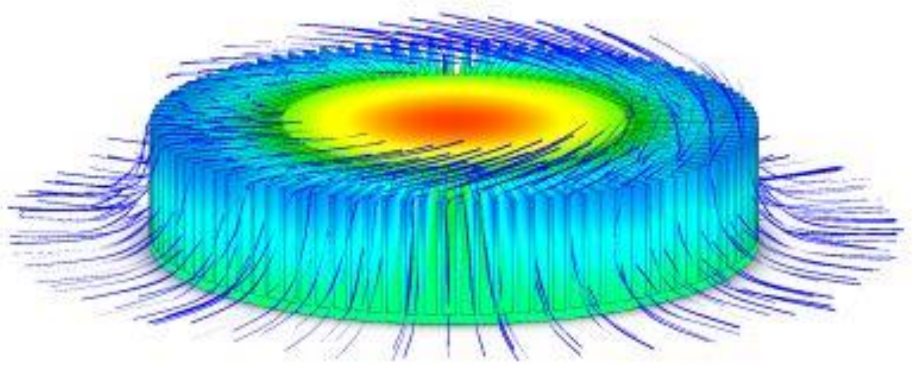
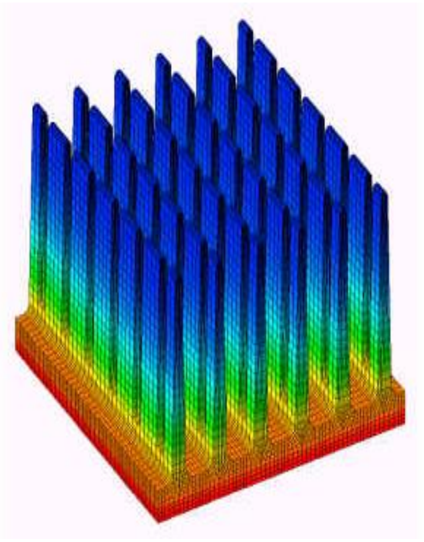
Ecuaţii cu derivate parţiale

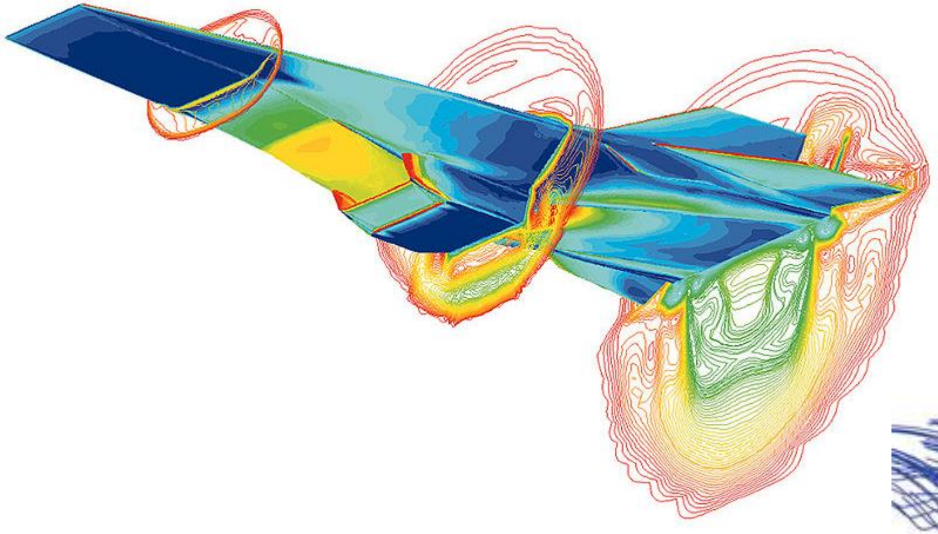
1. Noţiuni teoretice introductive

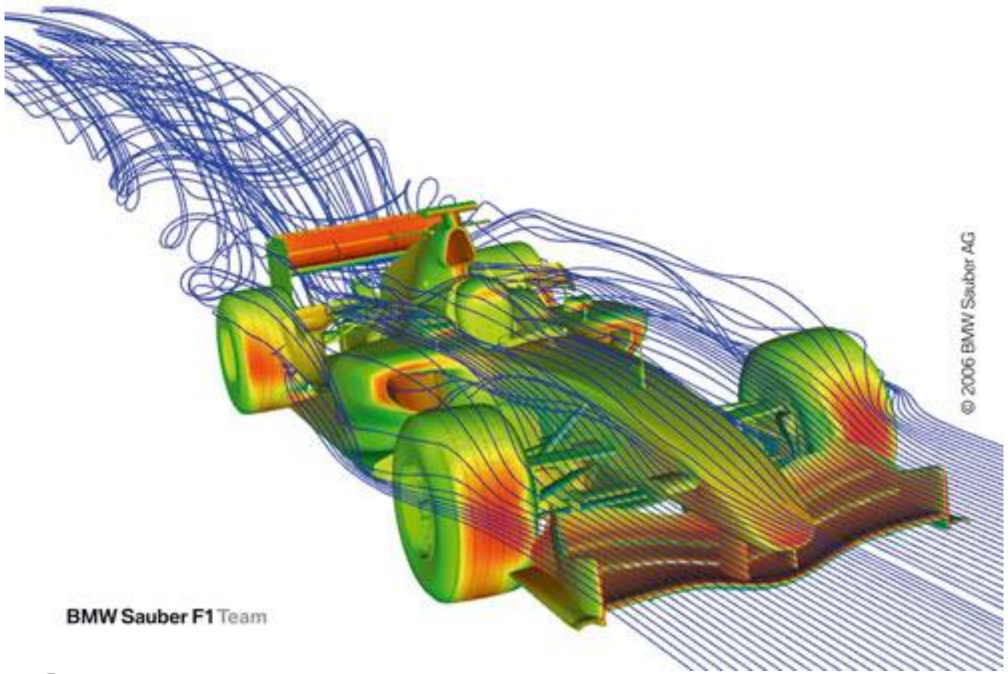
Cele mai multe aplicaţii tehnice sunt modelate cu ajutorul ecuaţiilor cu derivate parţiale- radiatorul unui procesor :





- curgerea aerului în jurul unui avion sau a unei maşini





Cea mai generală formă a unei ecuaţii cu derivate parţiale liniară de ordinul doi este:

$$a\frac{\partial^2 u}{\partial x^2}+b\frac{\partial^2 u}{\partial x\partial y}+c\frac{\partial^2 u}{\partial y^2}+d\frac{\partial u}{\partial x}+e\frac{\partial u}{\partial y}+fu=g(x,y),\quad (x,y)\in D\subset \mathbf{R}^2 \qquad (1)$$

unde $a, b, c, \ldots$ sunt constante, iar g este o funcţie cunoscută. În funcţie de coeficienţii părţii principale a operatorului (1):

$$a\frac{\partial^2 u}{\partial x^2}+b\frac{\partial^2 u}{\partial x\partial y}+c\frac{\partial^2 u}{\partial y^2}$$

putem clasifica ecuaţiile de tipul (1):

Eliptice: pentru $b^2-4ac<0$

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+Au=g(x,y) \qquad (2)$$

unde $A=0,\ \pm 1$.

Parabolice: pentru $b^2 - 4ac = 0$

$$\frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial y^2} = g(x, y) \qquad \text{(ecuaţia căldurii/difuziei)} \qquad (3)$$

Hieprbolice: pentru $b^2 - 4ac < 0$ şi se pot reduce la forma:

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} + Bu = g(x, y) \qquad (4)$$

unde $B = 0$ sau 1. Dacă $B = 0$ atunci se obţine ecuaţia undelor.

2. Ecuaţii parabolice cu o variabilă spaţială

Ecuaţia căldurii

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \ \ 0 \le x \le 1 \tag{5}$$

condiţie iniţiala

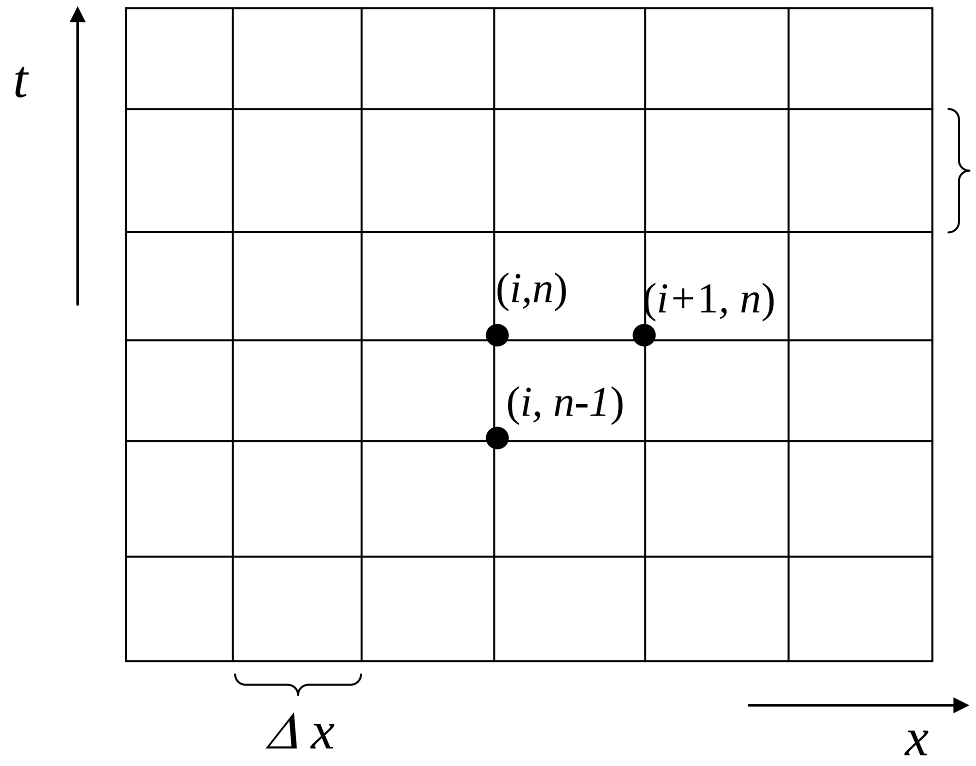
$$u(0, x) = u^0(x), \quad x \in [0,1] \tag{6}$$

