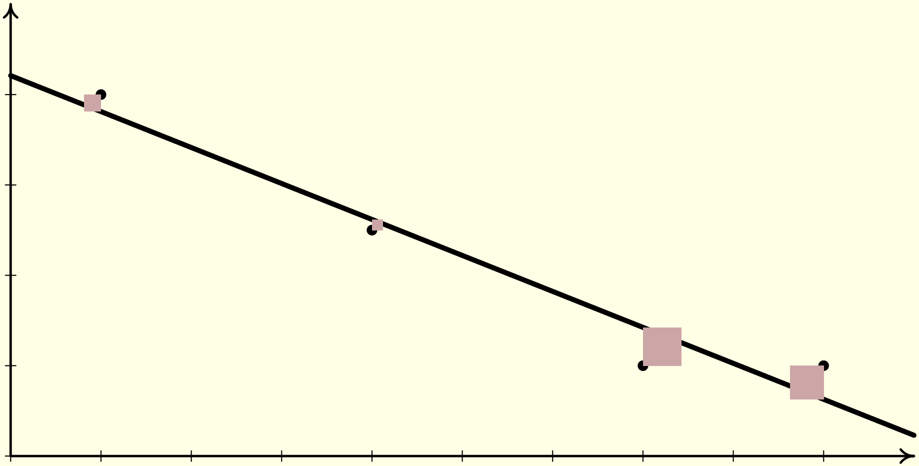


Soubor bodů

x	1	4	7	9
y	4	2.5	1	1

Snažíme se vystihnout chování bodů pomocí lineární závislosti. Přímka nebude pochopitelně procházet všemi body, chceme tedy alespoň, aby procházela co nejblíže okolo nich.
Za optimální přímku považujeme tu, která minimalizuje součet ploch červených čtverců.

Metoda nejmenších čtverců



Soubor bodů

x	1	4	7	9
y	4	2.5	1	1

Snažíme se vystihnout chování bodů pomocí lineární závislosti. Přímka nebude pochopitelně procházet všemi body, chceme tedy alespoň, aby procházela co nejblíže okolo nich.

Za optimální přímku považujeme tu, která minimalizuje součet ploch červených čtverců.

Tahle přímka je optimální. Jak ji najít?

2 Vzorec

Přímku proloženou metodou nejmenších čtverců hledáme za použití následující věty. K jejímu odvození je možné využít parciální derivace, jak je ukázáno ve dvou posledních kapitolách tohoto souboru, na straně 21. Studenti kteří jsou si jistí, že parciální derivace nemají v náplni svého kurzu (například LDF, 1. ročník), by toto odvození číst neměli.

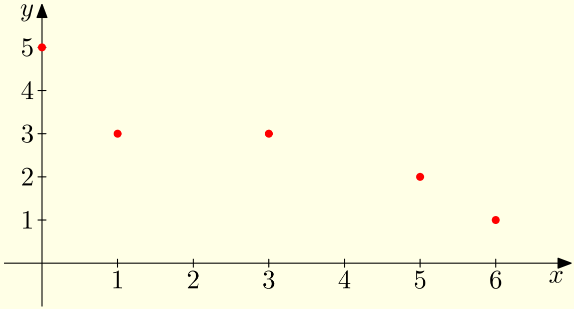
Věta [prokládání souboru bodů přímkou]. Přímka $y = ax + b$ je přímka, proložená metodou nejmenších čtverců souborem bodů $[x_1, y_1]$, $[x_2, y_2]$, ..., $[x_n, y_n]$, jestliže pro koeficienty a, b platí

$$\begin{aligned} a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i &= \sum_{i=1}^{n} x_i y_i, \\ a\sum_{i=1}^{n} x_i + b\, n &= \sum_{i=1}^{n} y_i. \end{aligned} \tag{1}$$

3 Příklad použití

Problém: Proložte přímku následujícím souborem bodů.

