Jak proložit přímku ?

Petr Vodstrčil

petr.vodstrcil@vsb.cz

Katedra aplikované matematiky, Fakulta elektrotechniky a informatiky,
Vysoká škola báňská–Technická univerzita Ostrava

28.1. 2010

Minimalizace kvadratické funkce

Příklad (Kvadratická funkce jedné proměnné)

Pro jaké $a \in \mathbb{R}$ je výraz

$$2a^2 - 6a + 1$$

minimální? $\left[a = \frac{3}{2}\right]$

Příklad (Kvadratická funkce dvou proměnných)

Pro kterou dvojici $(a, b) \in \mathbb{R}^2$ je výraz

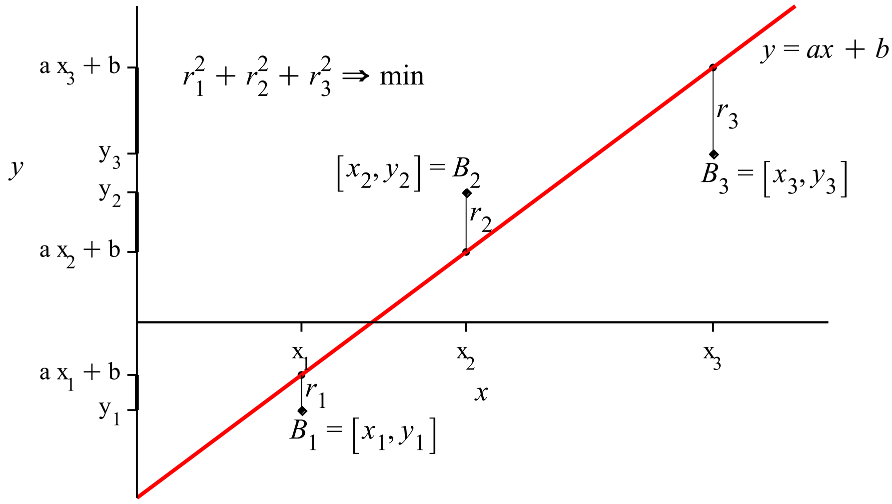
$$2a^2 + 21b^2 + 12ab - 20a - 72b + 11$$

minimální? $[a = -1,\ b = 2]$

Uvažujme body $B_1 = [x_1, y_1], B_2 = [x_2, y_2], \dots, B_n = [x_n, y_n]$ v rovině. Těmito body budeme chtít proložit tzv. regresní přímku. To je přímka, která co nejlépe „kopíruje“ zadané body.

Rovnice takovéto přímky je $y = ax + b$, přičemž koeficienty $a, b \in \mathbb{R}$ volíme tak, aby výraz $\sum_{i=1}^{n} \left(ax_i + b - y_i\right)^2$ byl minimální (metoda nejmenších čtverců).

Pro $n = 3$ je situace znázorněna na následujícím obrázku.



Poznámka

Danými body $B_1 = [x_1, y_1], B_2 = [x_2, y_2], \dots, B_n = [x_n, y_n]$ nemusíme vždy prokládat jenom přímku $y = ax + b$. Např. je možné těmito body proložit přímku $y = ax$ (procházející počátkem), popř. parabolu, exponenciálu, logaritmickou funkci, apod.

- V případě, že danými body chceme proložit přímku $y = ax$, musíme koeficient a zvolit tak, aby byl výraz $\sum_{i=1}^{n} (ax_i - y_i)^2$ minimální.
- Chceme-li našimi body proložit parabolu $y = ax^2 + bx + c$, zvolíme koeficienty a, b, c tak, aby výraz $\sum_{i=1}^{n} (ax_i^2 + bx_i + c - y_i)^2$ byl minimální.

My však budeme zadanými body prokládat výhradně přímku. Nejprve si ukážeme, jak proložit přímku $y = ax$ (procházející počátkem). Poté zkusíme proložit i přímku v obecné poloze, tj. přímku o rovnici $y = ax + b$.

Příklad

Jsou dány body $B_1 = [1, -1]$, $B_2 = [2, 3]$, $B_3 = [3, 4]$. Najděte regresní přímku procházející počátkem.

Řešení.

Přímka procházející počátkem má rovnici $y = ax$. Podle předchozího hledáme koeficient $a \in \mathbb{R}$ takový, aby výraz

$$\sum_{i=1}^{3} (ax_i - y_i)^2 = (a+1)^2 + (2a-3)^2 + (3a-4)^2 = 14a^2 - 34a + 26$$

byl minimální. To je však kvadratická funkce, kterou můžeme upravit na čtverec, tj.

