

Curs # 9

V. Minimizarea funcțiilor de două variabile.

V. 1. Metoda pasului descendente.

Fie $f: D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ o funcție de două variabile

$$z = f(x, y), (x, y) \in D$$

Def. Se numește linie de nivel curba $z = c$, $c = \text{const}$

$$\mathcal{C}: f(x, y) = c$$

Exemplu 1 Fie $z = x^2 + y^2$ și $A(3, 2)$

a) să se afle curba de nivel care trece prin $A(3, 2)$

b) să se afle curbile de nivel $z=1, z=4, z=9$

a) $\mathcal{C}_A : x^2 + y^2 = c$, cu $A \in \mathcal{C}_A \Rightarrow 3^2 + 2^2 = c \Rightarrow c = 13$

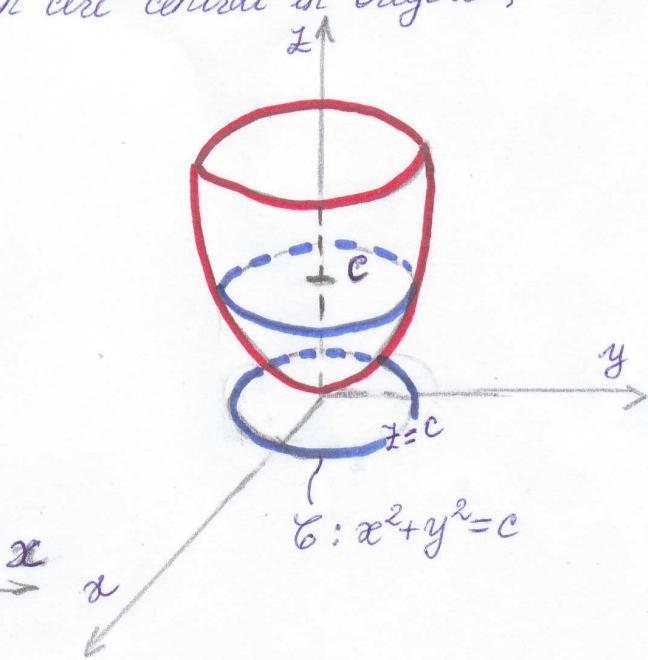
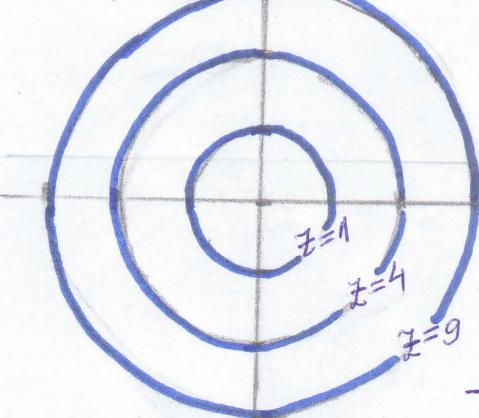
Astfel $\mathcal{C}_A : x^2 + y^2 = \sqrt{13}^2$

Obs. curba \mathcal{C}_A reprezintă un cerc centrat în origine și de rază $\sqrt{13}$

b) $\mathcal{C}_{z=1} : x^2 + y^2 = 1$

$\mathcal{C}_{z=4} : x^2 + y^2 = 2^2$

$\mathcal{C}_{z=9} : x^2 + y^2 = 3^2$



Obs. curba de nivel se obține la intersecția planului $z=c$ cu suprafața $z=f(x, y)$