x_i	0	1	3	5	6
y_i	5	3	3	2	1



Řešení: Body v souboru jsou $[0,5]$, $[1,3]$, $[3,3]$, $[5,2]$ a $[6,1]$. Celkem tedy máme pět bodů, tj. $n=5$. Výpočty potřebné pro nalezení koeficientů v soustavě provedeme v následující tabulce.

i	x_i	y_i		
1	0	5		
2	1	3		
3	3	3		
4	5	2		
5	6	1		
$\sum$				

$$\begin{aligned} a\sum x_i^2 \quad + \quad b\sum x_i \quad &= \quad \sum x_i y_i \\ a\sum x_i \quad + \quad b\, n \quad &= \quad \sum y_i \end{aligned}$$

i	x_i	y_i	x_i^2	
1	0	5	0	
2	1	3	1	
3	3	3	9	
4	5	2	25	
5	6	1	36	
$\sum$				

$$\begin{aligned} a\sum x_i^2 + b\sum x_i &= \sum x_i y_i \\ a\sum x_i + b\,n &= \sum y_i \end{aligned}$$

i	x_i	y_i	x_i^2	x_iy_i
1	0	5	0	0
2	1	3	1	3
3	3	3	9	9
4	5	2	25	10
5	6	1	36	6
$\sum$				

$$\begin{aligned} a\sum x_i^2 &+ b\sum x_i &= \sum x_iy_i \\ a\sum x_i &+ b\,n &= \sum y_i \end{aligned}$$

i	x_i	y_i	x_i^2	$x_i y_i$
1	0	5	0	0
2	1	3	1	3
3	3	3	9	9
4	5	2	25	10
5	6	1	36	6
$\sum$	15	14	71	28

$$\begin{aligned} a\sum x_i^2 + b\sum x_i &= \sum x_i y_i \\ a\sum x_i + b\, n &= \sum y_i \end{aligned}$$

i	x_i	y_i	x_i^2	x_iy_i
1	0	5	0	0
2	1	3	1	3
3	3	3	9	9
4	5	2	25	10
5	6	1	36	6
$\sum$	15	14	71	28

Sestavíme soustavu lineárních rovnic

$$\begin{aligned} 71a + 15b &= 28, \\ 15a + 5b &= 14. \end{aligned}$$

$$\begin{aligned} a\sum x_i^2 \quad &+ \quad b\sum x_i \quad &= \quad \sum x_iy_i \\ a\sum x_i \quad &+ \quad b\ n \quad &= \quad \sum y_i \end{aligned}$$

i	x_i	y_i	x_i^2	x_iy_i
1	0	5	0	0
2	1	3	1	3
3	3	3	9	9
4	5	2	25	10
5	6	1	36	6
$\sum$	15	14	71	28

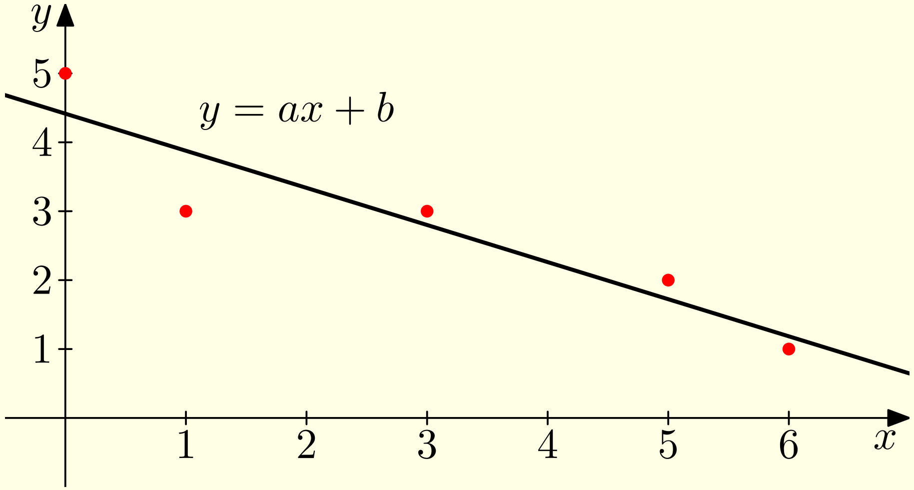
Sestavíme soustavu lineárních rovnic

$$\begin{aligned} 71a + 15b &= 28, \\ 15a + 5b &= 14. \end{aligned}$$

Řešením této soustavy je $a = -\dfrac{5}{13} \doteq -0.538$ a $b = \dfrac{387}{65} \doteq 4.415$. Nejlepší lineární aproximace souboru bodů je tedy přímka