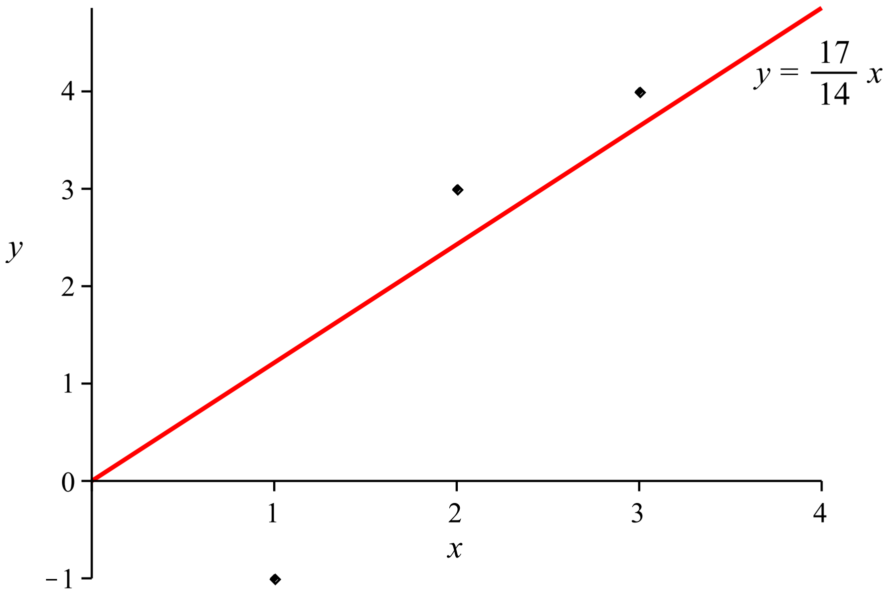
$$14a^2 - 34a + 26 = 14\left(a^2 - \frac{17}{7}a + \frac{13}{7}\right) =$$

$$= 14\left[\left(a - \frac{17}{14}\right)^2 - \left(\frac{17}{14}\right)^2 + \frac{13}{7}\right] = 14\left(a - \frac{17}{14}\right)^2 + \frac{75}{14}.$$

Řešení.

Vidíme tedy, že výraz je minimální pro $a = \frac{17}{14}$ a hledaná regresní přímka má rovnici $y = \frac{17}{14}x$. $\square$

Situace je znázorněna na následujícím obrázku.



Poznámka (Zobecnění pro n bodů)

Je-li dáno n bodů $B_1 = [x_1, y_1], B_2 = [x_2, y_2], \dots, B_n = [x_n, y_n]$, pak regresní přímka procházející počátkem má rovnici

$$y = ax,$$

kde

$$a = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}.$$

Příklad

Jsou dány body $B_1 = [1, -1]$, $B_2 = [2, 3]$, $B_3 = [3, 4]$ (stejné jako v předchozím příkladu). Najděte regresní přímku $y = ax + b$.

Řešení.

Rovnice hledané přímky je $y = ax + b$. Koeficienty $a, b \in \mathbb{R}$ přitom musíme zvolit tak, aby výraz

$$\sum_{i=1}^{3} (ax_i + b - y_i)^2 = (a + b + 1)^2 + (2a + b - 3)^2 + (3a + b - 4)^2 =$$

$$= 14a^2 + 3b^2 + 12ab - 34a - 12b + 26$$

byl minimální. Upravíme-li tento výraz na čtverec, dostaneme:

Řešení.

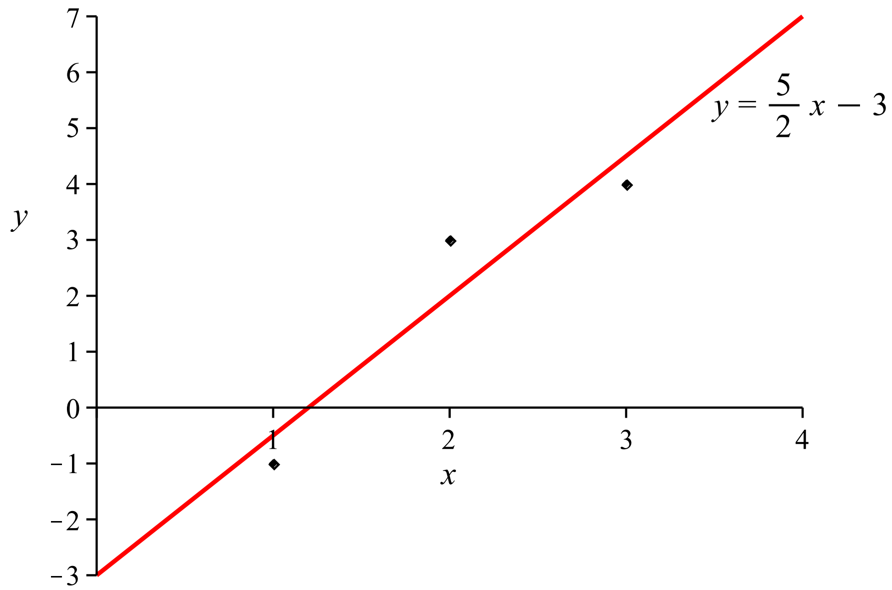
$$
\begin{aligned}
14a^2 + 3b^2 + 12ab - 34a - 12b + 26 &= (3b^2 + 12ab - 12b) + \\
+ (14a^2 - 34a + 26) &= 3\big(b^2 + b(4a-4)\big) + (14a^2 - 34a + 26) = \\
&= 3\left[\big(b + (2a-2)\big)^2 - (2a-2)^2\right] + (14a^2 - 34a + 26) = \\
&= 3(b + 2a - 2)^2 - 3(2a-2)^2 + (14a^2 - 34a + 26) = \\
&= 3(b+2a-2)^2 + (2a^2 - 10a + 14) = 3(b+2a-2)^2 + 2(a^2 - 5a + 7) = \\
&= 3(b+2a-2)^2 + 2\left[\left(a - \frac{5}{2}\right)^2 - \left(\frac{5}{2}\right)^2 + 7\right] = \\
&= 3(b+2a-2)^2 + 2\left(a - \frac{5}{2}\right)^2 + \frac{3}{2}.
\end{aligned}
$$

