

Vyrovňovací rovina (jednoduchý případ)

Představme si pravoúhlý souřadnicový systém, kde osy x a y vytvářejí horizontální rovinu a osa z vyjadřuje např. měřené relativní výšky vzhledem k rovině xy . Souřadnice x a y považujeme za bezchybné (nezávislé) a měřickými chybami bude zatížena pouze souřadnice z (závislá veličina). Vlastní rovina může být šikmá a je vyjádřena třemi parametry (výškou v počátku, směrnici ve směru osy x a směrnici ve směru osy y). Naším úkolem je proložit naměřenou skupinou bodů (třemi a více) rovinu tak, aby součet čtverců odchylek ve směru osy z byl minimální. $\sum v_z^2 = \min$. Označme v dalším $\Sigma = [\quad]$.

Zprostředkující funkce:

$$\tilde{z}_i = \beta_0 + \beta_1 \tilde{x}_i + \beta_2 \tilde{y}_i$$

Rovnice oprav (bezchybné x, y):

$$v_i = \beta_0 + \beta_1 x_i + \beta_2 y_i - z_i$$

$$\mathbf{v} = \mathbf{A}\boldsymbol{\beta} - \mathbf{l}$$

$$\mathbf{A} = \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \quad \mathbf{l} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix},$$

Normální rovnice

$$\mathbf{A}^T \mathbf{A} \boldsymbol{\beta} - \mathbf{A}^T \mathbf{l} = \mathbf{0}$$

$$\mathbf{N} \boldsymbol{\beta} - \mathbf{n} = \mathbf{0}$$

$$\begin{pmatrix} n & [x] & [y] \\ [x] & [xx] & [xy] \\ [y] & [xy] & [yy] \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} - \begin{pmatrix} [z] \\ [xz] \\ [yz] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Další postup je standardní řešení jako při vyrovnáním zprostředkujících měření.

Transformace některých funkcí na funkce lineární

Některé nelineární funkce lze jednoduchou transformací převést na funkce lineární a použít regresní analýzu metodou nejmenších čtverců (vyrovnávací přímkou) až u takto transformovaných lineárních funkcí. Při linearizaci se využívá zejména logaritmování původních funkcí.

$$y_i = a x_i^b,$$

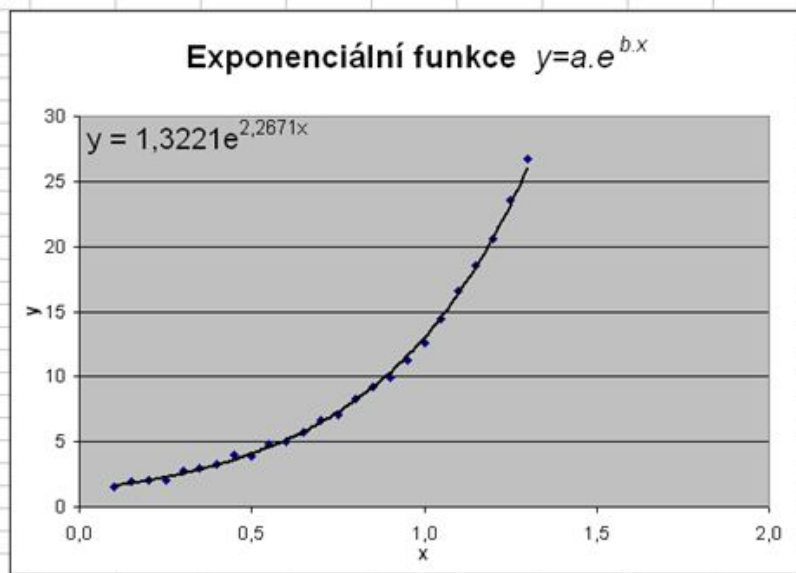
$$\ln y_i = \ln a + b \ln x_i, \quad u_i = \beta_0 + \beta_1 v_i, \quad \text{kde}$$

$$u_i = \ln y_i, \quad v_i = \ln x_i, \quad \beta_0 = \ln a, \quad \beta_1 = b$$

Obdobně lze linearizovat logaritmováním funkci:

$$y_i = a e^{bx_i}, \text{ nebo } y_i = a b^{x_i}, \quad b > 0, b \neq 1$$

Metodu považujeme za přibližnou, neboť minimalizuje součet čtverců logaritmovaných veličin a ne veličin původních.