$$\boxed{y = -0.538x + 4.415}.$$

Graf souboru bodů a výsledná přímka jsou zachyceny na obrázku.



Obrázek 1: Metoda nejmenších čtverců

4 Odvození vzorce

Uvažujme tři body $[x_1, y_1]$, $[x_2, y_2]$ a $[x_3, y_3]$. Svislé vzdálenosti těchto bodů od přímky $y = ax + b$ jsou $s_1 = |ax_1 + b - y_1|$, $s_2 = |ax_2 + b - y_2|$, $s_3 = |ax_3 + b - y_3|$ a chceme minimalizovat funkci

$$S(a, b) = (ax_1 + b - y_1)^2 + (ax_2 + b - y_2)^2 + (ax_3 + b - y_3)^2.$$

Lokální extrém diferencovatelné funkce může nastat pouze ve stacionárním bodě. Platí

$$\begin{aligned}\frac{\partial S}{\partial a} &= 2(ax_1 + b - y_1)x_1 + 2(ax_2 + b - y_2)x_2 + 2(ax_3 + b - y_3)x_3 \\ &= 2a(x_1^2 + x_2^2 + x_3^2) + 2b(x_1 + x_2 + x_3) - 2(x_1y_1 + x_2y_2 + x_3y_3)\end{aligned}$$

a

$$\begin{aligned}\frac{\partial S}{\partial b} &= 2(ax_1 + b - y_1) + 2(ax_2 + b - y_2) + 2(ax_3 + b - y_3) \\ &= 2a(x_1 + x_2 + x_3) + 2 \cdot 3b - 2(y_1 + y_2 + y_3).\end{aligned}$$

Položíme-li derivace rovny nule, dostáváme po vydělení dvěma a přesunu členů bez neznámých na pravou stranu rovnice

$$\begin{aligned} a(x_1^2 + x_2^2 + x_3^2) &+ b(x_1 + x_2 + x_3) &&= x_1y_1 + x_2y_2 + x_3y_3, \\ a(x_1 + x_2 + x_3) &+ b \cdot 3 &&= y_1 + y_2 + y_3. \end{aligned}$$

4 Odvození vzorce

Uvažujme tři body $[x_1, y_1]$, $[x_2, y_2]$ a $[x_3, y_3]$. Svislé vzdálenosti těchto bodů od přímky $y = ax + b$ jsou $s_1 = |ax_1 + b - y_1|$, $s_2 = |ax_2 + b - y_2|$, $s_3 = |ax_3 + b - y_3|$ a chceme minimalizovat funkci

$$S(a, b) = (ax_1 + b - y_1)^2 + (ax_2 + b - y_2)^2 + (ax_3 + b - y_3)^2.$$

Lokální extrém diferencovatelné funkce může nastat pouze ve stacionárním bodě. Platí

$$\begin{aligned}\frac{\partial S}{\partial a} &= 2(ax_1 + b - y_1)x_1 + 2(ax_2 + b - y_2)x_2 + 2(ax_3 + b - y_3)x_3 \\ &= 2a(x_1^2 + x_2^2 + x_3^2) + 2b(x_1 + x_2 + x_3) - 2(x_1y_1 + x_2y_2 + x_3y_3)\end{aligned}$$

a

$$\begin{aligned}\frac{\partial S}{\partial b} &= 2(ax_1 + b - y_1) + 2(ax_2 + b - y_2) + 2(ax_3 + b - y_3) \\ &= 2a(x_1 + x_2 + x_3) + 2 \cdot 3b - 2(y_1 + y_2 + y_3).\end{aligned}$$

