LABORATOR 5

METODE NUMERICE ÎN MODELAREA MATEMATICĂ

1. Regresia liniară

Fiind dat sistemul de puncte (x_i, y_i), $i = 1..n$, se cere să se determine dreapta $y = ax + b$ (a şi b nedeterminaţi) care trece aproximativ prin (sau printre) aceste puncte.

Aplicând "*metoda celor mai mici pătrate*", se consideră funcţia

$$F(a,b) = \sum_{i=1}^{n} (ax_i + b - y_i)^2 .$$

Impunând condiţia ca $F(a,b)$ să fie minimă, rezultă $\frac{\partial F}{\partial a} = 0$ şi $\frac{\partial F}{\partial b} = 0$ se obţine sistemul

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

$$a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i$$

din care se obţin coeficienţii a şi b. În practică se calculează parametri modelării

- Mediile de selecţie: $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$; $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
- Abaterile medii pătratice de selecţie

$$S_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2} \; ; \quad S_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2}$$

- Coeficientul de corelaţie dintre variabilele x şi y

$$r_{xy} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{S_x S_y} \quad \Leftrightarrow \quad r_{xy} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i y_i) - \overline{x}\cdot\overline{y}}{S_x S_y}$$

exprimă intensitatea dependenței liniare dintre cei doi parametrii şi variază între -1 şi +1.

- Abaterea medie pătratică de la dreapta de regresie

$$S_{xy} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - ax_i - b)^2}$$

care se utilizează pentru a măsura împrăştierea punctelor în jurul dreptei de regresie.

Problemă. Să se determine ecuația dreptei de regresie corespunzătoare măsurătorilor

x	-2	-1	0	1	2
y	-1.2	0.2	1.3	1.9	3.2

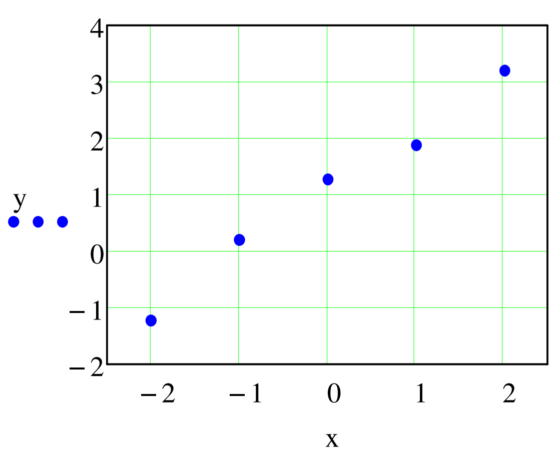
Program:

ORIGIN $\equiv$ 1

$$x := \begin{pmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{pmatrix} \qquad y := \begin{pmatrix} -1.2 \\ 0.2 \\ 1.3 \\ 1.9 \\ 3.2 \end{pmatrix}$$

$n := 5$ numarul perechilor de puncte date

Reprezentarea grafica a punctelor:



Calculul coeficientilor a si b:

Matricea sistemului:

$$S := \begin{bmatrix} \sum_{i=1}^{n} (x_i)^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{bmatrix}$$

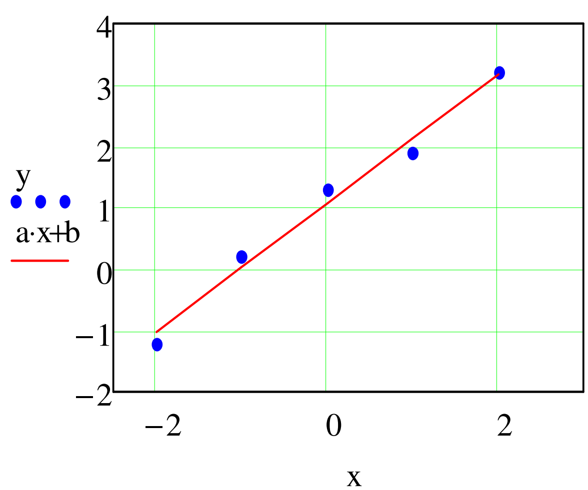
Matricea termenilor liberi:

$$T := \begin{bmatrix} \sum_{i=1}^{n} (x_i \cdot y_i) \\ \sum_{i=1}^{n} y_i \end{bmatrix}$$

$$\begin{pmatrix} a \\ b \end{pmatrix} := S^{-1} \cdot T$$

$a = 1.05 \qquad b = 1.08$

Reprezentarea dreptei de regresie:



Aplicații: Să se determine dreapta de regresie pentru următoarele sisteme de puncte, calculând parametrii modelărilor

a)

x	-1	0	1
y	2.3	1.5	0.6

b)

x	-11	-3	-1	5	7	9
y	-5	2	4	11	15	24

c)

x	-11	-5	-3	5	8	9
y	20	7	1	-5	-10	-14

2. ***Parabola de regresie***

Fiind dat sistemul de puncte (x_i, y_i), $i=1..n$, se cere să se determine parabola $y = ax^2 + bx + c$ (a, b şi c nedeterminați) care modelează fenomenul. Aplicând "*metoda celor mai mici pătrate*", se consideră funcția

$$F(a,b,c) = \sum_{i=1}^{n} \left(ax_i^2 + bx_i + c - y_i\right)^2 .$$

Impunând condiția ca $F(a,b,c)$ să fie minimă, rezultă $\frac{\partial F}{\partial a} = 0$, $\frac{\partial F}{\partial b} = 0$ şi $\frac{\partial F}{\partial c} = 0$ se obține sistemul

$$a\sum_{i=1}^{n} x_i^4 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i^2 y_i$$

$$a\sum_{i=1}^{n} x_i^3 + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i\, y_i$$

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i + cn = \sum_{i=1}^{n} y_i$$

din care se obțin coeficienții a, b şi c.

Problemă. Să se determine ecuația parabolei de regresie corespunzătoare următoarelor date, calculând parametrii modelării

x	-7	-5	-3	1	3	6
y	20	11	5	-1	7	19

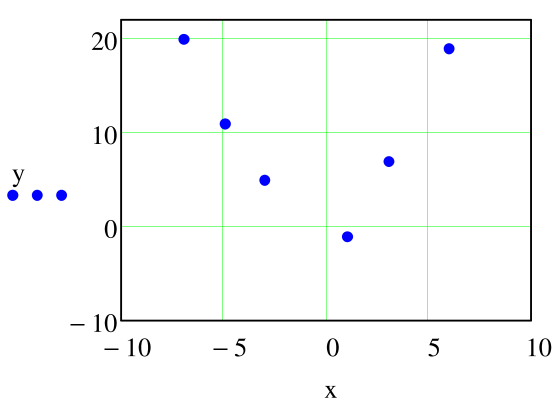
Program:

ORIGIN ≡ 1

$$x := \begin{pmatrix} -7 \\ -5 \\ -3 \\ 1 \\ 3 \\ 6 \end{pmatrix} \qquad y := \begin{pmatrix} 20 \\ 11 \\ 5 \\ -1 \\ 7 \\ 19 \end{pmatrix}$$

n := 6 numarul perechilor de puncte date

Reprezentarea grafica a punctelor:



Calculul coeficientillor a, b si c:

$i := 1 .. n$

Matricea sistemului:

$$S := \begin{bmatrix} \sum_i (x_i)^4 & \sum_i (x_i)^3 & \sum_i (x_i)^2 \\ \sum_i (x_i)^3 & \sum_i (x_i)^2 & \sum_i x_i \\ \sum_i (x_i)^2 & \sum_i x_i & n \end{bmatrix}$$

Matricea termenilor liberi:

$$T := \begin{bmatrix} \sum_i \left[(x_i)^2 \cdot y_i\right] \\ \sum_i (x_i \cdot y_i) \\ \sum_i y_i \end{bmatrix}$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} := S^{-1} \cdot T$$