Def. Definim $\nabla f(x, y)$ vectorul format din derivatele parțiale

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right)^T$$

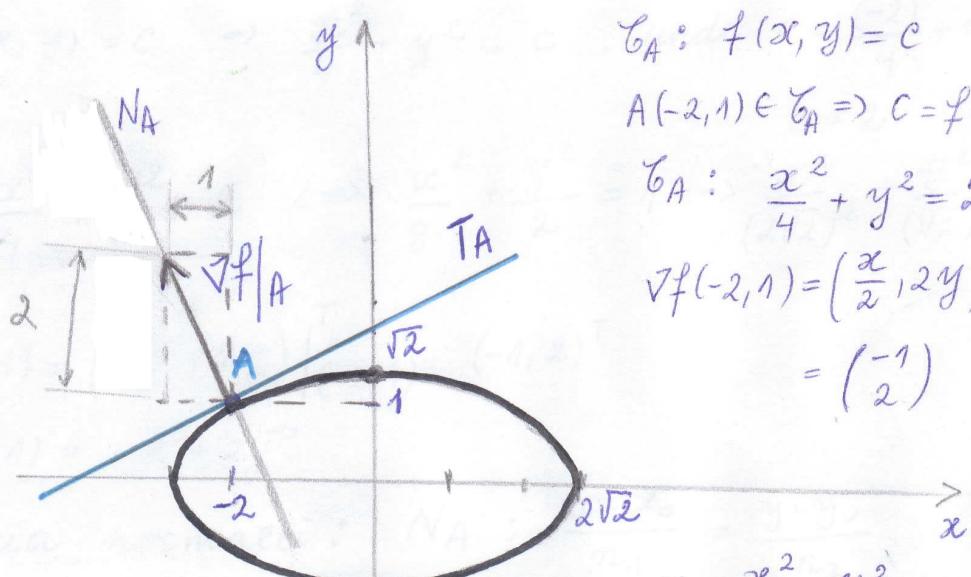
și îl numim gradientul funcției f în punctul (x, y)

Exemplu 2. Fie $f(x, y) = x^2 + y^2$, să se afle $\nabla f(3, 2)$

$$\nabla f(x, y) = (2x, 2y)^T$$

$$\nabla f(3, 2) = (6, 4)^T$$

Exemplu 3. Fie $f(x, y) = \frac{x^2}{4} + y^2$ și C_A curba de nivel a funcție f care trece prin punctul $A(-2, 1)$. Să se calculeze $\nabla f(-2, 1)$, ecuația tangentei la curbă în punctul A și ecuația normalăi la curbă în A .



$$C_A : f(x, y) = c$$

$$A(-2, 1) \in C_A \Rightarrow c = f(-2, 1)$$

$$C_A : \frac{x^2}{4} + y^2 = 2$$

$$\nabla f(-2, 1) = \left(\frac{x}{2}, 2y \right)^T \Big|_{(-2, 1)} = \left(\begin{matrix} -1 \\ 2 \end{matrix} \right)$$

$$C : \frac{x^2}{8} + \frac{y^2}{2} = 1$$

$$C : \frac{x^2}{(2\sqrt{2})^2} + \frac{y^2}{(\sqrt{2})^2} = 1$$

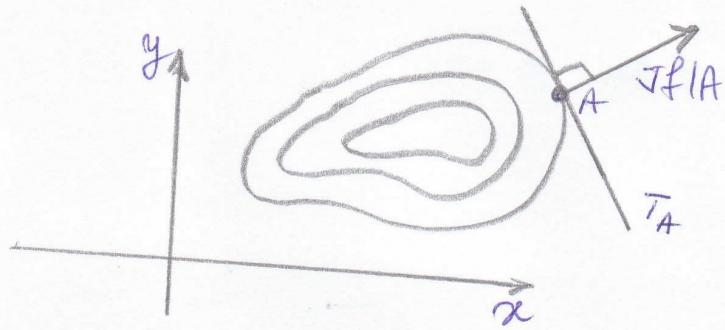
$$N_A : \frac{x - x_0}{n_1} = \frac{y - y_0}{n_2}, \vec{n} = \nabla f|_A$$

$$N_A : \frac{x+2}{-1} = \frac{y-1}{2} \Rightarrow N_A : y-1 = -2 \cdot (x+2)$$

Intrucat $\nabla f|_A \perp T_A \Rightarrow T_A : n_1 \cdot (x-x_0) + n_2 \cdot (y-y_0) = 0$, ceea ce $\vec{n} = \nabla f|_A$,

deci $T_A : -1 \cdot (x+2) + 2 \cdot (y-1) = 0$ sau $T_A : y = \frac{1}{2}x + 2$

Obs. Vectorul $\nabla f|_A$ este normal la curba de nivel care trece prin punctul A.