condiţii pe frontieră:

$$u(t,0) = u(t,1) = 0, \ \ t > 0 \tag{7}$$

Scheme explicite

Vom aproxima soluţia problemei (5) – (7) folosind diferenţe finite. Pentru aceasta este necesară o divizarea domeniului $[0, t_f] \times [0,1]$ în subdomenii ce formează o reţea sau o grilă bidimensională.



Considerând o grilă echidistantă atunci nodurile sau punctele grilei se pot obţine astfel:

$$\left(x_i = i\Delta x,\ t_n = n\Delta t\right),\ \ i = 0,1,\ldots, N_x,\ \ n = 0,1,\ldots, N_t\Big)$$

Vom nota prin

$$U_i^n \approx u(x_i, t_n) \tag{8}$$

valoarea aproximativă a soluţiei problemei (5) – (7) în nodul (i,n). Aproximăm în continuare derivata temporală folosind diferenţe finite progresive, iar derivata spaţială folosind diferenţe finite centrale

$$\frac{\partial u}{\partial t}(x_i, t_n) \approx \frac{u(x_i, t_{n+1}) - u(x_i, t_n)}{\Delta t} \tag{9}$$

$$\frac{\partial^2 u}{\partial x^2}(x_i, t_n) \approx \frac{u(x_{i+1}, t_n) - 2u(x_i, t_n) + u(x_{i-1}, t_n)}{(\Delta x)^2} \tag{10}$$

apoi înlocuind în (3) şi ţinând cont de notaţia (8) obţinem:

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} = \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{(\Delta x)^2}$$

de unde putem exprima:

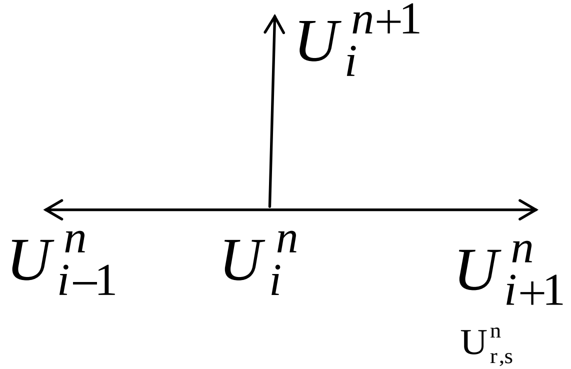
$$U_i^{n+1} = U_i^n + \nu\left(U_{i+1}^n - 2U_i^n + U_{i-1}^n\right), \quad \nu = \frac{\Delta t}{(\Delta x)^2} \tag{11}$$

Se observă că valoarea lui u de la pasul de timp t_{n+1} se calculează folosind doar valori de la pasul de timp t_n. În acest caz vom spune că avem o metodă explicită cu **diferenţe finite**. Folosind condiţia iniţială (6) şi condiţiile pe frontieră (7) care se vor scrie în cazul discret:

$$U_i^0 = u^0(x), \quad i = 1, 2, \ldots, N_x - 1 \tag{12}$$

$$U_0^n = U_{N_x}^n = 0, \quad n = 0, 1, 2, \ldots \tag{13}$$

putem calcula toate valorile interioare la paşii următori de timp.



Vom aplica schema explicită pentru condiţia iniţială

$$u^0(x)=\begin{cases} x, & x\in\left[0,\dfrac{1}{2}\right] \\ 1-x, & x\in\left[\dfrac{1}{2},1\right] \end{cases} \tag{14}$$

şi în acest cazul particular soluţia este

$$u(x,t)=\frac{4}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2}\sin\left(\frac{1}{2}k\pi\right)\sin(k\pi\, x)e^{-(k\pi)^2 t} \tag{15}$$

Rezolvăm problema dată de ecuaţiile (5), (14) şi (7) analitic folosind 100 de termeni in seria (20) şi numeric pentru $\Delta t=0.0013$ şi $\Delta x=0.05$.

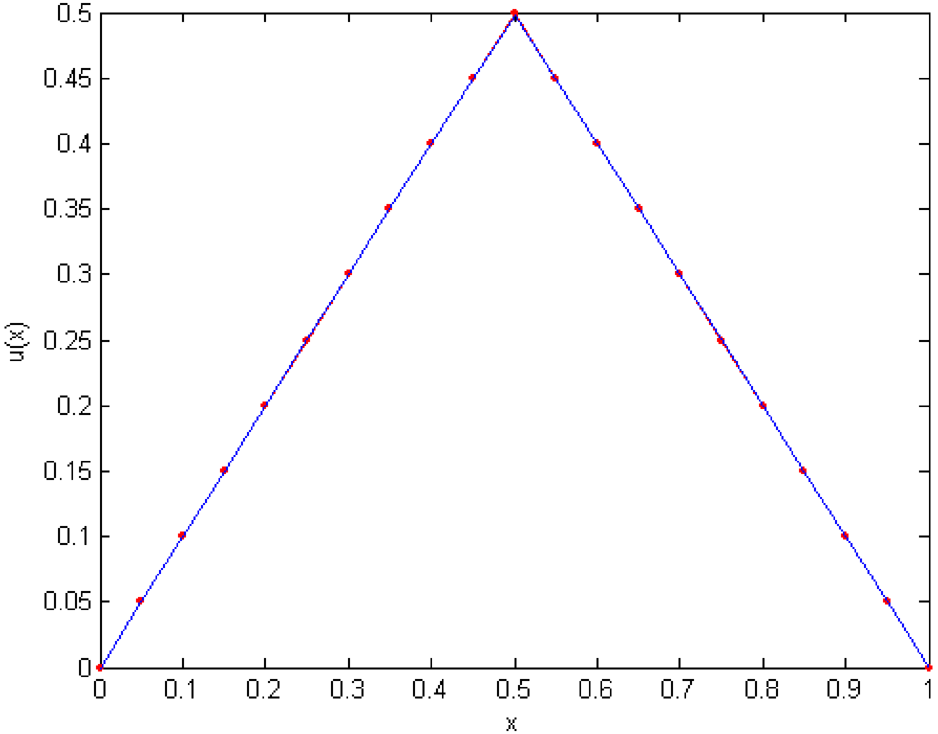
```
dt=0.0013; dx=0.05;
Nx=1/(dx)+1;
x=0:dx:1;
tf=0.1;
Nt=tf/dt;
Uo=zeros(Nx,1);
Un=zeros(Nx,1);
```

```
%initializare
for i=1:round(Nx/2)
    Uo(i)=(i-1)*dx;
end
for i=round(Nx/2):Nx
    Uo(i)=1-(i-1)*dx;
end
```