Položíme-li derivace rovny nule, dostáváme po vydělení dvěma a přesunu členů bez neznámých na pravou stranu rovnice

$$\begin{aligned} a(x_1^2 + x_2^2 + x_3^2) &+ b(x_1 + x_2 + x_3) &&= x_1y_1 + x_2y_2 + x_3y_3, \\ a(x_1 + x_2 + x_3) &+ b \cdot 3 &&= y_1 + y_2 + y_3. \end{aligned}$$

4 Odvození vzorce

Uvažujme tři body $[x_1, y_1]$, $[x_2, y_2]$ a $[x_3, y_3]$. Svislé vzdálenosti těchto bodů od přímky $y = ax + b$ jsou $s_1 = |ax_1 + b - y_1|$, $s_2 = |ax_2 + b - y_2|$, $s_3 = |ax_3 + b - y_3|$ a chceme minimalizovat funkci

$$S(a, b) = (ax_1 + b - y_1)^2 + (ax_2 + b - y_2)^2 + (ax_3 + b - y_3)^2.$$

Lokální extrém diferencovatelné funkce může nastat pouze ve stacionárním bodě. Platí

$$\begin{aligned}\frac{\partial S}{\partial a} &= 2(ax_1 + b - y_1)x_1 + 2(ax_2 + b - y_2)x_2 + 2(ax_3 + b - y_3)x_3 \\ &= 2a(x_1^2 + x_2^2 + x_3^2) + 2b(x_1 + x_2 + x_3) - 2(x_1y_1 + x_2y_2 + x_3y_3)\end{aligned}$$

a

$$\begin{aligned}\frac{\partial S}{\partial b} &= 2(ax_1 + b - y_1) + 2(ax_2 + b - y_2) + 2(ax_3 + b - y_3) \\ &= 2a(x_1 + x_2 + x_3) + 2 \cdot 3b - 2(y_1 + y_2 + y_3).\end{aligned}$$

Položíme-li derivace rovny nule, dostáváme po vydělení dvěma a přesunu členů bez neznámých na pravou stranu rovnice

$$\begin{aligned}a(x_1^2 + x_2^2 + x_3^2) &+ b(x_1 + x_2 + x_3) &= x_1y_1 + x_2y_2 + x_3y_3, \\ a(x_1 + x_2 + x_3) &+ b \cdot 3 &= y_1 + y_2 + y_3.\end{aligned}$$

Pro obecný případ množiny n bodů $[x_1, y_1]$, ..., $[x_n, y_n]$ je cenová funkce $S(a,b) = \sum_{i=1}^{n} (ax_i + b - y_i)^2$ a parciální derivace jsou

$$\frac{\partial S}{\partial a} = 2\sum_{i=1}^{n}(ax_i + b - y_i)x_i = 2\left(a\sum_{i=1}^{n}x_i^2 + b\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}x_iy_i\right),$$

$$\frac{\partial S}{\partial b} = 2\sum_{i=1}^{n}(ax_i + b - y_i) = 2\left(a\sum_{i=1}^{n}x_i + b\sum_{i=1}^{n}1 - \sum_{i=1}^{n}y_i\right).$$

Protože $\sum_{i=1}^{n} 1 = \underbrace{1 + 1 + \cdots + 1}_{n\text{-krát}} = n$, získáváme nulováním parciálních derivací soustavu rovnic

$$\begin{aligned} a\sum_{i=1}^{n}x_i^2 + b\sum_{i=1}^{n}x_i &= \sum_{i=1}^{n}x_iy_i, \\ a\sum_{i=1}^{n}x_i + b\,n &= \sum_{i=1}^{n}y_i. \end{aligned} \tag{2}$$

Po dosazení získáváme soustavu lineárních rovnic s neznámými a, b, kterou vyřešíme a získáme koeficienty do rovnice přímky $y = ax + b$.