Def. (Derivata după o direcție)

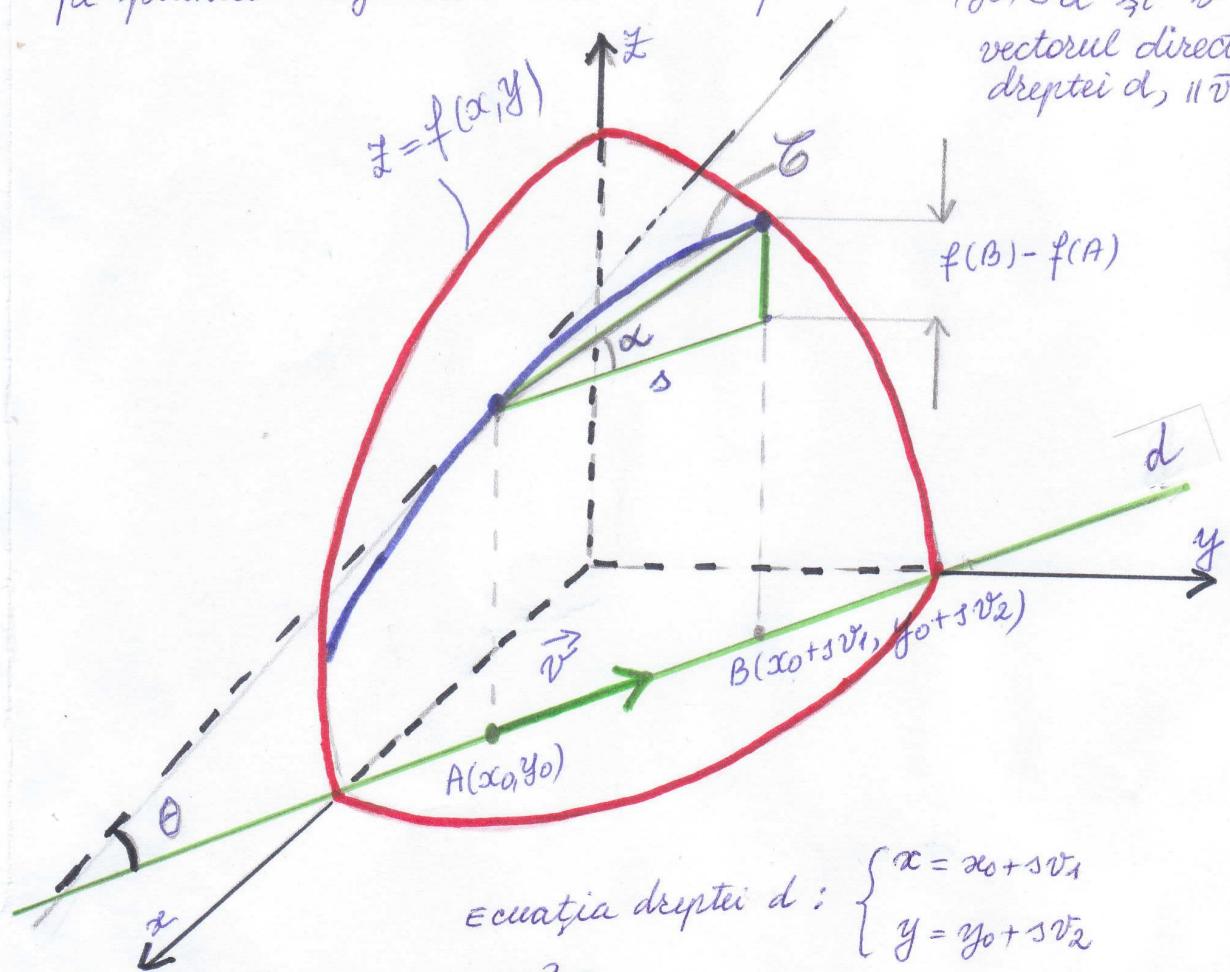
Fie $\vec{v} = v_1 \vec{i} + v_2 \vec{j}$, $\|\vec{v}\|=1$. Numărul $D_{\vec{v}} f|_A$ definit prin