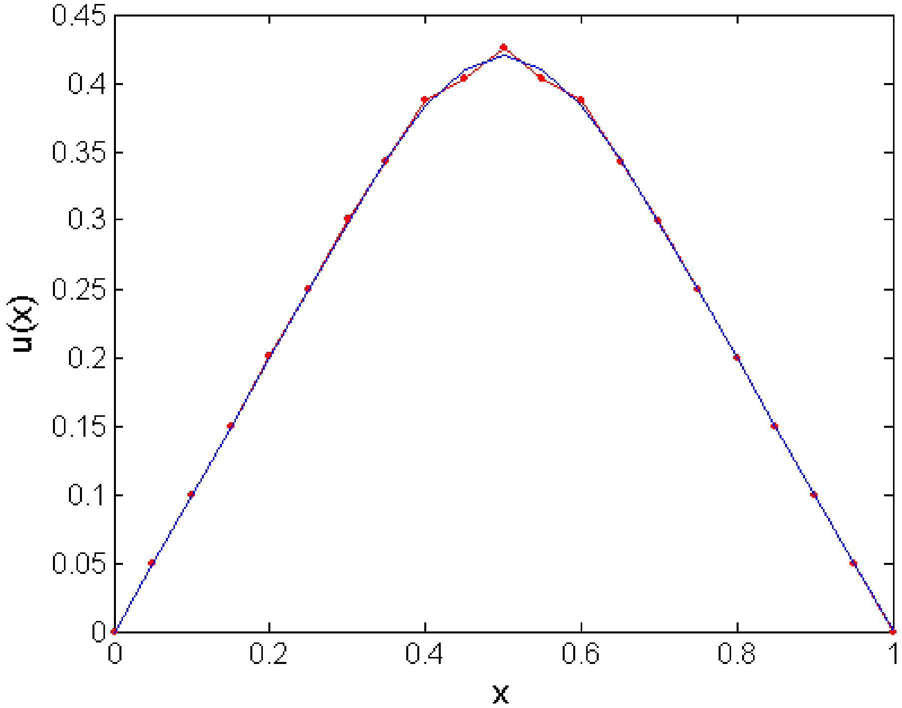
```
n=0;
while (n<Nt)
    n=n+1;
    for i=2:Nx-1
        Un(i)=Uo(i)+dt/(dx*dx)*(Uo(i-1)-2*Uo(i)+Uo(i+1));
    end
    Uo=Un;
end

plot(x,Uo,'.-r')
hold on
%solutia analitica
U=HeatAnalytic(x,tf);
plot(x,U,'b')
```

```
function rez=HeatAnalytic(x,t)
rez=0;
for k=1:100
    rez=rez+4/(k*pi)^2*sin(k*pi/2)*sin(k*pi*x)*exp(-k^2*pi^2*t);
end
```

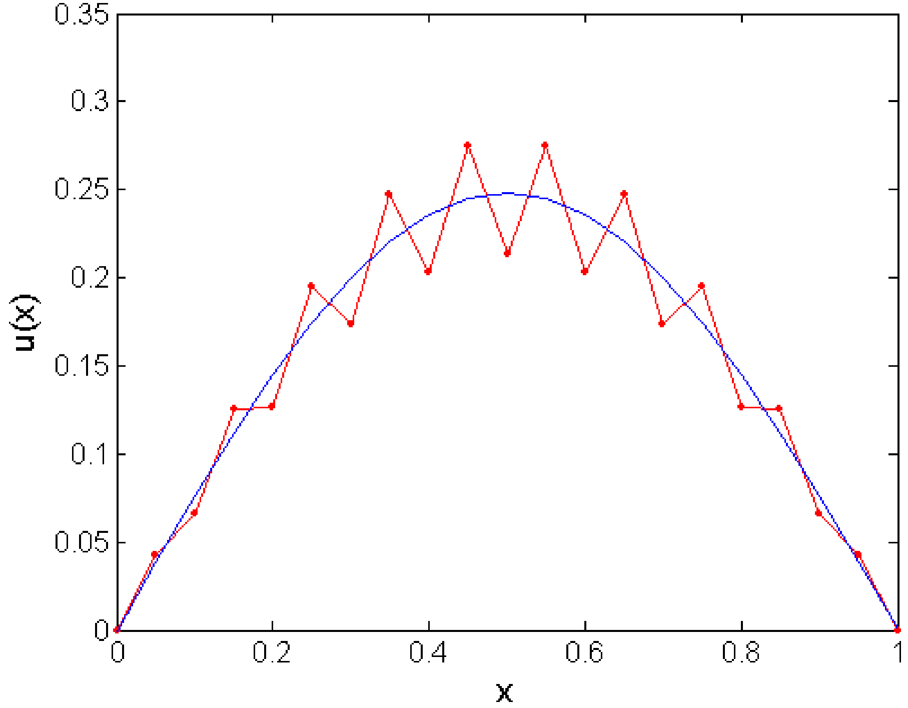
\



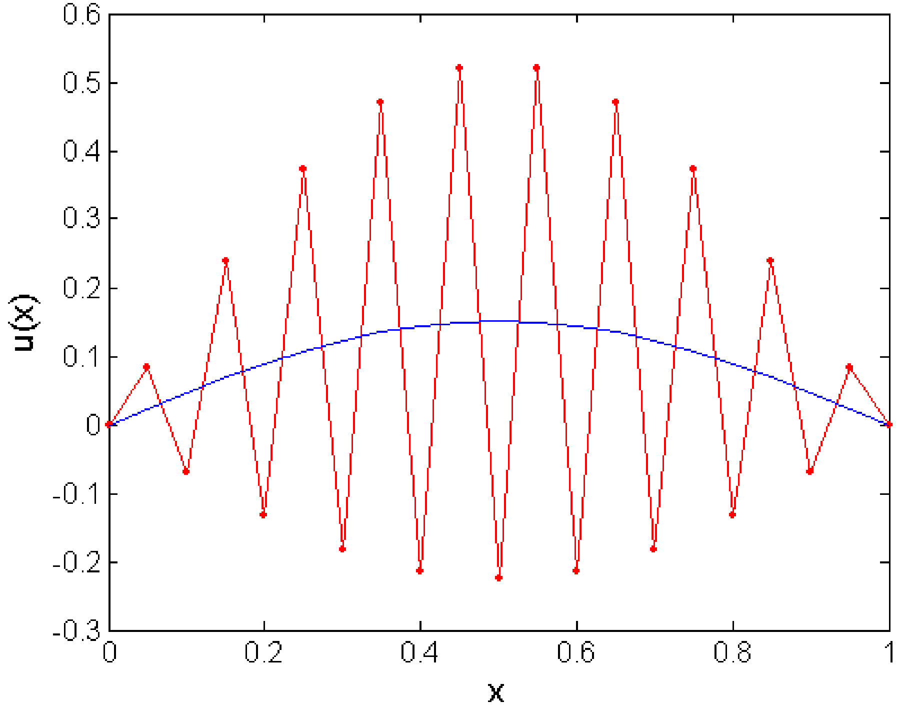
$t = 0$



$t = 0.005$



$t = 0.05$



$t = 0.1$

Se poate arăta că pentru convergentă e necesar ca:

$$\nu = \frac{\Delta t}{(\Delta x)^2} \le \frac{1}{2} \tag{16}$$

O schemă implicită

Restricţia (16) este o restricţie severă. Dacă dorim obţinerea unei aproximări bune avem nevoie de paşi mici în spaţiu ceea ce conduce la folosirea unui număr foarte mare de paşi în timp. Pentru a remedia acest impediment vom folosi în schema numerică diferenţe regresive în timp păstrând aproximarea cu diferenţe centrale a derivatei spaţiale. Atunci avem

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} = \frac{U_{i+1}^{n+1} - 2U_i^{n+1} + U_{i-1}^{n+1}}{(\Delta x)^2} \tag{17}$$

Se observă că pentru a calcula valoarea lui U de la pasul de timp t_{n+1} este necesar să înaintăm în timp pornind de la condiţia iniţială

$$U_i^0 = u^0(x), \quad i = 1, 2, \ldots, N_x - 1 \tag{18}$$

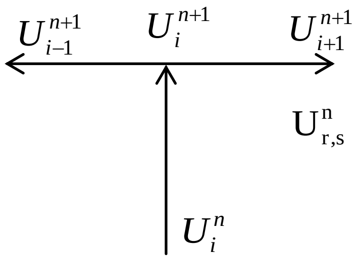
rezolvând un sistem de ecuaţii tridiagonal de forma:

$$-\nu U_{i-1}^{n+1} + (1 + 2\nu) U_i^{n+1} - \nu U_{i+1}^{n+1} = U_i^n \tag{19}$$

completat cu condiţiile la frontieră (13).

În acest caz vom spune că metoda este **implicită,** iar sistemul scris pe componente are

$$\begin{bmatrix} 1 & & & & & & & & \\ -\nu & 1+2\nu & -\nu & & & & & & \\ & \ddots & \ddots & \ddots & & & & & \\ & & -\nu & 1+2\nu & -\nu & & & & \\ & & & \ddots & \ddots & \ddots & & & \\ & & & & -\nu & 1+2\nu & -\nu & & \\ & & & & & & 1 & & \end{bmatrix} \times \begin{bmatrix} U_0 \\ U_1 \\ \dots \\ U_i \\ \dots \\ U_{Nx-1} \\ U_{Nx} \end{bmatrix}^{(n+1)} = \begin{bmatrix} 0 \\ U_1 \\ \dots \\ U_i \\ \dots \\ U_{Nx-1} \\ 0 \end{bmatrix}^{(n)}$$