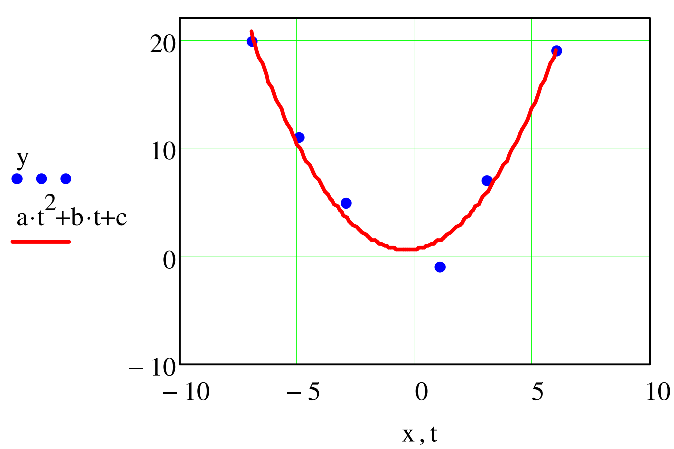
$a = 0.456$

$b = 0.323$

$c = 0.627$

Reprezentarea parabolei de regresie:

$t := \min(x), \min(x) + 0.1 .. \max(x)$



Mediile de selectie:

$$xm := \frac{\sum_{i=1}^{n} x_i}{n} \qquad ym := \frac{\sum_{i=1}^{n} y_i}{n}$$

$$xm = -0.833 \qquad ym = 10.167$$

Abaterile medii patratice de selectie:

$$Sx := \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - xm)^2} \qquad Sx = 4.561$$

$$Sy := \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (y_i - ym)^2} \qquad Sy = 7.493$$

Coeficientul de corelatie:

$$\rho := \sqrt{1 - \frac{\sum_{i=1}^{n} \left[y_i - a \cdot (x_i)^2 - b \cdot x_i - c\right]^2}{\sum_{i=1}^{n} (y_i - ym)^2}} \qquad \rho = 0.985$$

Abaterea medie patratica:

$$S := \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} \left[y_i - a \cdot (x_i)^2 - b \cdot x_i - c\right]^2} \qquad S = 1.282$$

Aplicații: Să se determine parabola de regresie şi parametrii modelărilor pentru următoarele sisteme de puncte

a)

x	-2	0	1	2
y	-3.3	1.5	0.6	-3.5

b)

x	-1	0	1	2
y	2.3	1.5	1.6	5.3

c)

x	10	11.7	37	48	9	43	47.5
y	61	59	40	35	50	63	30

3. <u>Modelarea</u> <u>exponențială</u>

Fiind dat sistemul de puncte (x_i, y_i), $i = 1..n$, se caută o funcție de forma

$$y = a\,e^{b\,x}$$

(unde a, b sunt necunoscute), care să modeleze cel mai bine norul de puncte.

Prin logaritmarea ecuației anterioare se obține

$$\ln y = \ln a + b\,x\,,$$

care, prin notațiile $A = \ln a$ şi $Y = \ln y$, devine

$$Y = A + bx\,,$$

necunoscutele fiind A şi b.

Aceste necunoscute se vor determina impunând condiția de minimizare a funcției

$$F(a,b) = \sum_{i=1}^{n} [A + bx_i - \ln y_i]^2\,,$$

Se calculează în continuare derivatele parțiale în raport cu A şi b

$$F_A^{'} = \sum_{i=1}^{n} 2(A + bx_i - \ln y_i)\,,$$

$$F_b^{'} = \sum_{i=1}^{n} 2(A + bx_i - \ln y_i) \cdot x_i\,,$$

care, prin anulare, conduc la sistemul

$$\begin{cases} A\sum_{i=1}^{n} 1 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \ln y_i \\ A\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \ln y_i. \end{cases}$$

Parametri modelării sunt

- coeficientul de corelație

$$\rho = \sqrt{1 - \frac{\sum_{i=1}^{n} (ae^{b\,x_i} - y_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2}},$$

- abaterea medie pătratică

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (ae^{b\,x_i} - y_i)^2}.$$

Problemă. Să se realizeze o aplicație MathCAD pentru modelarea exponențială a datelor următoare

x	1.2	0.8	1.3	1.9	3.2
y	-2	-1	0	1	2

calculând parametrii modelării.