$$D_{\vec{v}} f|_A := \lim_{s \rightarrow 0} \frac{f(x_0 + sv_1, y_0 + sv_2) - f(x_0, y_0)}{s}, \quad A(x_0, y_0)$$

S.n. derivata după o direcție în punctul A.

Fie o curbă C de pe suprafața funcției f și dreapta d
proiecția acesteia pe planul xy. Considerăm un punct A(x_0, y_0) ∈ d și \vec{v}

vectorul director al dreptei d, $\|\vec{v}\|=1$.



$$\operatorname{tg} \alpha = \frac{f(B) - f(A)}{d(A, B)}$$

$$d(A, B) = \sqrt{(x_0 + sv_1 - x_0)^2 + (y_0 + sv_2 - y_0)^2} = s \sqrt{v_1^2 + v_2^2} = s \cdot \underbrace{\|\vec{v}\|}_1 = s$$

Prin urmare: $\operatorname{tg} \alpha = \frac{f(x_0 + sv_1, y_0 + sv_2) - f(x_0, y_0)}{s}$

Încănd la limită după $s \rightarrow 0$ se obține

$$\boxed{\operatorname{tg} \theta = \lim_{s \rightarrow 0} \frac{f(x_0 + sv_1, y_0 + sv_2) - f(x_0, y_0)}{s}}$$

deci $\operatorname{tg} \theta = D_v f|_A$

cu alte cuvinte, $D_v f|_A$ reprezintă panta tangentei la curba dată parametric $\gamma: \begin{cases} x = x_0 + sv_1 \\ y = y_0 + sv_2 \\ z = f(x_0 + sv_1, y_0 + sv_2) \end{cases}$

Prop. Derivata după o direcție se calculează conform formulei

$$\boxed{D_v f = \nabla f \cdot \vec{v}}$$

Exemplu 4 Să se afle $D_v f|_{(3,2)}$ pentru $f(x, y) = x^2 + y^2$, $\vec{v} = -\vec{i} + \vec{j}$

$$\begin{aligned} D_v f|_{(3,2)} &= \nabla f|_{(3,2)} \cdot \vec{v} = (2x, 2y)|_{(3,2)} \cdot (-1, 1)^T = \\ &= \left\langle \begin{pmatrix} 6 \\ 4 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle = -6 + 4 = -2 \end{aligned}$$

Se pune următoarea problemă: să se afle direcția după care derivata $D_v f$ să fie maximă (minimă)

$$\begin{aligned} D_v f &= \nabla f \cdot \vec{v} = \|\nabla f\| \cdot \underbrace{\|\vec{v}\|}_1 \cdot \cos \varphi (\nabla f, \vec{v}) = \\ &= \|\nabla f\| \cdot \cos \varphi (\nabla f, \vec{v}) \end{aligned}$$

Aleg \vec{v} a.i. $D_{\vec{v}}f$ să fie maximă (resp. minimă).