Pro obecný případ množiny n bodů $[x_1, y_1]$, ..., $[x_n, y_n]$ je cenová funkce $S(a,b) = \sum_{i=1}^{n}(ax_i + b - y_i)^2$ a parciální derivace jsou

$$\frac{\partial S}{\partial a} = 2\sum_{i=1}^{n}(ax_i + b - y_i)x_i \quad = 2\left(a\sum_{i=1}^{n}x_i^2 + b\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}x_iy_i\right),$$

$$\frac{\partial S}{\partial b} = 2\sum_{i=1}^{n}(ax_i + b - y_i) \quad = 2\left(a\sum_{i=1}^{n}x_i + b\sum_{i=1}^{n}1 - \sum_{i=1}^{n}y_i\right).$$

Protože $\sum_{i=1}^{n} 1 = \underbrace{1 + 1 + \cdots + 1}_{n\text{-krát}} = n$, získáváme nulováním parciálních derivací soustavu rovnic

$$\begin{aligned} a\sum_{i=1}^{n}x_i^2 &+ b\sum_{i=1}^{n}x_i &= \sum_{i=1}^{n}x_iy_i, \\ a\sum_{i=1}^{n}x_i &+ b\,n &= \sum_{i=1}^{n}y_i. \end{aligned} \tag{2}$$

Po dosazení získáváme soustavu lineárních rovnic s neznámými a, b, kterou vyřešíme a získáme koeficienty do rovnice přímky $y = ax + b$.

5 Otázky pozorného čtenáře

Pozorný čtenář si všiml, že jsme hledání lokálního extrému cenové funkce $S(a, b)$ poněkud odbyli a napadnou ho hned dvě otázky.

- *Našli jsme stacionární body funkce $S(a, b)$. Kde je však dokázáno, že se jedná o minimum?* Ano, otázku zda ve stacionárním bodě skutečně nastává extrém a jaký jsme neřešili. Z povahy úlohy je totiž zřejmé, že nějaká optimální přímka bude existovat a protože nám vyšel jediný stacionární bod, jedná se právě o toto optimální řešení, tj. minimum cenové funkce.

- *Má nalezená soustava (2) skutečně řešení? (Vím, že řešení soustavy dvou lineárních rovnic o dvou neznámých existovat nemusí, nebo jich může být nekonečně mnoho.)* Podrobnější rozbor ukazuje, že případ kdy řešení systému (2) není určeno jednoznačně nebo neexistuje nastane pokud všechny body mají stejnou x-ovou souřadnici. Toto je zcela zřejmě případ, který zůstává stranou "rozumných" aplikací této metody. V praktických případech je tedy jednoznačná řešitelnost systému (2) zaručena.

5 Otázky pozorného čtenáře

Pozorný čtenář si všiml, že jsme hledání lokálního extrému cenové funkce $S(a, b)$ poněkud odbyli a napadnou ho hned dvě otázky.

- *Našli jsme stacionární body funkce* $S(a, b)$*. Kde je však dokázáno, že se jedná o minimum?* Ano, otázku zda ve stacionárním bodě skutečně nastává extrém a jaký jsme neřešili. Z povahy úlohy je totiž zřejmé, že nějaká optimální přímka bude existovat a protože nám vyšel jediný stacionární bod, jedná se právě o toto optimální řešení, tj. minimum cenové funkce.

- *Má nalezená soustava* (2) *skutečně řešení? (Vím, že řešení soustavy dvou lineárních rovnic o dvou neznámých existovat nemusí, nebo jich může být nekonečně mnoho.)* Podrobnější rozbor ukazuje, že případ kdy řešení systému (2) není určeno jednoznačně nebo neexistuje nastane pokud všechny body mají stejnou x-ovou souřadnici. Toto je zcela zřejmě případ, který zůstává stranou "rozumných" aplikací této metody. V praktických případech je tedy jednoznačná řešitelnost systému (2) zaručena.