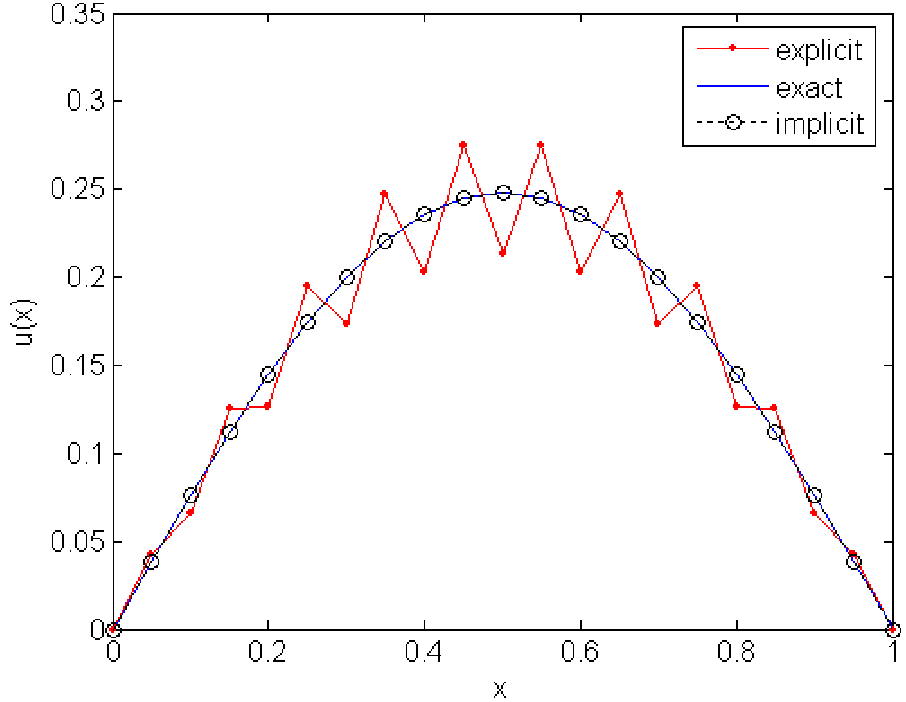


Reprezentare schematică a dependenţei temporale în schema implicită

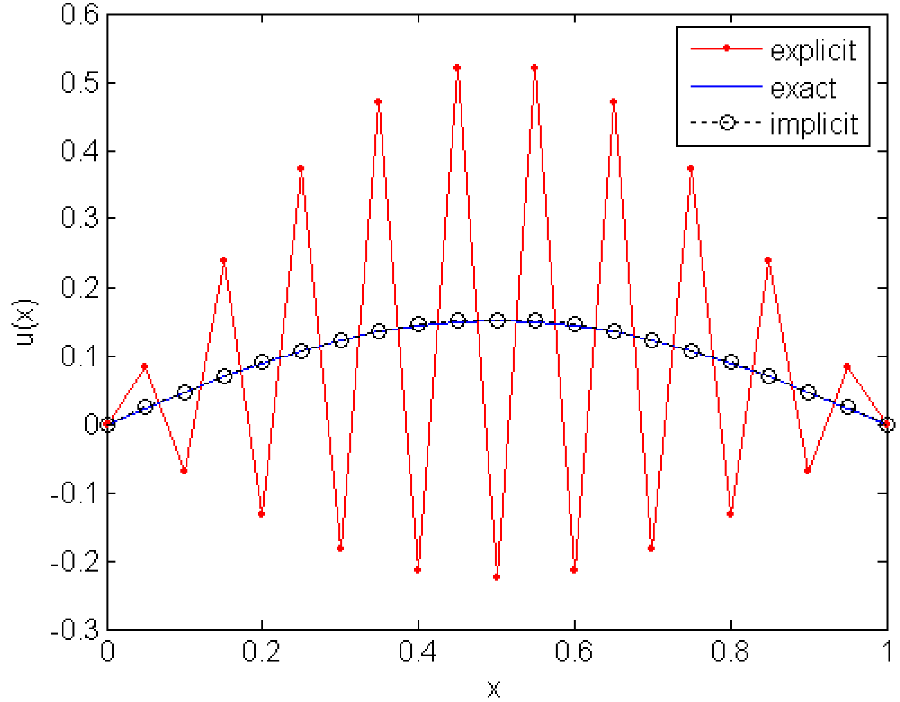
```
dt=0.0013;dx=0.05;
niu=dt/(dx*dx);
Nx=1/(dx)+1;
x=0:dx:1;
tf=0.1;
Nt=tf/dt;

Uo=zeros(Nx,1); Un=zeros(Nx,1);
```

```
%initializare
for i=1:round(Nx/2)
    Uo(i)=(i-1)*dx;
end
for i=round(Nx/2):Nx
    Uo(i)=1-(i-1)*dx;
end
A=zeros(Nx,Nx);b=zeros(Nx,1);
```

```
n=0;
while (n<Nt)
    n=n+1;
        A(1,1)=1;b(1)=0;
    for i=2:Nx-1
    A(i,i-1)=-niu;A(i,i)=1+2*niu;A(i,i+1)=-niu;b(i)=Uo(i);
    end
    A(Nx,Nx)=1;b(Nx)=0;
        Un=A\b;
    Uo=Un;
end

plot(x,Uo,'o:k')
hold on
```



$t = 0.05$



$t = 0.1$

Se poate arăta că metoda implicită este stabilă necondiţionat.

3. Ecuaţii eliptice

$$\left[\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right]= f(x, y), \quad (x, y)\in D\subset \mathbf{R}^2 \tag{20}$$

- condiţii pe frontieră Dirichlet: $u(x, y)\big|_{\partial D}$ sau Neumann: $\left.\frac{\partial u(x, y)}{\partial n}\right|_{\partial D}$

Schema cu diferenţe finite

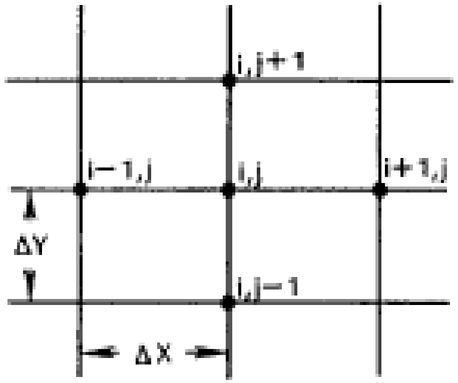
Considerăm reţeaua de puncte cu pasul Δx şi Δy în direcţiile Ox şi Oy, N_x şi N_y fiind numărul nodurilor în cele două direcţii. Notăm aproximaţia soluţiei iniţiale într-un punct (i, j) al reţelei cu

$$U_{i,j} \approx u\left(x_i, y_j\right), \quad i = 0,1,\ldots, N_x, \quad j = 0,1,\ldots, N_y.$$

Utilizând aproximarea cu diferenţe finite centrale pentru derivatele spaţiale obţinem o schemă cu 5 noduri:

$$\frac{U_{i+1,j}-2U_{i,j}+U_{i-1,j}}{\Delta x^2}+\frac{U_{i,j+1}-2U_{i,j}+U_{i,j-1}}{\Delta y^2}=f\left(x_i,y_j\right)\ , \tag{21}$$

$$i=2,\ 3,\ \ldots,N_x-1,\quad j=2,\ 3,\ \ldots,N_y-1$$



Nodurile implicate în formula schema (22)

Observăm că am obţinut un sistem de ecuaţii cu necunoscutele $U_{i,j}$, sistem ce se poate rezolva dacă se adaugă condiţiile pe frontieră.
Scriem în continuare sistemul (22) pentru f = 0 într-o formă simplificată:

$$U_{i+1,j} - 2U_{i,j} + U_{i-1,j} + \left(\frac{\Delta x}{\Delta y}\right)^2 \left(U_{i,j+1} - 2U_{i,j} + U_{i,j-1}\right) = 0 \ ,$$
$$i = 2, 3, ..., N_x - 1, \ \ j = 2, 3, ..., N_y - 1$$

sau

$$U_{i+1,j} + U_{i-1,j} + \beta^2 U_{i,j+1} + \beta^2 U_{i,j-1} - 2\left(1 + \beta^2\right) U_{i,j} = 0 \ , \quad i = 2, 3, ..., N_x - 1, \ \ j = 2, 3, ..., N_y - 1 \tag{23}$$

Exemple
Considerăm generarea de căldură într+un domeniu dreptunghiular:

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + 1 = 0, \quad (x, y) \in D = [0,1] \times [0,1]$$
$$T(x, y)\big|_{\partial D} = 0$$

care se discretizează în felul următor:

$$\frac{U_{i+1,j}-2U_{i,j}+U_{i-1,j}}{\Delta x^2}+\frac{U_{i,j+1}-2U_{i,j}+U_{i,j-1}}{\Delta y^2}+1=0\quad,$$

$$i=2,\ 3,\ \ldots,N_x-1,\quad j=2,\ 3,\ \ldots,N_y-1$$

Vom aplica în continuare metoda Jacobi şi metoda Gauss-Seidel. Programele Matlab sunt:

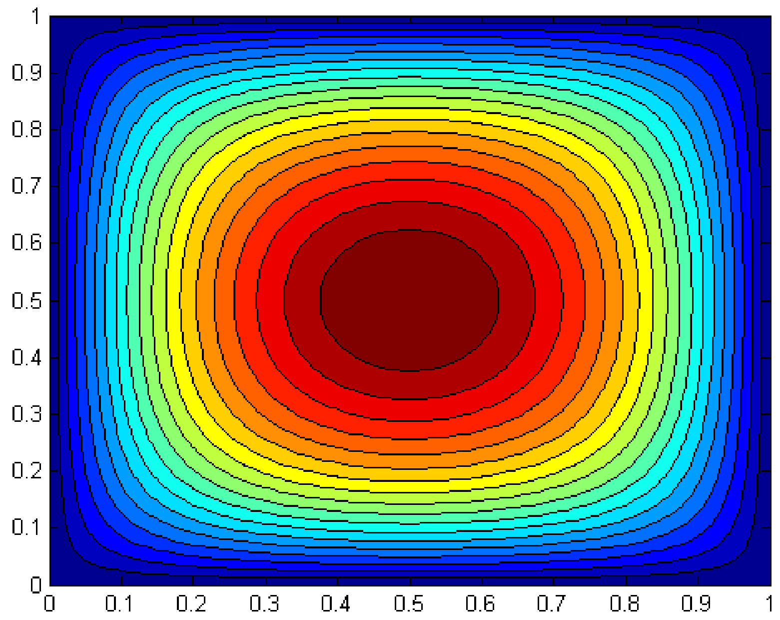
```
tic
N=101; %number of nodes in x- direction
   h=1/(N-1);
   T=zeros(N,N);
   Tnew=T
nr_it=0;
stop=0;
while (stop~=1)
    nr_it=nr_it+1; errT=0;
    for i=2:N-1
        for j=2:N-1
%SECVENTA JACOBI
           Tnew(i,j)=0.25*(T(i+1,j)+T(i-1,j)+T(i,j+1)+T(i,j-1)+h*h);
           if abs(T(i,j)-Tnew(i,j))>errT
```

```
                errT=abs(T(i,j)-Tnew(i,j));
           end
%SFARSIT SECVENTA JACOBI

%SECVENTA GAUSS-SEIDEL
            Tnew(i,j)=0.25*(T(i+1,j)+T(i-1,j)+T(i,j+1)+T(i,j-1)+h*h);
            if abs(T(i,j)-Tnew(i,j))>errT
                errT=abs(T(i,j)-Tnew(i,j));
            end

            T(i,j)=Tnew(i,j);
%SFARSIT SECVENTA GAUSS-SEIDEL
        end
    end;
    if errT<1e-6
             stop=1;
    end
    if mod(nr_it,250)==0
       fprintf('nr_it=%g err=%g\n', nr_it,errT));
    end
T=Tnew;
end
x=0:h:(N-1)*h;y=x;
contour(x,y,T',20)
axis equal
axis([0,1,0,1])
```

```
toc
```

Se observă că metoda Gauss-Seidel este mai rapidă:

Grid	Jacobi			Gauss-Seidel		
	T_{max}	Numărul iteraţiilor	Timp de calcul (secunde)	T_{max}	Numărul iteraţiilor	Timp de calcul (secunde)
51 x 51	0.073142	2577	0.625	0.073396	1465	0.422
101 x 101	0.071640	7502	6.969	0.0726535	4454	4.594



Distribuţia temperaturii

4. Ecuaţii hiperbolice într-o variabilă spaţială (Ecuaţia undelor)

Considerăm cea mai simplă ecuaţie de acest tip

$$\frac{\partial u}{\partial t} + a(x,t)\frac{\partial u}{\partial x} = 0, \qquad x \in [a,b], \; t \ge 0 \tag{24}$$

$$u(x,0) = u^0(x) \tag{25}$$

Condiţia lui Courant, Friedrichs şi Lewy (*CFL*)

Reamintim că pentru o ecuaţie de forma:

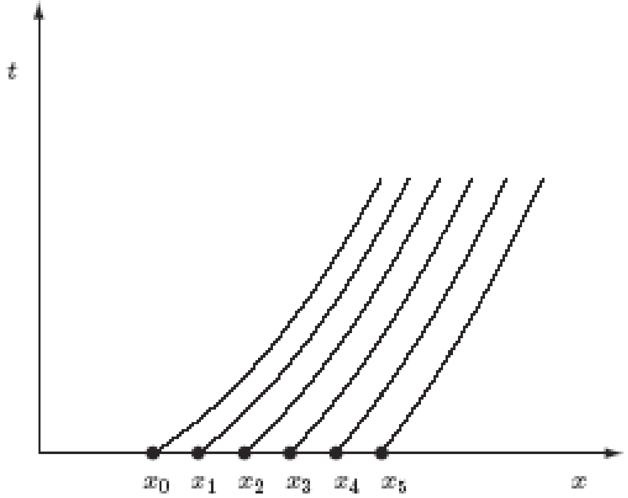
$$a(x,t)\frac{\partial u}{\partial t} + b(x,t)\frac{\partial u}{\partial x} = c(x,t)$$

caracteristicile sunt soluţiile ecuaţiei

$$a\,dy - b\,dt = 0$$

iar $u(x,t)$ se află din $c\,dt - a\,du = 0$

Dacă $a(x,t) \underset{notatie}{=} a = const.$ atunci caracteristicile sunt $x - at = const.$, iar soluţia ecuaţiei (24) are forma: $u(x,t) = u^0(x - at)$ (vezi figura)



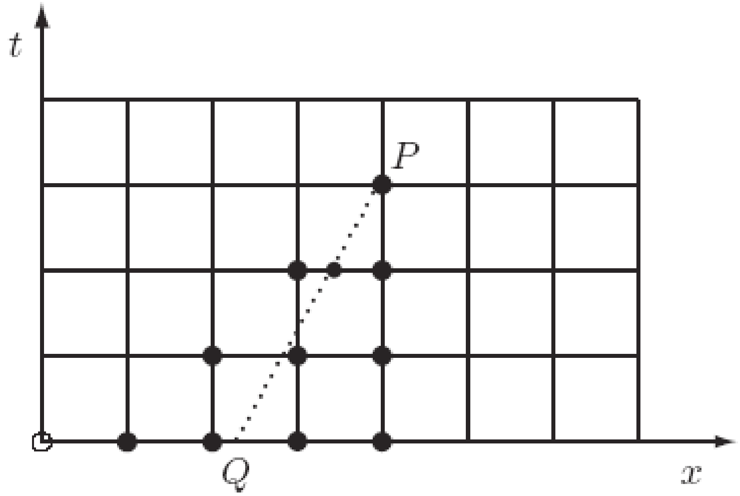
Pentru a rezolva numeric ecuaţia (25) putem utiliza o schemă explicită cu diferenţe finite progresive în timp şi regresive în spaţiu:

$$\frac{U_i^{n+1}-U_i^n}{\Delta t}+a\frac{U_i^n-U_{i-1}^n}{\Delta x}=0 \qquad (26)$$

sau

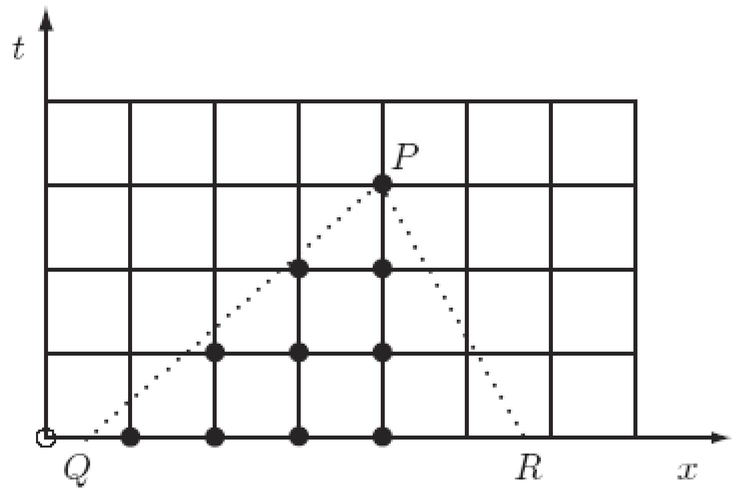
$$U_j^{n+1}=U_j^n-\frac{a\Delta t}{\Delta x}\left(U_i^n-U_{i-1}^n\right)=(1-\nu)U_i^n+\nu U_{i-1}^n \quad , \nu=\frac{a\Delta t}{\Delta x} \qquad (27)$$

Se observă că valoarea U_j^{n+1} depinde de două valori calculate la pasul de timp n, iar acestea la rândul lor depind de câte două valori calculate la pasul de timp $n-1$. Astfel se poate construi un domeniu de dependenţă a datelor pentru schema numerică, de formă triunghiulară, similar cu cel din figura



Domeniul de dependenţă corespunzător ecuaţiei cu derivate parţiale este chiar curba caracteristică ce trece prin punctul (x_i, t_n). **Condiţia CFL spune că o schemă numerică este convergentă dacă domeniul de dependenţă a ecuaţiei diferenţiale (caracteristica) aparţine domeniului de dependenţă al schemei numerice.**

Figura de mai jos prezintă cazul în care condiţia CFL este violată, ambele caracteristici PQ şi PRsituându-se în afara domeniului de dependenţă al schemei numerice, iar convergenţa în punctul P nu poate fi atinsă.