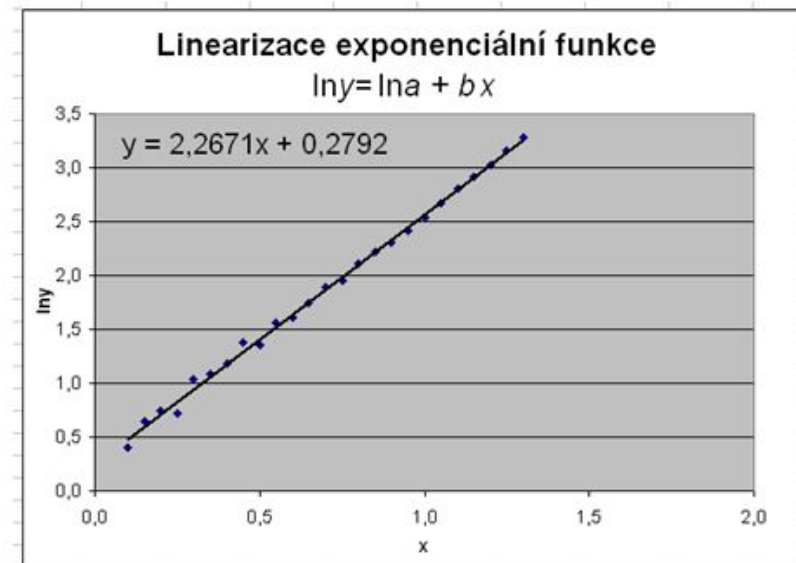


Ukázka transformace nelineární (exponenciální) funkce na funkci lineární.

První graf ukazuje průběh exponenciální funkce v původní podobě.

U těchto funkcí lze parametry rovnic určit přímo v Excelu (viz. rovnice v grafu).

Na druhém grafu je vidět průběh stejné funkce po její linearizaci logaritmováním. Parametry rovnice přímky jsou rovněž určeny přímo v Excelu.



Hledané parametry exponenciální funkce získáme až zpětnou transformací parametrů určených při lineárním řešení:
 $a = e^{2,2671} = 1,3221$, $b = 2,2671$

Pozor: Excel označil rovnici regresní přímky jako $y = \dots$, i když se vlastně jedná o ln y.

Další přibližné metody pro určení parametrů vyrovnávací přímky

A) Metody grafické:

u těchto metod vyneseme jednotlivé body v příslušném měřítku na papír (např. milimetrový) a přímku proložíme od oka pomocí pravítka. Určované parametry poté odměříme nebo vypočteme.

B) Metody graficko-početní

tyto metody jsou obdobou předcházejících grafických metod, ale můžeme přímku např. otáčet kolem předem vypočteného těžiště, řešit grafické metody s různými měřítky na jednotlivých osách, na logaritmických papírech apod.

C) Skupinová metoda

u této metody se naměřený soubor rozdělí na dvě přibližně stejné četné skupiny a v každé z nich se určí těžiště. Přímka je pak jednoznačně určena dvěma body-těžišti jednotlivých skupin. Střední odchylka v první skupině by měla být přibližně stejně velká jako v druhé skupině. Pozor na takové rozdělení skupin, u kterých by se obě těžiště nacházely velmi blízko sebe, určení parametrů takovéto přímky je pak velmi nespolehlivé (může se vyskytnout např. při dělení na skupinu lichých a skupinu sudých měření seřazených podle pořadí zkoumaného trendu).