Maximul (sau minimul) se atinge dacă $\cos(\nabla f, \vec{v}) = \pm 1$

- $D_{\vec{v}}f$ maximă $\Leftrightarrow m(\nabla f, \vec{v}) = 1 \Leftrightarrow \vec{v} = \frac{\nabla f}{\|\nabla f\|}$ ($\vec{v} \parallel \nabla f$)

- $D_{\vec{v}}f$ minimă $\Leftrightarrow m(\nabla f, \vec{v}) = -1 \Leftrightarrow \vec{v} = -\frac{\nabla f}{\|\nabla f\|}$

În acest caz $D_{\vec{v}}f = \pm \|\nabla f\|$

Exemplu 5 Fie $f(x, y) = x^2 + y^2$, $A(3, 2)$. Să se afle direcția \vec{v} după care $D_{\vec{v}}f|_A$ este maximă sau minimă

$$\nabla f = (2x, 2y)^T$$

$$\nabla f|_{(3,2)} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

$$\vec{v} = \frac{\nabla f|_{(3,2)}}{\|\nabla f|_{(3,2)}\|} = \frac{(6, 4)^T}{\|(6, 4)^T\|} = \frac{(6, 4)^T}{\sqrt{36+16}} = \frac{1}{\sqrt{52}} \cdot \begin{pmatrix} 6 \\ 4 \end{pmatrix} = \frac{1}{2\sqrt{13}} \cdot \begin{pmatrix} 6 \\ 4 \end{pmatrix} = \frac{1}{\sqrt{13}} \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$$D_{\vec{v}}f|_A = \|\nabla f|_{(3,2)}\| = \sqrt{36+16} = 2\sqrt{13} \text{ (maximă)}$$

$$\vec{v} = -\frac{1}{\sqrt{13}} \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$$D_{\vec{v}}f|_A = -\|\nabla f|_{(3,2)}\| = -2\sqrt{13}.$$

Obs. Orice altă direcție s-ar alege, derivata după acea direcție nu va depăși valoarea maximă $2\sqrt{13}$.

Obs. spunem că funcția $z = f(x, y)$ se schimbă cel mai rapid pe direcția gradientului.

Fie $A \in M_2(\mathbb{R})$ simetrică și pozitiv definită și $b \in \mathbb{R}^2$

Fie forma pătratică

$$f(x, y) = \frac{1}{2} \langle A \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle - \left\langle \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle$$

Funcția f scrisă desfășurată are forma

$$f(x, y) = \frac{1}{2}a_{11}x^2 + a_{12} \cdot x \cdot y + \frac{1}{2}a_{22}y^2 - b_1x - b_2y$$

Intrucăt A este simetrică și poz. definită, funcția admite un punct de minim. Fie acesta (x^*, y^*) . Punctul (x^*, y^*) este punct critic pentru f , i.e. $\frac{\partial f}{\partial x}(x^*, y^*) = 0$, $\frac{\partial f}{\partial y}(x^*, y^*) = 0$ sau $\nabla f|(x^*, y^*) = 0$. Aflăm în continuare derivatele parțiale:

$$\begin{cases} \frac{\partial f}{\partial x} = a_{11} \cdot x + a_{12} y - b_1 \\ \frac{\partial f}{\partial y} = a_{12} x + a_{22} \cdot y - b_2 \end{cases}$$

$$\text{deci } \begin{cases} a_{11}x^* + a_{12}y^* - b_1 = 0 \\ a_{12}x^* + a_{22}y^* - b_2 = 0 \end{cases} \Leftrightarrow \begin{cases} a_{11}x^* + a_{12}y^* = b_1 \\ a_{12}x^* + a_{22}y^* = b_2 \end{cases} \Rightarrow A \begin{pmatrix} x^* \\ y^* \end{pmatrix} = b$$

Se observă că punctul de minim este soluție pentru sistemul $A \cdot \begin{pmatrix} x \\ y \end{pmatrix} = b$.