În cazul schemei (27) se observă că nu avem convergenţă pentru $a<0$ deoarece în acest caz caracteristicile sunt de forma PR. În cazul în care $a>0$ pentru a avea asigurată convergenţa este necesar ca $a\Delta t/\Delta x \le 1$.
Dacă vom folosi o discretizare cu diferenţe centrale pentru derivata spaţială a ecuaţiei (25) obţinem următoarea schemă cu diferenţe finite:

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + a\frac{U_{i+1}^n - U_{i-1}^n}{2\Delta x} = 0 \tag{28}$$

Alegând convenabil paşii Δt şi Δx, această schemă satisface condiţia CFL

$$c = |a|\frac{\Delta t}{\Delta x} \leq 1 \qquad \text{(numărul lui Courant)} \tag{29}$$

atât pentru valori negative ale lui a cât şi pentru valori pozitive.

Folosind diverse tipuri de discretizări se pot obţine mai multe scheme cu diferenţe finite. Una dintre schemele explicite cele mai eficiente este **metoda Lax-Wendorff**:

$$u_i^{n+1} = u_i^n - a\Delta t\left[\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}\right] + \frac{1}{2}a^2(\Delta t)^2\left[\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2}\right] \tag{30}$$

Schema (30) are ordinul de exactitate doi şi este stabilă pentru $c \leq 1$.

Exemplu: Considerăm ecuaţia:

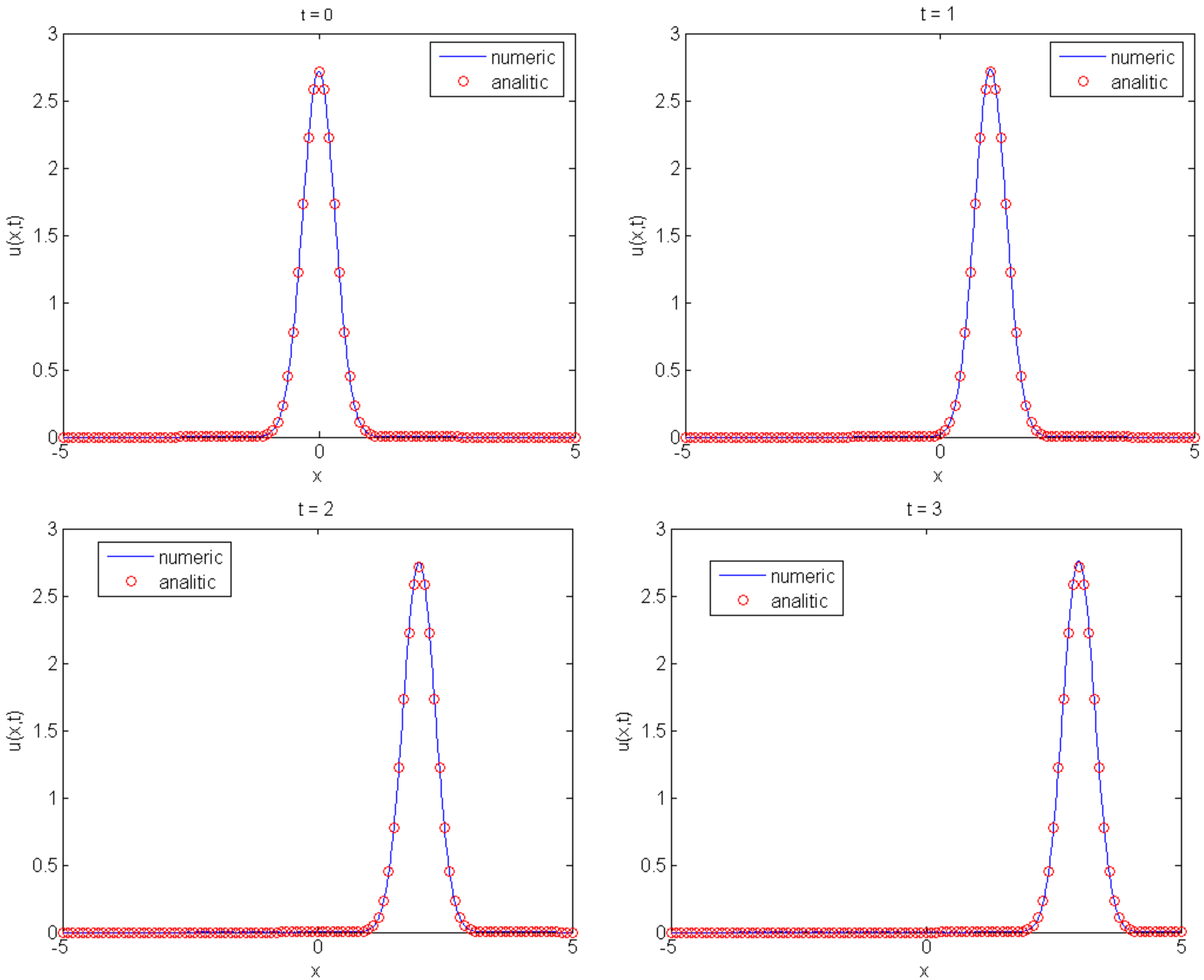
$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x \in [-5,5], \; t \geq 0$$

$$u(x,0) = \exp(1 - 5x^2)$$

care are soluţia analitică $u(x,t) = \exp(1 - 5(x-t)^2)$. Vom rezolva ecuaţia numeric folosind schema Lax-Wendorff.

```
a=-5;
b=5;
dt=0.001;
dx=0.01;
x=a:dx:b;
xa=a:10*dx:b;
N=length(x);
uo=exp(1-5*x.^2);%conditia initiala
tf=3;
t=0;
uex=exp(1-5*(xa-t*ones(1,length(xa))).^2);%solutia exacta
nr_it=0;
plot(x,uo,'b',xa,uex,'or')
pause
k=1;
M(k)=getframe;
```

```
while t<tf
 t=t+dt;
 nr_it=nr_it+1;

 un(1)=uo(1)-dt/dx*(uo(2)-uo(1));
 for i=2:N-1
   un(i)=uo(i)-dt/dx/2*(uo(i+1)-uo(i-1))+0.5*dt*dt/dx/dx*(uo(i+1)-2*uo(i)+uo(i-1));
 end
 un(N)=uo(N)-dt/dx*(uo(N)-uo(N-1));

 if mod(nr_it,10)==0
   k=k+1;
   uex=exp(1-5*(xa-t*ones(1,length(xa))).^2);
   plot(x,un,'b',xa,uex,'or');
   M(k)=getframe;
 end
uo=un;
end
movie(M)
```