Lineární regrese v Excelu (1)

Excel nabízí nástroje pro lineární i nelineární regresi:

Určujeme-li parametry pouze jedné přímky, je výhodné postupovat následovně:

1. Nezávislou proměnnou x (bezchybnou) uspořádáme do sloupcového vektoru
Závislou proměnnou y (zatíženou chybami) uspořádáme do dalšího sloupcového vektoru stejné délky. Excel určuje přímku s označením parametrů $y = mx + b$.
2. Označíme 2 sousední buňky (v řádku) a napíšeme do příkazového řádku funkci: `=LINREGRESE(pole_y;pole_x)`. Příkaz ukončíme trojhmatem pro maticové operace CTRL+SHIFT+ENTER.
3. V označených buňkách bude v první buňce násobný parametr m (naše značení B, β_1), v druhé buňce součtový parametr b (naše značení A, β_0).

Pozn: pokud má přímka procházet počátkem, bude příkaz vypadat následovně: `=LINREGRESE(pole_y;pole_x;0)`. Poslední operátor je logický a vyjadřuje veličinu NEPRAVDA.

Součtový parametr můžeme do předem označené buňky vypočítat též příkazem: `=INDEX(LINREGRESE(pole_y;pole_x);2)` nebo příkazem: `=INTERCEPT(pole_y;pole_x)`.

Násobný parametr do další označené buňky příkazem:

`=INDEX(LINREGRESE(pole_y;pole_x);1)` nebo příkazem:

`=SLOPE(pole_y;pole_x)`. Protože je výsledkem vždy jen jedno číslo, nemusíme používat trojhmat.

Ukázka příkazů pro lineární regresi v Excelu (2)

B20 $\text{=LINREGRESE}(C7:C16;B7:B16)$

Sešit2					
	A	B	C	D	E
1	LINEÁRNÍ REGRESE				
2					
3					
4		DATA			
5		x	y		
6		[m]	[mm]		
7		2	0,50		
8		4	0,80		
9		6	0,80		
10		8	1,25		
11		10	1,50		
12		12	2,20		
13		14	2,45		
14		16	2,40		
15		18	3,15		
16		20	3,50		
17					
18		Výsledek			
19		m=B	b=A		
20		0,168939	-0,00333		
21					

B20 $\text{=INDEX}(\text{LINREGRESE}(C7:C16;B7:B16);1)$

Sešit2						
	A	B	C	D	E	F
1	LINEÁRNÍ REGRESE					
2						
3						
4		DATA				
5		x	y			
6		[m]	[mm]			
7		2	0,50			
8		4	0,80			
9		6	0,80			
10		8	1,25			
11		10	1,50			
12		12	2,20			
13		14	2,45			
14		16	2,40			
15		18	3,15			
16		20	3,50			
17						
18		Výsledek				
19		m=B	b=A			
20		0,168939				
21						

C20 $\text{=INDEX}(\text{LINREGRESE}(C7:C16;B7:B16);2)$

Sešit2						
	A	B	C	D	E	F
1	LINEÁRNÍ REGRESE					
2						
3						
4		DATA				
5		x	y			
6		[m]	[mm]			
7		2	0,50			
8		4	0,80			
9		6	0,80			
10		8	1,25			
11		10	1,50			
12		12	2,20			
13		14	2,45			
14		16	2,40			
15		18	3,15			
16		20	3,50			
17						
18		Výsledek				
19		m=B	b=A			
20			-0,00333			
21						

B20 $\text{=LINREGRESE}(C7:C16;B7:B16;0)$

Sešit2						
	A	B	C	D	E	F
1	LINEÁRNÍ REGRESE					
2						
3						
4		DATA				
5		x	y			
6		[m]	[mm]			
7		2	0,50			
8		4	0,80			
9		6	0,80			
10		8	1,25			
11		10	1,50			
12		12	2,20			
13		14	2,45			
14		16	2,40			
15		18	3,15			
16		20	3,50			
17						
18		Výsledek				
19		m=B	b=A			
20		0,168701	0			
21						