Metoda pasului descendente

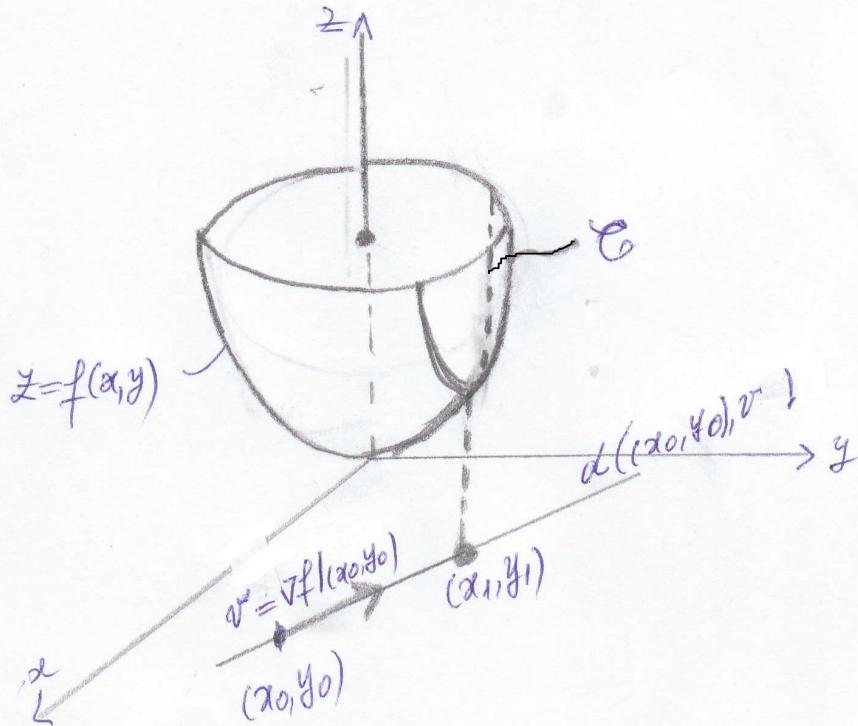
Metoda pasului descendente presupune construirea unui sir (x_k, y_k) , $k=0, 1, \dots$ care să tindă la punctul de minim (x^*, y^*) . La fiecare iteratie pentru calculul noului punct ne vom deplasa pe direcția gradientului, în raport cu care funcția f descrește cel mai rapid.

Fie (x_0, y_0) un punct de pornire. Noul punct

Fie (x_0, y_0) un punct de pornire. Noul punct
 (x_1, y_1) se alege de pe dreapta d , care trece prin (x_0, y_0) și are drept vector director vectorul $\vec{v} = \nabla f|(x_0, y_0)$

$$d: \begin{cases} x = x_0 - \alpha v_1 \\ y = y_0 - \alpha v_2 \end{cases}, \text{ unde } \vec{v} = \nabla f|(x_0, y_0)$$

Fie curba C : $\begin{cases} x = x_0 - \alpha v_1 \\ y = y_0 - \alpha v_2 \\ z = f(x_0 - \alpha v_1, y_0 - \alpha v_2) \end{cases}$



Se caută pe dreapta d punctul (x_1, y_1) corespunzător valoiei minime a lui z pe curba C . Valoarea lui d se obține minimizând funcția $g(d) = f(x_0 - \alpha v_1, y_0 - \alpha v_2)$. Fie acest punct notat cu d_0 . Se rezolvă ecuația $g'(d) = 0$.

$$g(d) = f(x_0 - \alpha v_1, y_0 - \alpha v_2) = \frac{1}{2} \langle A \cdot \begin{pmatrix} x_0 - \alpha v_1 \\ y_0 - \alpha v_2 \end{pmatrix}, \begin{pmatrix} x_0 - \alpha v_1 \\ y_0 - \alpha v_2 \end{pmatrix} \rangle$$

$$- \langle \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \begin{pmatrix} x_0 - \alpha v_1 \\ y_0 - \alpha v_2 \end{pmatrix} \rangle = \frac{1}{2} \langle A(x_0 - \alpha v), (x_0 - \alpha v) \rangle$$

$$- \langle b, (x_0 - \alpha v) \rangle, \text{ unde } x = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Pentru derivarea produsului scalar vom folosi următoarea formulă: $\frac{d}{dt} (\langle a(t), b(t) \rangle) = \langle \frac{d}{dt} a(t), b(t) \rangle + \langle a(t), \frac{d}{dt} b(t) \rangle$