t = 0
t = 1
t = 2
t = 3
numeric
analitic
u(x,t)
x
3
2.5
2
1.5
1
0.5
0
-5
0
5

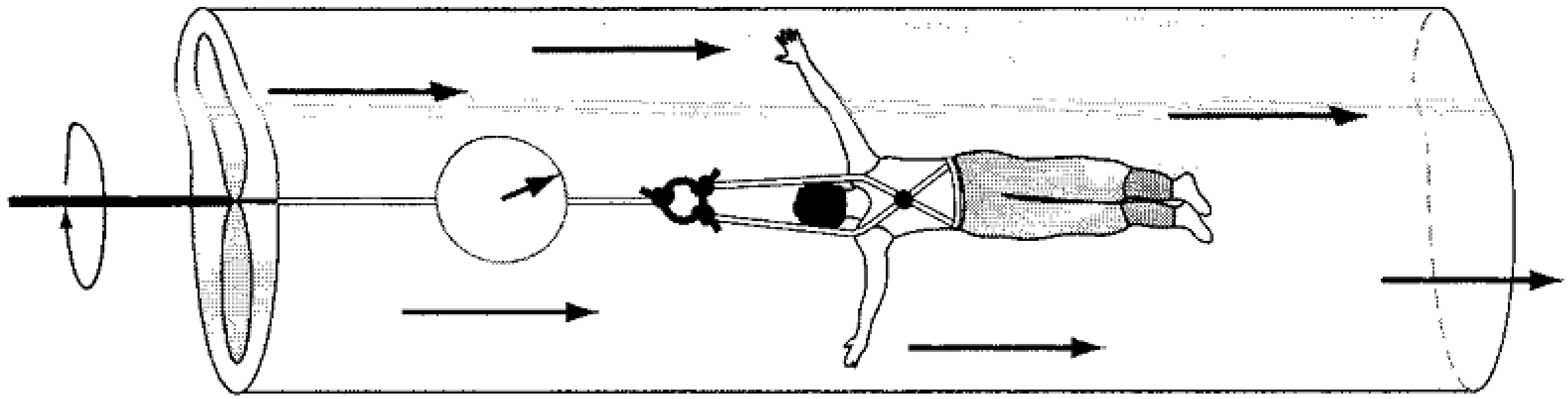
Regresii liniare

1. Noţiuni teoretice introductive

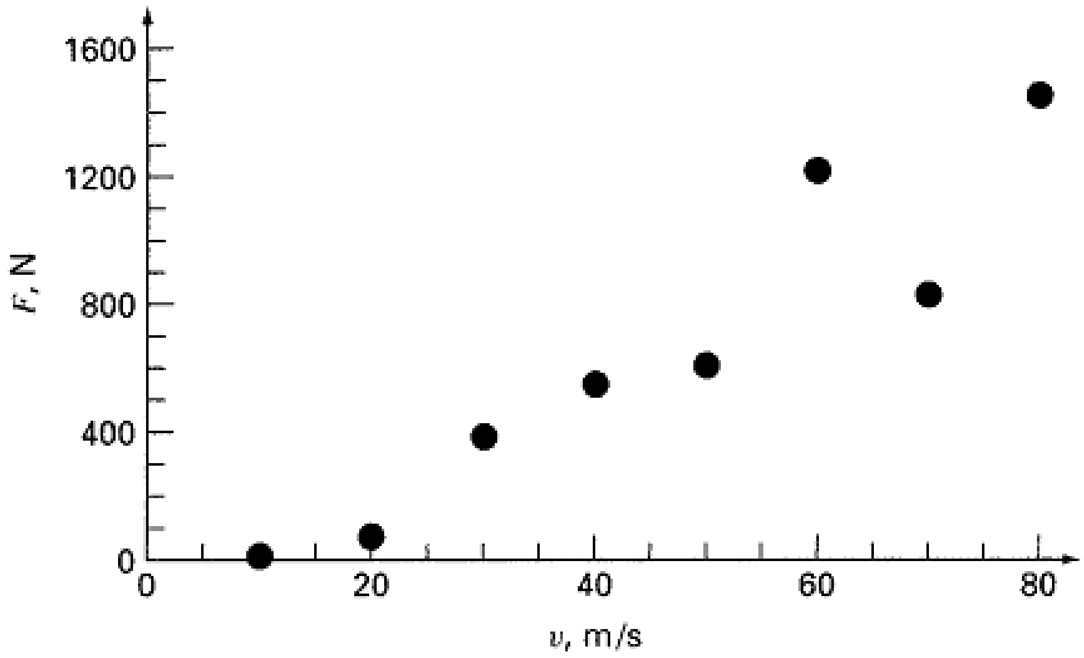
Se ştie ca teoretic, forţa de rezistenţă ce o întampină un obiect la mişcarea prin aer este:

$$F_U = c_d v^2$$

where F_U = the upward force of air resistance [N = kg m/s^2], c_d = a drag coefficient (kg/m), and v = velocity [m/s].



v, m/s	10	20	30	40	50	60	70	80
F, N	25	70	380	550	610	1220	830	1450



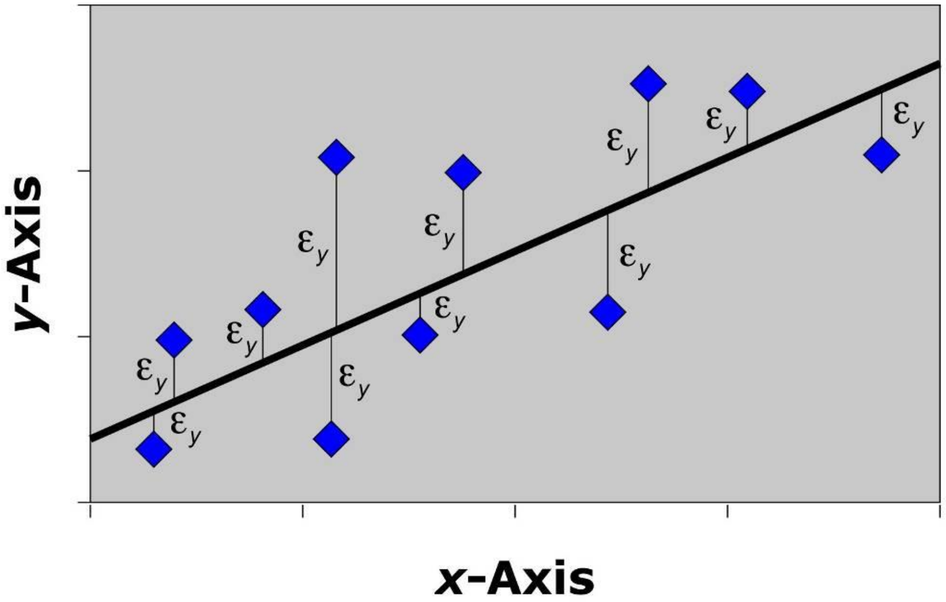
Se pune problema gasirii unei curbe ce aproximează cât mai bine datele obţinute experimental („norul de puncte”).

Metoda celor mai mici pătrate.

Fie curba $y=f(x) = ax+b$ care aproximează norul de puncte. Se formează suma:

$$S(a,b) = \sum_{i=1}^{n} [f(x_i) - y_i]^2$$

reprezentând suma pătratelor distanţelor de la punctele experimentale la punctele curbei y = f(x).



Dorim sa minimizăm pe S(a,b)

Calculăm derivatele parţiale ale lui S în raport cu a şi b şi determinăm extremul funcţiei S(a, b) din sistemul de ecuaţii:

$$\begin{cases} \dfrac{S(a,b)}{\partial a} = 0 \\ \dfrac{S(a,b)}{\partial b} = 0 \end{cases}$$

Verificăm dacă valorile determinate (a, b) reprezintă într-adevăr un minim pentru funcţia S. Se verifică inegalităţile:

$\Delta > 0$;

$r > 0$.

Cu a şi b determinate trasăm drepta de ecuaţie $y=ax+b$ care va trece “prin interiorul” norului de puncte astfel încât distanţa de la aceste puncte la dreptă să fie minimă.

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

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

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

Obţinem

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

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

unde necunoscutele sunt coeficienţii *a* şi *b*. Avem

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

Pentru exemplul de mai sus căutăm o curbă de forma:

$$y = a_0 + a_1 x$$

i	x_i	y_i	x_i^2	$x_i y_i$
1	10	25	100	250
2	20	70	400	1,400
3	30	380	900	11,400
4	40	550	1,600	22,000
5	50	610	2,500	30,500
6	60	1,220	3,600	73,200
7	70	830	4,900	58,100
8	80	1,450	6,400	116,000
$\sum$	360	5,135	20,400	312,850

The means can be computed as

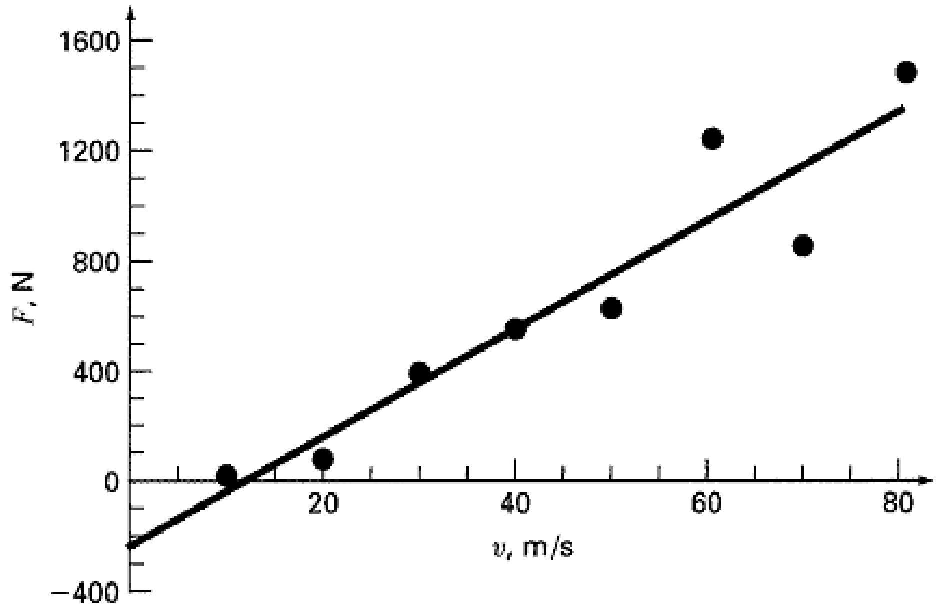
$$\bar{x} = \frac{360}{8} = 45 \qquad \bar{y} = \frac{5,135}{8} = 641.875$$

$$a_1 = \frac{8(312{,}850) - 360(5{,}135)}{8(20{,}400) - (360)^2} = 19.47024$$

$$a_0 = 641.875 - 19.47024(45) = -234.2857$$

Using force and velocity in place of y and x, the least-squares fit is

$$F = -234.2857 + 19.47024v$$



Observaţie: Rezultatele, cel putin pentru viteze mici, nu sunt corecte deoarece avem valori negative ale forţei de rezistenţă.

Pentru a verifica „cât de bună” este aproximarea noastră introducem mărimile:

$$S_r = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)^2$$

$$S_t = \sum (y_i - \bar{y})^2$$ unde $$\bar{y} = \frac{\sum y_i}{n}$$ este media valorilor.