B20 $\text{=SLOPE}(C7:C16;B7:B16)$

Sešit2					
	A	B	C	D	E
1	LINEÁRNÍ REGRESE				
2					
3					
4		DATA			
5		x	y		
6		[m]	[mm]		
7		2	0,50		
8		4	0,80		
9		6	0,80		
10		8	1,25		
11		10	1,50		
12		12	2,20		
13		14	2,45		
14		16	2,40		
15		18	3,15		
16		20	3,50		
17					
18		Výsledek			
19		m=B	b=A		
20		0,168939			
21					

C20 $\text{=INTERCEPT}(C7:C16;B7:B16)$

Sešit2					
	A	B	C	D	E
1	LINEÁRNÍ REGRESE				
2					
3					
4		DATA			
5		x	y		
6		[m]	[mm]		
7		2	0,50		
8		4	0,80		
9		6	0,80		
10		8	1,25		
11		10	1,50		
12		12	2,20		
13		14	2,45		
14		16	2,40		
15		18	3,15		
16		20	3,50		
17					
18		Výsledek			
19		m=B	b=A		
20			-0,00333		
21					

Pokračování lineární regrese v Excelu (3)

Funkce LINREGRESE nabízí i výpočet směrodatných odchylek určených parametry. Po zadání výstupní matice 2x2 prvků napíšeme příkaz: =LINREGRESE(pole_y;pole_x;;PRAVDA), který vypočítá v prvním řádku násobný a součtový parametr a v druhém řádku jejich směrodatné odchylky, viz první obrázek. Protože je výstupem matice, musíme zadání zakončit trojúhánkem. Na druhém obrázku je ukázána funkce =LINTREND(pole_y;pole_x), která do vybraného pole vypočítá přímo vyrovnané hodnoty = vyrovnané body ležící na vyrovnávací přímce. Uvedené funkce poskytují ještě další možnosti, např. statistického zhodnocení.

B21		fx (=LINREGRESE(C7:C16;B7:B16;;1))	
A	B	C	D
1	LINEÁRNÍ REGRESE		
2			
3			
4	DATA		
5	x	y	
6	[m]	[mm]	
7	2	0,50	
8	4	0,80	
9	6	0,80	
10	8	1,25	
11	10	1,50	
12	12	2,20	
13	14	2,45	
14	16	2,40	
15	18	3,15	
16	20	3,50	
17			
18	Výsledek		
19	m=B	b=A	
20	σB	σA	
21	0,168939	-0,003333	
22	0,010461	0,129812	
23			
24			

E7	fx (=LINTREND(C7:C16;B7:B16))				
A	B	C	D	E	F
1	LINEÁRNÍ REGRESE				
2					
3					
4	DATA				
5	x1	y		y - vyrovnané	
6	[m]	[mm]		[mm]	
7	2	0,50		0,334545	
8	4	0,80		0,672424	
9	6	0,80		1,010303	
10	8	1,25		1,348182	
11	10	1,50		1,686061	
12	12	2,20		2,023939	
13	14	2,45		2,361818	
14	16	2,40		2,699697	
15	18	3,15		3,037576	
16	20	3,50		3,375455	
17					
18	Výsledek				
19	m=B	b=A			
20	0,168939	-0,003333			
21					

Pokračování lineární regrese v Excelu (4)

Silnou stránkou Excelu je i tvorba grafů. Na prvním obrázku je ukázka výběru typu grafu XYbodový se zákresem naměřených dat. Na dalším obrázku proložení regresní přímky s rovnicí s vypočtenými parametry. Rovnice se vkládá v záložce *Možnosti*, kde je možno volit průchod přímky počátkem či výpočet koeficientu determinace (bude objasněno později).