Mai mult, sunt utile și formulele:

$$\bullet \langle A(a+b), c \rangle = \langle Aa, c \rangle + \langle Ab, c \rangle , \quad A \in M_2(\mathbb{R}) \\ a, b, c \in \mathbb{R}^2$$

$$\bullet \langle Aa, b \rangle = \langle a, A^T b \rangle = \underset{A\text{-sim}}{\langle a, Ab \rangle}$$

$$\bullet \langle a, b \rangle = \langle b, a \rangle \quad \bullet \langle da, b \rangle = d \langle a, b \rangle$$

$$g'(\alpha) = \frac{d}{d\alpha} \cdot \left(\frac{1}{2} \langle A(x_0 - \alpha v), (x_0 - \alpha v) \rangle - \langle b, (x_0 - \alpha v) \rangle \right)$$

$$= \frac{1}{2} \left\langle \frac{d}{d\alpha} (A(x_0 - \alpha v)), (x_0 - \alpha v) \right\rangle + \frac{1}{2} \langle A(x_0 - \alpha v), \frac{d}{d\alpha} (x_0 - \alpha v) \rangle$$

$$- \langle b, \frac{d}{d\alpha} (x_0 - \alpha v) \rangle =$$

$$= \frac{1}{2} \langle -Av, x_0 - \alpha v \rangle + \frac{1}{2} \langle A(x_0 - \alpha v), -v \rangle - \langle b, -v \rangle$$

$$= -\frac{1}{2} \langle v, A(x_0 - \alpha v) \rangle - \frac{1}{2} \langle v, A(x_0 - \alpha v) \rangle + \langle b, v \rangle$$

$$= -\langle A(x_0 - \alpha v), v \rangle + \langle b, v \rangle$$

$$= -\langle Ax_0, v \rangle + \alpha \langle Av, v \rangle + \langle b, v \rangle$$

$$g'(\alpha) = 0 \quad (\Rightarrow) \quad -\langle Ax_0, v \rangle + \alpha \langle Av, v \rangle + \langle b, v \rangle = 0$$

$$\Leftrightarrow \alpha_v = \frac{\langle Ax_0 + b, v \rangle}{\langle Av, v \rangle}$$

De unde rezultă coordonatele noului punct și anume:

$$\begin{cases} x_1 = x_0 - \alpha_v v_1 \\ y_1 = y_0 - \alpha_v v_2 \end{cases}$$

Se obține următoarea schemă numerică:

$x_0 \in \mathbb{R}^2$ - ales de regulă în vecinătatea punctului de minim

$$k \geq 0 \quad v = \nabla f / \| \nabla f \|$$

$$\alpha_k = \frac{\langle v, Ax_k - b \rangle}{\langle v, Av \rangle}$$

$$x_{k+1} = x_k - \alpha_k v$$

$$\text{obs. } x_{k+1} = x_k - \alpha_k v \Leftrightarrow$$

$$\begin{cases} x_{k+1} = x_k - \alpha_k v_1 \\ y_{k+1} = y_k - \alpha_k v_2 \end{cases}$$

Algoritmul se oprește în momentul în care $\|\nabla f|_{x_k}\| < \varepsilon$
Obs. se poate demonstra iterativ că $\langle \nabla f|_{x_k}, \nabla f|_{x_{k-1}} \rangle = 0$
cu alte cuvinte, la fiecare iterație nouă direcția a
gradientului este perpendiculară pe direcția de la
iterația anterioară. Putem spune că fiecare iterație
produce o deplasare în zig-zag până se ajunge suficient
de aproape de punctul de minim.