Se calculează eroarea standard a estimaţiei

$$s_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

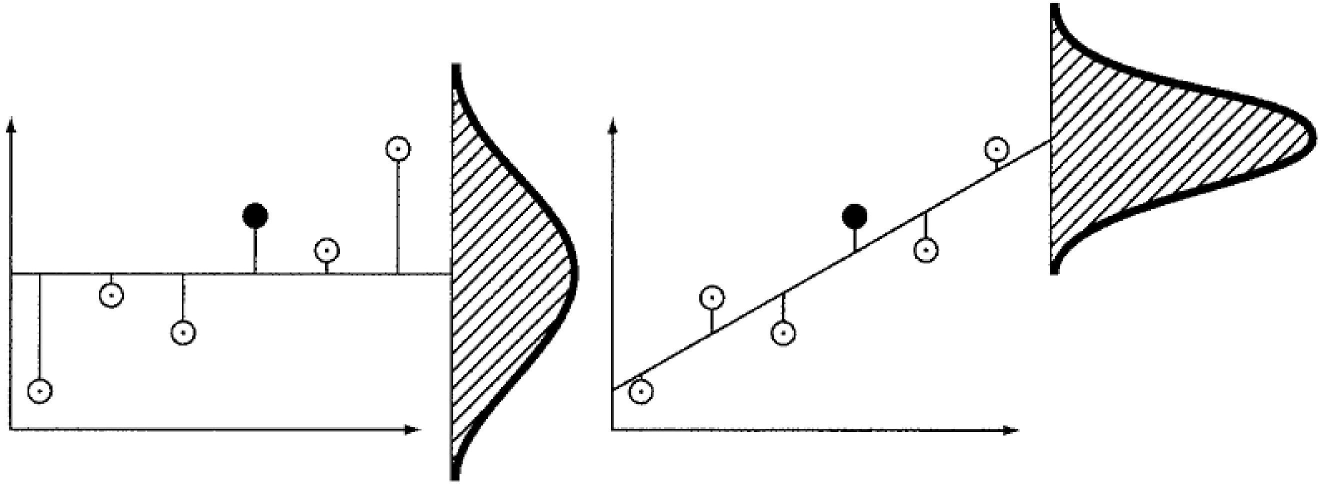
unde notaţia y/x semnifică faptul că eroarea se referă la o valoare preconizată a lui y corespunzând unei valori particulare a lui x. Numitorul n-2 semnifică faptul ca s-au pierdut două grade de libertate pentru calculul valorii lui S_r (prin determinarea coeficienţilor a_0 şi a_1).

Reamintim că abaterea medie pătratică dată de

$$s_y = \sqrt{\frac{S_t}{n-1}}$$

măsoară dispersia datelor.
Putem face o analogie între abaterea medie pătratică şi eroarea standard a estimaţiei:

Regression data showing (*a*) the spread of the data around the mean of the dependent variable and (*b*) the spread of the data around the best-fit line. The reduction in the spread in going from (*a*) to (*b*), as indicated by the bell-shaped curves at the right, represents the improvement due to linear regression.



(*a*) (*b*)

Parametrul care ne indică „cât de bună” este aproximarea noastră este *coeficientul de determinare*

$$r^2 = \frac{S_t - S_r}{S_t}$$

unde

$$r = \frac{n\sum(x_i y_i) - \left(\sum x_i\right)\left(\sum y_i\right)}{\sqrt{n\sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{n\sum y_i^2 - \left(\sum y_i\right)^2}}$$

este *coeficientul de corelaţie*. Cu cât coeficientul de determinare este mai aproape de 1, aproximarea noastră este mai bună.

Pentru exemplul de mai sus avem

$$s_y = \sqrt{\frac{1,808,297}{8-1}} = 508.26 \quad > \quad s_{y/x} = \sqrt{\frac{216,118}{8-2}} = 189.79$$

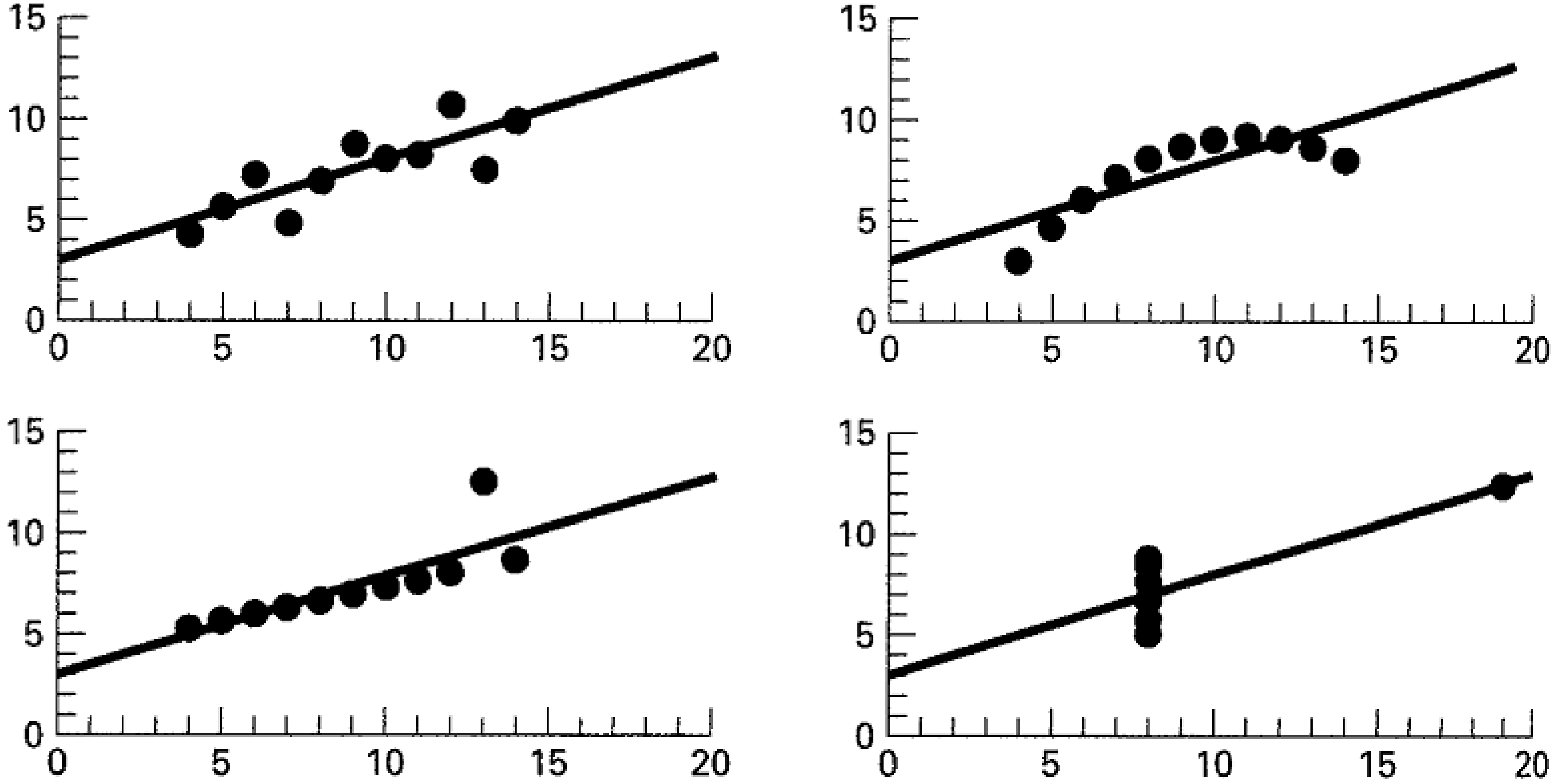
şi atunci regresia liniară aproximează corect datele.

i	x_i	y_i	$a_0 + a_1x_i$	$(y_i - \bar{y})^2$	$(y_i - a_0 - a_1x_i)^2$
1	10	25	−39.58	380,535	4,171
2	20	70	155.12	327,041	7,245
3	30	380	349.82	68,579	911
4	40	550	544.52	8,441	30
5	50	610	739.23	1,016	16,699
6	60	1,220	933.93	334,229	81,837
7	70	830	1,128.63	35,391	89,180
8	80	1,450	1,323.33	653,066	16,044
$\sum$	360	5,135		1,808,297	216,118

$$r^2 = \frac{1{,}808{,}297 - 216{,}118}{1{,}808{,}297} = 0.8805$$

Putem spune că 88% din incertitudinile initiale sunt sunt explicate de acest model liniar.

Totuşi nu ne putem baza doar pe calculul coeficientului de determinare, curba obţinută trebuie verificată şi vizual. În exmplele de mai jos, toate datele sunt aproximate cu aceeaşi dreaptă y =3 + 0.5x şi au acelaşi coeficient de determinare, $r^2 = 0.674$.


15
10
5
0
0 5 10 15 20
15
10
5
0
0 5 10 15 20
15
10
5
0
0 5 10 15 20
15
10
5
0
0 5 10 15 20