Tento výraz je zřejmě minimální, pokud $(b + 2a - 2) = 0$ a $\left(a - \frac{5}{2}\right) = 0$, odkud $a = \frac{5}{2}$ a $b = -3$.

Řešení.

Hledaná regresní přímka má rovnici $y = \frac{5}{2}x - 3$. □

Celou situaci opět vystihuje obrázek.



Poznámka (Zobecnění pro n bodů)

Je-li dáno n bodů $B_1 = [x_1, y_1], B_2 = [x_2, y_2], \dots, B_n = [x_n, y_n]$, pak regresní přímka má rovnici

$$y = ax + b,$$

kde

$$a = \frac{n\left(\sum\limits_{i=1}^{n} x_i y_i\right) - \left(\sum\limits_{i=1}^{n} x_i\right)\left(\sum\limits_{i=1}^{n} y_i\right)}{n\left(\sum\limits_{i=1}^{n} x_i^2\right) - \left(\sum\limits_{i=1}^{n} x_i\right)^2}$$

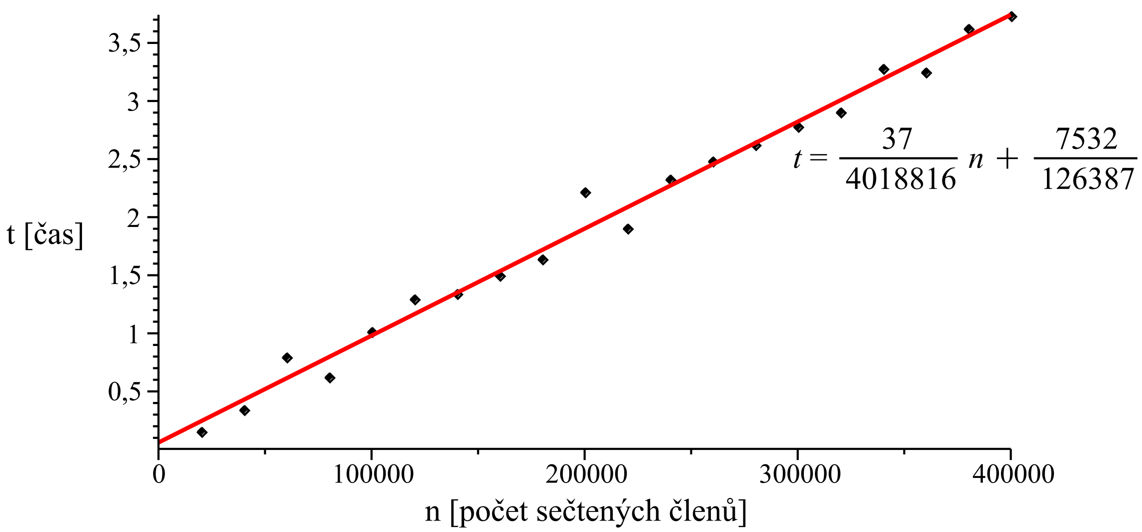
a

$$b = \frac{\left(\sum\limits_{i=1}^{n} y_i\right)\left(\sum\limits_{i=1}^{n} x_i^2\right) - \left(\sum\limits_{i=1}^{n} x_i\right)\left(\sum\limits_{i=1}^{n} x_i y_i\right)}{n\left(\sum\limits_{i=1}^{n} x_i^2\right) - \left(\sum\limits_{i=1}^{n} x_i\right)^2}.$$

Nyní si ukážeme příklad reálného měření, kdy jsme na počítači sčítali prvních n členů řady

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots$$

a sledovali jsme čas potřebný k výpočtu. Výslednými daty jsme pak proložili regresní přímku.



Děkuji za pozornost !!!